3-D Model Search and Retrieval From Range Images Using Salient Features

Georgios Stavropoulos, Panagiotis Moschonas, Konstantinos Moustakas, *Member, IEEE*, Dimitrios Tzovaras, *Member, IEEE*, and Michael Gerassimos Strintzis, *Fellow, IEEE*

***Abstract*—This paper presents a novel framework for partial matching and retrieval of 3-D models based on a query-by-range-image approach. Initially, salient features are extracted for both the query range image and the 3-D target model. The concept behind the proposed algorithm is that, for a 3-D object and a corresponding query range image, there should be a virtual camera with such intrinsic and extrinsic parameters that would generate an optimum range image, in terms of minimizing an error function that takes into account the salient features of the objects, when compared to other parameter sets or other target 3-D models. In the context of the developed framework, a novel method is also proposed to hierarchically search in the parameter space for the optimum solution. Experimental results illustrate the efficiency of the proposed approach even in the presence of noise or occlusion.**

***Index Terms*—3-D search, partial matching, range image, salient features.**

I. Introduction

OBJECT recognition and matching is a very challenging research area that has been extensively addressed during the last decades. It has numerous application areas, including computer vision, CAD, autonomous navigation, etc. In general, object matching is the process of identifying the correspondence between images, surfaces, points, 3-D models, etc. Especially the problem of 3-D model search and retrieval is a topic that has recently received increasing interest [1].

Partial 3-D matching is a special case of 3-D object matching that involves finding correspondences on parts of 3-D models. One of the most challenging problems in partial matching is to search and retrieve similar 3-D models, when the information describing the query objects is not complete (e.g., a view of an object or a range map is available). The wide availability of range scanners and 3-D digitizers and the emergence of next-generation technologies in 3-D graphics and computational equipment has significantly increased the interest for partial matching algorithms using as input range data. The present paper proposes a "query by range image" 3-D model search algorithm based on a novel partial matching method that utilizes salient points of the range image and the 3-D model.

Manuscript received August 21, 2009; revised January 19, 2010 and May 17, 2010; accepted June 01, 2010. Date of publication June 17, 2010; date of current version October 15, 2010. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Qibin Sun.

The authors are with the Informatics and Telematics Institute, Centre for Research and Technology Hellas, Thessaloniki, Greece (e-mail: moustak@iti.gr).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TMM.2010.2053023

A. Background

The problem of finding correspondences between complete 3-D objects has been successfully addressed by many researchers in the past [2]–[6] while extensive surveys can be found in [7] and [8]. The most prevalent methods are using the query-by-example approach. The major problem of this kind of approaches is to provide translation, rotation, and scaling invariant descriptors. In [2], a method for search and retrieval of 3-D models with the use of the spherical trace transform is described. It is based on tracing the volume of a 3-D model with radial segments and 2-D planes, tangential to concentric spheres. Then using three sets of functional with specific properties, completely rotation invariant descriptor vectors are produced. Elad *et al.* [3] used moments of 3-D objects as the feature vector. The similarity measure is a weighted Euclidean distance between feature vectors.

On the other hand, few approaches have been presented in the past that deal with the problem of recognizing a 3-D object when only a part of its shape is available as query. Some approaches focus on face alignment [9]–[12] for registering two face surfaces. However, the problem of alignment aims to register two surfaces that are *a priori* known to be identical. On contrary, the problem of partial matching and retrieval of similar objects aims to identify similarities and retrieve objects that are in the general case not identical, but exhibit some similarity, to the query partial view. Some full object 3-D matching methods also support partial matching using however as query the full object. In [13]–[17], techniques that use *Reeb graphs* are proposed in order to find similarities between two 3-D objects.

Reeb graphs are topological and skeletal structures that are used as a search key that represents the features of a 3-D shape. In [18], utilizing shock graph matching, indexing using topological signature vectors is applied to implement view-based similarity matching more efficiently.

Biasotti *et al.* [16] compare Reeb graphs obtained by using different quotient functions and highlight how their choice determines the final matching result. For instance, the integral geodesic distance as quotient function is especially suited for articulated objects, while the distance to the barycenter should be preferred when the aim is to distinguish between different poses of an articulated object.

Other commonly used methods for 3-D matching that also support partial matching use *Local features* as described in [19]–[24]. Finally, partial matching also can be achieved with the use of *model graphs* [25]–[27].

In [28], the light field descriptor is presented. The concept behind this method is that if two objects correspond, then they

should also correspond from every viewpoint. A similar approach is presented in [29], the so-called "depth buffer" and in [30] that also includes depth information. However, these approaches [28], [29], [31] do not deal with partial matching.

Germann *et al.* [32] initially precalculate a number of range images from different points of view. Other presented approaches deal with local surfaces applied to the problem of pairwise registration of range images [33] and to 3-D model-based recognition [34], [36]–[38].

Shum *et al.* [22] map the surface curvature of 3-D objects to the unit sphere with the use of a spherical coordinate system. By searching over a spherical rotation space, a distance between two curvature distributions is computed and used as a measure for the similarity of two objects. Unfortunately, this method is limited to objects which contain no holes, i.e., have genus zero.

Most 3-D matching approaches scale linearly with respect to the number of models in the database. Matei *et al.* [35] presented a method based on locality sensitive hashing so as to perform the similarity tests in sublinear time.

The aforementioned approaches have a few drawbacks. In many cases, they use *a priori* information for registering the partial view with the complete 3-D model, i.e., in the case of face matching, the tip of the nose or the center of the eyes can easily be registered. In the proposed method, there is no need of preregistering the views, thus allowing matching between an arbitrary range image and a 3-D model. Methods that rely on *local feature similarity* either cannot be applied on objects that contain holes, or require nontrivial preprocessing for the meshes. On the other hand, *Reeb graph methods* are very sensitive to topological changes and cannot be applied to arbitrary meshes, since topological problems, like missing faces, disturb the computation of the graph.

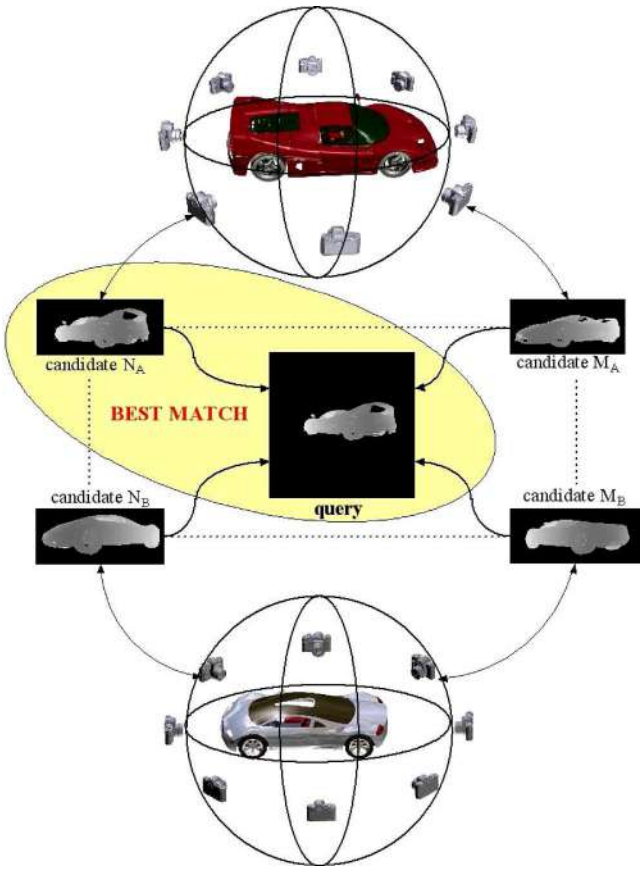


Fig. 1. Proposed framework: Assuming a query range image, the algorithm searches for the best match in parameter space that consists of all possible positions and orientations of the camera. Since this is a computationally very expensive procedure, a hierarchical approach is designed to search for the best match.

B. Proposed Method

In this paper, a novel method for identifying the correspondence between a range image and a full 3-D model is presented.

Similarities between the query image and a model in the database is performed by searching for the camera parameters (camera viewpoint, orientation, scale, and internal geometry) that would generate an image similar to that of the query image as illustrated in Fig. 1. Rather than match the whole image, only the salient points are compared. The matching relies on the fact that the salient points in the correctly matching image should have a similar spatial distribution to those in the query image. Instead of exhaustively searching all possible camera parameters, the search is conducted in a hierarchical, coarse-to-fine manner by first partitioning the parameter search space coarsely, and then pursuing the best matching region of the parameter space at progressively finer levels of quantization.

The proposed scheme is very efficient in terms of recognition rate, and while using the hierarchical approach, the computational complexity is reduced by several orders of magnitude.

The major steps of the proposed method are three: the *3-D model preprocessing*, the *2.5-D geometry processing*, and the *Hierarchical matching*. The system uses as input a range image (I), and a database of 3-D models $M_i. \quad i = 1 \ldots . N$.

At the preprocessing step, a set of features is extracted for each model M_i, the set of *salient points* (S_i). During run-time, a 3-D mesh is initially extracted from the input range image, and then the set of salient points for the extracted mesh is identified. This is achieved by selecting a set of feature points on the range image and creating a triangulated 3-D mesh using Delaunay triangulation and the depth information contained in the range image. Then the salient features of the 3-D mesh are extracted.

The final step is the matching procedure. Matching is performed in a hierarchical manner (Section IV-C), by selecting a set of camera parameters (Section II), and calculating the value of an error function between the salient features extracted from the query range image and the ones extracted from the 3-D model (Section IV-B). The hierarchical matching algorithm proceeds by searching for the minimum error set of parameters $(\mathbf{q}_l)$ at the "l" level of the hierarchy, in the neighborhood of the minimum error set of parameters $(\mathbf{q}_{l-1})$ of the upper level of the hierarchy.

The main contributions of the proposed framework are the representation of the objects using their salient features, that is compact and can be efficiently used for realistic scenarios like the presence of occlusions; the hierarchical approach to reduce the computational cost of searching in the feature space and the

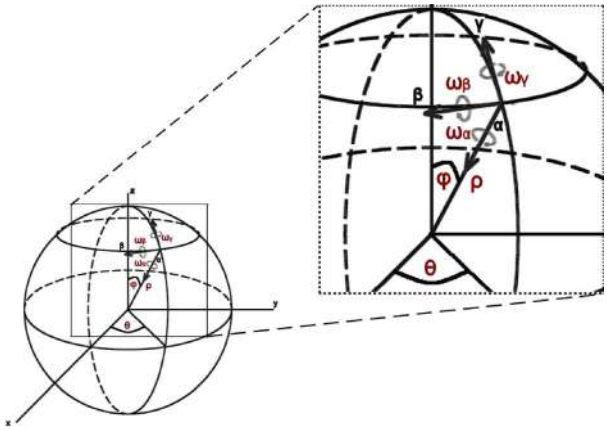


Fig. 2. Camera parameters of the proposed framework. Perspective projection is assumed.

general concept of the virtual camera that, when having the correct parameters, should produce a similar range image to the query that also exhibits a similar salient feature spatial distribution. Moreover, it should be emphasized that no information about the setup and the parameters of the camera that captured the query is *a priori* known. Thus, the proposed scheme goes one step further from existing methods by utilizing range images and salient features to perform 3-D search from partial queries. The proposed framework does not aim to identify the optimal alignment of objects known to be similar but to search and retrieve objects that exhibit similarity to the query image but are not identical in the general case. The framework is experimentally seen to provide very good retrieval performance even in the presence of noise or occlusions as described in the experimental results in Section V.

II. Camera Model

Perspective projection using the standard pinhole camera model is adopted in the proposed method. The intrinsic parameters of the camera are its resolution [r_w (width), r_h (height)] and the focal length (f). The position of the camera is described using a spherical coordinate system with the following parameters: radius (ρ), longitude (ϕ), and latitude (θ).

In the model used in the context of the proposed framework, the camera looks at the center of mass of the object that is placed at the origin of the coordinate system. Vector $\mathbf{cv}$ (Fig. 2) corresponds to the direction in which the camera looks. In the proposed framework, the camera typically looks at the origin of the coordinate system. Parameters roll (ω_α) yaw and pitch $(\omega_\beta, \omega_\gamma)$ are used to refine where the camera looks. Assuming the local coordinate system of the camera (α, β, γ), the angles ω_α, ω_β, and ω_γ correspond to rotation around the α-, β-, and γ-axis, respectively, as illustrated in Fig. 2.

Some of the camera parameters can be estimated prior to the matching process, in order to reduce the dimensionality of the parameter space and thus save computational power. The resolution of the camera is set equal to the resolution of the input image. The focal length can be explicitly estimated as described in the following paragraph. Thus, the parameters that need to be estimated are reduced to the following six $(\rho, \phi, \theta, \omega_\alpha, \omega_\beta, \omega_\gamma)$.

Focal Length Estimation: The focal length of the camera, which is unknown since there is no *a priori* information about the camera that captured the query, can be directly estimated using the range image. Assuming that the virtual camera is placed correctly, the object on the produced range image should have equal dimensions with the query range image. So, instead of including one more dimension for f in the parameter space, the focal length is estimated utilizing the 2-D bounding box of the input range image and the 2-D bounding box of the projection of the 3-D model.

This is performed by projecting the 3-D model onto the camera plane using a reference focal length f, and then comparing the axis aligned bounding box (AABB) of the projected 3-D model $(\Lambda\Lambda\mathrm{BB}(\widehat{M}))$ with the AABB of the query range image $(\Lambda\Lambda\mathrm{BB}(I))$. The estimate of the focal length $\hat{f}$ stems from (1):

$$\hat{f} = f \cdot \frac{\Lambda\Lambda\mathrm{BB}(\widehat{M})_w}{\Lambda\Lambda\mathrm{BB}(I)_w}. \tag{1}$$

III. Feature Extraction

As will be discussed in a following section, the identification of similar areas is computationally very expensive if matching is performed by comparing the query range image with another range image extracted from the 3-D model. Therefore, a subset of features for the range image and the 3-D model should eventually be used that should be easy to handle and representative for each object. In the present paper, salient features are used that lie in general in the most protruding areas of a 3-D surface.

A. Salient Feature Extraction

The developed method for salient feature extraction that corresponds to sharp protruding areas of the object's surface [39], [40] is based on Hoffman and Singh's theory of salience [41]. In order to make this paper self-contained, a brief description of the method follows. Initially, the dual graph $G = (V, E)$ of the given triangulated surface is generated [39], where V and E are the dual vertices and edges. A dual vertex is the center of mass of a triangle, and a dual edge links two adjacent triangles. The degree of protrusion for each dual vertex results from the following equation:

$$p(\mathbf{u}) = \sum_{i=1}^{N} g(\mathbf{u}, \mathbf{v}_i) \cdot \mathrm{area}(\mathbf{v}_i) \tag{2}$$

where N is the number of dual vertices in the entire surface, $p(\mathbf{u})$ is the protrusion degree for the dual vertex $\mathbf{u}$. $g(\mathbf{u}, \mathbf{v}_i)$ is the geodesic distance of $\mathbf{u}$ from dual vertex $\mathbf{v}_i$, and $\mathrm{area}(\mathbf{v}_i)$ is the area of the triangle $\mathbf{v}_i$.

Using simple gradient-based methods (i.e., steepest descent), all local maxima of the protrusion map $p(u)$ are obtained. Geodesic windows are then applied and only the global maxima inside the window are considered as salient. A

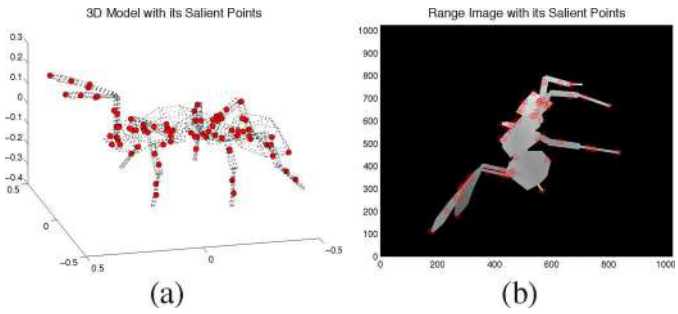


Fig. 3. Salient features of an Ant model extracted using its (a) entire 3-D model and (b) 2-D projected range image.

geodesic window (GW) centered at the dual vertex u is defined as follows:

$$\mathrm{GW}_{\mathbf{u}} = \{\mathbf{v} | \forall \mathbf{v} \in V. g(\mathbf{u}. \mathbf{v}) < c\} \tag{3}$$

where c defines the window size.

B. 3-D Model Preprocessing

The 3-D models of most databases are in general in various scales. In order to be able to easily compare the similarities between a query range image and a set of models, the models should be normalized to a common scale. As mentioned earlier, the spherical coordinate system is used, so the best choice would be to normalize each model to the *unit sphere* and extract the *salient points* of the 3-D models using the technique described in the previous section.

It should be noticed that the database models are normalized to a common scale so as to ease the implementation of the matching algorithm. The normalization is not a crucial process. Database objects could also stay at their original scale; in this case, the estimated camera parameters, more precisely the radius "ρ" of the position of the camera, would be different so as to compensate for the difference in scale.

C. Range Image

In order to extract salient features from the range image, a 3-D surface should initially be formed. The surface is created using only a subset of the points of the image, so as to reduce the redundancy and size of the triangulated surface to be generated. Features on the range image are selected, as the ones with maximum minimal eigenvalue in a predefined window [42]. After the 3-D surface is formed, the salient features are extracted in the same way that they are extracted for the 3-D models. Fig. 3 illustrates the salient features extracted from the 3-D model of an Ant and from its 2-D projected range image.

IV. Matching

As mentioned in the overview (Section I-B), the basis of the proposed framework is a virtual camera, assumed to lie in the space of the examined 3-D model. The proposed method searches in the parameter space for the set of camera parameters that capture a surface as similar as possible (identical in the ideal case) to the surface of the query range image.

The above is encapsulated in the following hypothesis.

Hypothesis: Assuming that the 3-D model M and the partial surface that is described through the range image I do correspond, then

$$\exists C_{P_i,P_c} \quad \text{so} \quad \text{that} \quad C_{P_i,P_c}(M) = I \tag{4}$$

where C_{P_i,P_c} is a range sensor (camera) with specific intrinsic (P_i) and extrinsic (P_c) parameters.

Consider the query range image I of an object M. If the object $\hat{M}$ (identical to M) is captured using a virtual camera $\hat{C}$ that captures a range image $\hat{I}$ using the camera parameters $\hat{P}_i. \hat{P}_c$, then $\hat{I} = \hat{C}_{\hat{P}_i,\hat{P}_c}(\hat{M})$. Under the assumption that the objects M and $\hat{M}$ are identical and if the correct camera parameters $\hat{P}_i. \hat{P}_c$ are estimated (i.e., $\hat{P}_i = P_i. \hat{P}_c = P_c$), then trivially $\hat{I} = I$.

For the non-ideal case of non-identical objects $(\hat{M} \approx M)$, in the presence of noise or occlusions $(I \approx C_{P_i,P_c}(M))$, or with an approximate only estimation of the camera parameters $(\hat{P}_i \approx P_i. \hat{P}_c \approx P_c)$, the images $\hat{I}$ and I cannot be identical.

Utilizing the above hypothesis, the problem of identifying the correspondence between a range image and a 3-D model is reduced in finding the correct camera parameters, P_i and P_c, that minimize the error function:

$$\{P_i. P_c\} = \arg_{\min}\{\mathcal{E}(P_i. P_c)\} \tag{5}$$

where

$$\mathcal{E}(P_i. P_c) = f(C_{P_i,P_c}(\hat{M}). I). \tag{6}$$

A. Matching Using the Range Image

The simplest way to match a given range image with a 3-D model utilizing the hypothesis of Section IV is to create a range image for every camera in space, for every possible set of parameters, and compare the captured range image with the query range image. Although this method would be in the ideal case perfectly accurate, it is computationally unacceptable, while it is prone to errors in real use cases (e.g., in the presence of occlusion or noise).

Search in the parameter space is performed by using *a spherical coordinate system* that describes the camera position. The spherical coordinate system provides a trivial way to ensure that the virtual camera will "look" at the model that is placed in the center of the coordinate system.

Then, the range image of a particular view is generated using graphics hardware, in particular OpenGL functions.

After the range image is created, it is compared with the query range image by calculating the error for the current camera parameters using the following equation:

$$\mathcal{E}_{\mathcal{R}}(P_i. P_c) = \frac{1}{N_\Omega} \sum_{\forall \{i,j\} \in \Omega} (I(i. j) - \hat{C}_{P_i,P_c}(\hat{M})(i. j))^2 \tag{7}$$

where Ω is the support set of the pixels of the range image that have depth value both in the query and the generated range image and N_Ω the size of Ω. The set of camera parameters

$\{P_i.P_c\}$ that minimizes (7) is considered as the best matching set of parameters.

The range image error function includes all the pixels that belong both to the query and the database object's image, only if the size of Ω is larger than a predefined threshold that is set to be 0.65 with respect to the size of the query object. Thus, if less than the 65% of the query object match with the target, then this set of camera parameters is considered as nonmatching set.

It should be noted that Ω can also be defined as the union of the query and the generated range images that would result in a further penalty for surfaces with nonmatching contours. This degrades however the performance of the algorithm in the presence of occlusions, which is a real-world scenario and is extensively researched in the experimental results section.

B. Matching Using Salient Features

A way to accelerate the matching process is to use only a number of feature points of the range image and the 3-D model. In the context of the proposed method, salient points are used, as described in Section I-B.

If the virtual camera described in the hypothesis of Section IV exists and the query range image corresponds to the 3-D model, then if a set of salient points is extracted from the query range image and another one from the surface that the virtual camera captures, then these two sets of points should also have corresponding subsets of features.

The proposed method proceeds similarly to the exhaustive matching approach (Section IV-A). A set of salient points (S^I) is extracted for the query range image (as described in Section III-C), and another one (S^M) for the 3-D model. The latter is transformed for each set of camera parameters $(P_i.P_c)$, in order to be comparable with the set of salient points (S^I), thus producing the set of salient points $S^M(P_i.P_c)$. These sets of points are used to calculate the similarity between the query range image and the 3-D model.

More specifically, the distance function $\mathcal{E}_S$ used to calculate the difference between the two sets of salient points takes into account the 3-D distance between the elements of the two sets of salient points and stems from (8):

$$\mathcal{E}_S(P_i.P_c) = \frac{1}{N_s}\sum_{k=0}^{N_s}\left(\min_j\left(\text{dist}\left(S_k^I.S_j^M\right)\right)\right). \quad \forall j \qquad (8)$$

where N_s is the number of the salient points of the range image and $\text{dist}(A.B)$ is the 3-D distance between points A and B. Its value has been experimentally selected to be $N_s = 35$. As seen over the performed experiments, the further increase of this parameter does not lead to more efficient retrieval performance. The error function of (8) actually computes the sum of the distances of the range image salient points from their closest salient points of the processed 3-D model.

As in the case of exhaustive matching, the set of camera parameters that minimizes the error function $\mathcal{E}_S$ is considered as the best match:

$$\{P_i.P_c\} = \arg_{\min}\{\mathcal{E}_S(P_i.P_c)\}. \qquad (9)$$

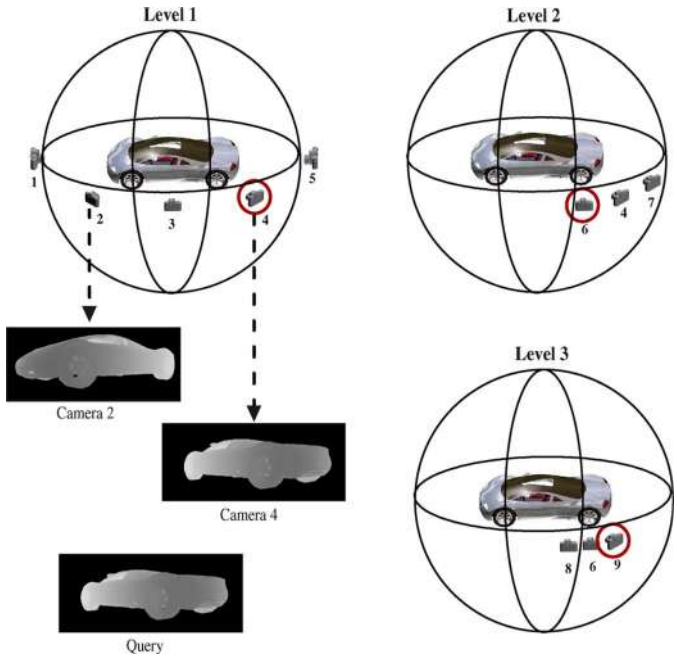


Fig. 4. Hierarchical search in the parameter space. The procedure is illustrated only for parameter θ.

C. Hierarchical Search

Even if the use of salient points reduces the computational complexity of matching full range images, the matching procedure remains computationally expensive when exhaustively searching in the parameter space. In the proposed framework, a hierarchical approach is developed that reduces the computational complexity by several orders of magnitude.

Consider the parameter space described in Section II. The range of all the angles $\theta.\phi.\omega_\alpha.\omega_\beta.\omega_\gamma$ is sampled using a 0.2° step, while for the camera position ρ, the step is set to 0.1. Notice also that: $\theta.\omega_\alpha \in [-\pi.\pi]$. $\phi \in [-\pi/2.\pi/2]$, and $\omega_\beta.\omega_\gamma \in [-\pi/9.\pi/9]$, while $\rho \in [1.5]$.

All the sampling densities and the maximum-minimum values for $\omega_\beta.\omega_\gamma.\rho$ are experimentally selected for the ITI 3-D model database [1] and are valid for every 3-D model database.

The hierarchical search algorithm builds initially a coarse sampling of the 6-D parameter space and evaluates for each sample the error function as described in Section IV. The sample that produces the minimum error is considered as best match, and then the algorithm proceeds to the second layer of the hierarchy. In this layer, the error function is evaluated around the local neighborhood of the "winning" sample, and the new sample that produces the minimum error is considered as best match for the second layer of the hierarchy. This procedure is repeated until the final layer of the hierarchy is reached that corresponds to the maximum accuracy as previously described.

At this point, it should be mentioned that the sampling of the first layer of the hierarchy is of high importance since a very coarse sampling would possibly result in missing a corresponding view, while dense sampling would inhibit the performance of the algorithm as also seen with experiments in the ITI database [1]. In the context of the proposed framework, the initial sampling for angles $\theta.\phi$, and ω_α has been chosen to be

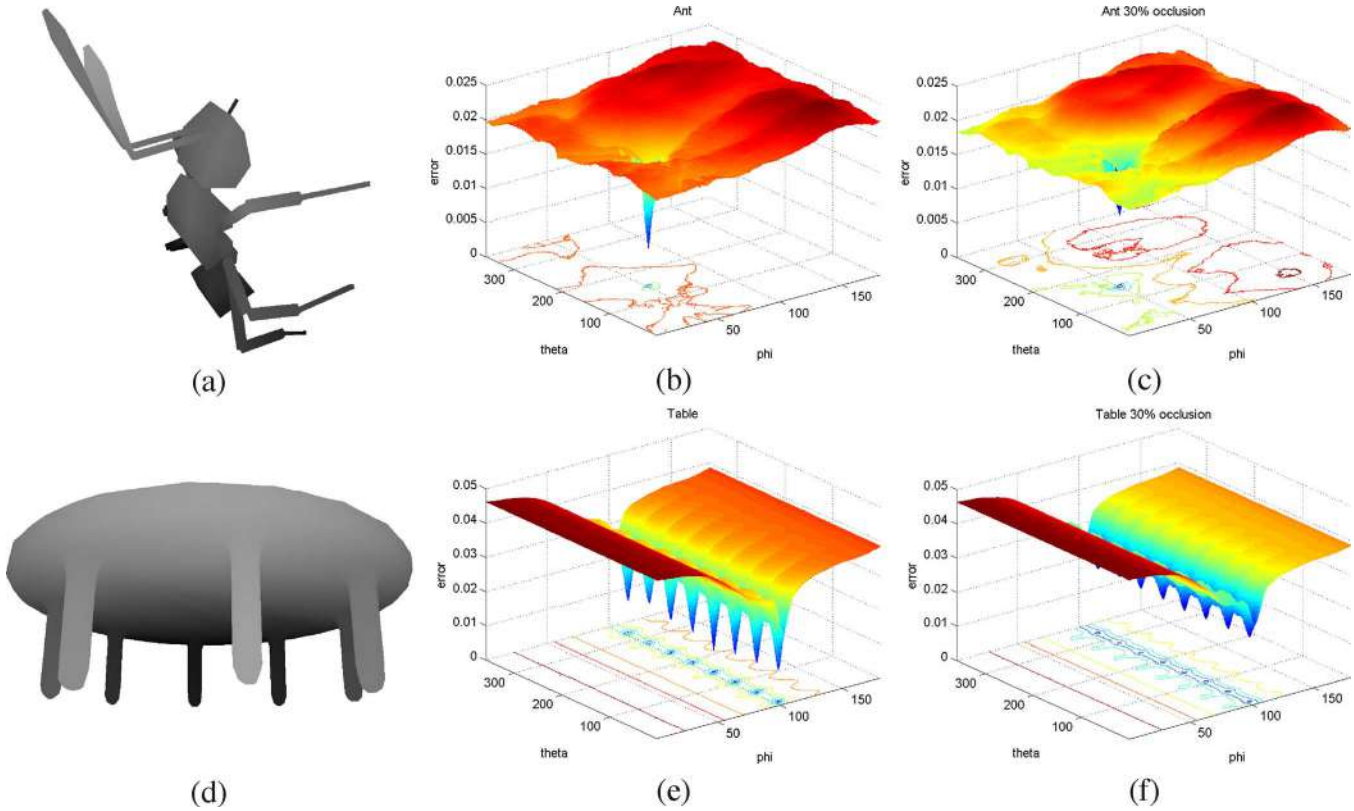


Fig. 5. (a) Ant model. (b) Distribution of the error function for the Ant and for different values of θ and ϕ. (c) Distribution of the error function for the Ant for 30% occlusion. (d) Table model. (e) Distribution of the error function for the Table. (f) Distribution of the error function for the Table for 30% occlusion.

$\pi/6$, while for ρ, ω_3, and ω_4, their sampling includes the three values of the set $\{\min, (\max - \min)/2, \max\}$.

For all other layers of the hierarchy, all parameters are handled in an equal manner, thus resulting in the computation of the error function for two more samples for each parameter. More specifically, the error function is evaluated for the samples of the set $\{w - \delta/2, w, w + \delta/2\}$, where w and δ are the minimum error sample and the sampling step of the previous layer, respectively.

Fig. 4 illustrates an indicative diagram that illustrates the concept behind the hierarchical approach. Given a query range image, initially, in the first level of the hierarchy, range images and their salient features are extracted and the error function is evaluated for all cameras of the initial setup. In the second level, local search is performed in the area near the camera that produced the lowest error (camera 4). This procedure is repeated iteratively until the final level is reached that equals to an accuracy of 0.2° for all the angles, and 0.1 for r.

D. Convergence Analysis

An important issue of the proposed method, and of hierarchical methods in general, is to avoid the convergence of the algorithm in a local minimum of the function to be minimized. In the context of the proposed framework, this issue has been carefully addressed and the parameter space is sampled in such a way so as to minimize the possibility of convergence in a local minimum.

Fig. 5 illustrates four diagrams that concern the Ant [Fig. 5(a)] and the Table [Fig. 5(d)] model. In Fig. 5(b), the distribution of the error function is depicted for all possible different values of θ and ϕ, while Fig. 5(c) illustrates the same distribution in the presence however of 30% occlusion. Notice that the global minimum can be clearly identified, while other local minima have much higher value. Similar characteristics can be observed in Fig. 5(c). This is an indication for robustness in the presence of occlusion as also will be discussed later in the experimental results section.

Moreover, the valley of the global minimum extends to more than 50° for both the ϕ and θ parameters. This behavior is observed for all examined objects of the database and for the remaining parameters of the parameter space. Therefore, in the proposed algorithm, the initial sampling of the parameter space for the first level of the hierarchy is chosen to be as described in Section IV-C. Coarser sampling is seen to inhibit the performance of the algorithm; more precisely using initial sampling of $\pi/4$ for the angles θ and ϕ reduces the "Rank 1" retrieval performance of the algorithm by 15%. Denser sampling is not seen to significantly increase the accuracy of the results, while slowing down the process; more precisely using initial sampling of $\pi/12$ for the angles θ and ϕ is seen to increase the "Rank 1" retrieval performance of the algorithm by 1%.

The diagrams of Fig. 5(e) and (f) illustrate the error distribution for the symmetrical table object. Notice that in this case, there exist several local minima that are however very close in value to the global minimum. As can be seen in Fig. 5(e) and (f), their only major difference is that the minimum error in the

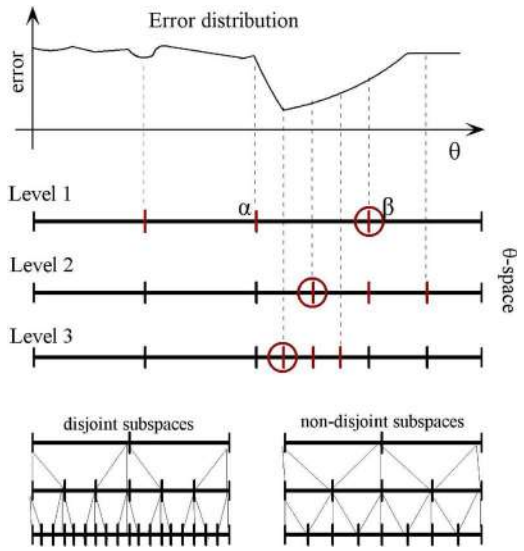


Fig. 6. Recovery of small errors of the proposed hierarchical approach. On the top of the figure, an indicative error distribution function is illustrated and just below, three levels of the hierarchical approach. The minimum error can be reached by either selecting α or β in Level 1, since the proposed approach does subdivide the parameter space into nondisjoint subspaces (bottom-right). In case of disjoint space subdivision (bottom-left), the minimum error could not be reached by selecting β in Level 1.

valleys is increased due to the missing occluded part. A more thorough discussion on symmetric models is presented in a following section.

Moreover, the hierarchical algorithm is designed so as to be able to recover for small errors in the estimate of the parameters. Consider, without loss of generality, three levels of the hierarchical algorithm for a single parameter as illustrated in Fig. 6.

The proposed hierarchical algorithm does not divide the parameter space into disjoint regions. Thus, any value between α and β in Fig. 6 can be reached by selecting in the first level as minimum error sample either α or β. Typical hierarchical algorithms that divide the parameter space into disjoint subspaces that are asymptotically adjoined (Fig. 6) would, contrary to the proposed approach, not be able to handle error distributions like the one illustrated in Fig. 6, or would require a much denser sampling of the parameter space in the initial level.

V. Experimental Results

The proposed method was tested on the 3-D model database of the *Watertight model Track of Shape Retrieval Contest'07* (SHREC) [43] and the Princeton Shape Benchmark [44].

The SHREC database consists of 400 models organized in 20 categories. From the database, 400 range images where created from different views (one for each model) using random angle parameters, and entered as query in the matching algorithm. Each range image was compared with all the 3-D models.

The Princeton Shape Benchmark database consists of 1814 3-D models. The applied classification was based on the proposed "Coarse1" categorization. A total of 49 classes were used. A range image was created from every model using random angle parameters. As before, each produced range image was compared to all the 3-D models. The resolution of the synthesized range images was 400×400 for both benchmark databases.

A. Evaluation on Benchmarking Databases

To obtain comparable results, the approach of Germann *et al.* [32] is used since it is one of the very few approaches for partial matching using as query partial object views, although utilizing a very different method when compared to the proposed approach and making assumptions that are not made in the proposed approach like known object scale.

The evaluation is performed by computing the ranking during retrieval and using precision-recall diagrams, where precision is defined as the ratio of the relevant retrieved elements against the total number of the retrieved elements, and recall is the ratio of the relevant retrieved elements against the total relevant elements in the database.

Fig. 7 illustrates the precision versus recall (P-R) curves for the proposed hierarchical matching approach using only salient features and using the full range image for matching compared to [32], for both the Princeton Shape Benchmark [Fig. 7(a)] and the SHREC database [Fig. 7(b)].

Moreover, Fig. 8 illustrates the accuracy of the aforementioned methods in terms of their ranking efficiency. The proposed method using salient features outperforms the approach of [32] and is also clearly superior to the same approach using the full range image for matching. This is expected when considering that the salient features comprise a robust set of attributes of the 3-D object that describes its characteristics based on the theory of visual saliency [41].

Additionally, this set of features is also robust in the presence of occlusion and noise. Fig. 9(a)–(c) indicatively depicts a clear range image query, an occluded query, and an occluded query with additive Gaussian noise, respectively. The occluded queries are obtained by eliminating all connected local surfaces (triangles) that are closest to a salient point, without splitting the object into nonconnected parts. The noise queries are generated by adding Gaussian noise on the range image. As illustrated in the diagrams of Fig. 10, the performance of the proposed method does not drop significantly even for 20% or 30% of occlusion, contrary to the approach of [32]. The robustness of the proposed algorithm in the presence of occlusion is expected since the "descriptor" of the object using salient features is not altered for an occluded object, but actually only reduced in terms of its size, i.e., the salient features that correspond to the occluded area are not included in the "descriptor".

Fig. 11(a) and (b) demonstrates the robustness of the proposed framework in the presence of occlusion and additive Gaussian noise in the query range image. The performance of the proposed method remains high, while the performance of the approach in [32] drops significantly.

Fig. 12 illustrates the first five retrieved results for four queries that are depicted in the first column of the image. The results of the three first queries are all from the same class, and the highly ranked retrieved model is the model from which the query range image is originated. For the last query, the fourth retrieved result is from a different class. This is consistent

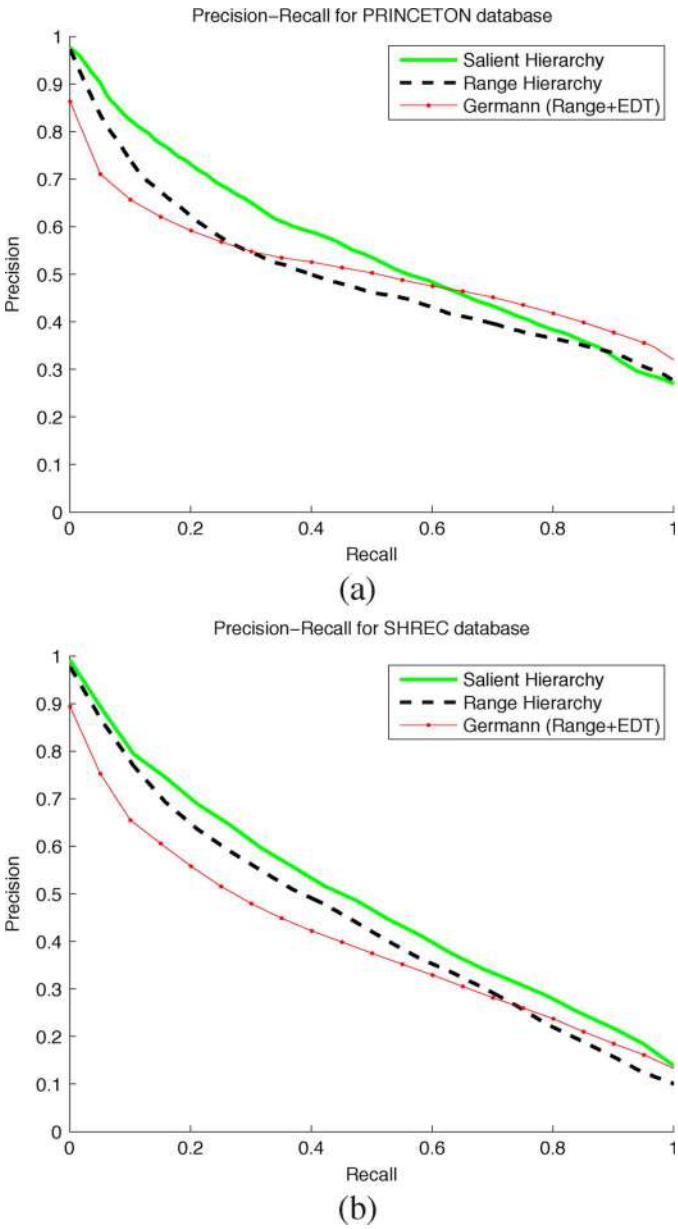


Fig. 7. Precision-recall curves for the proposed hierarchical matching approach using salient features and the entire range image compared with the approach of [32] for the (a) Princeton Shape Benchmark and (b) SHREC database.

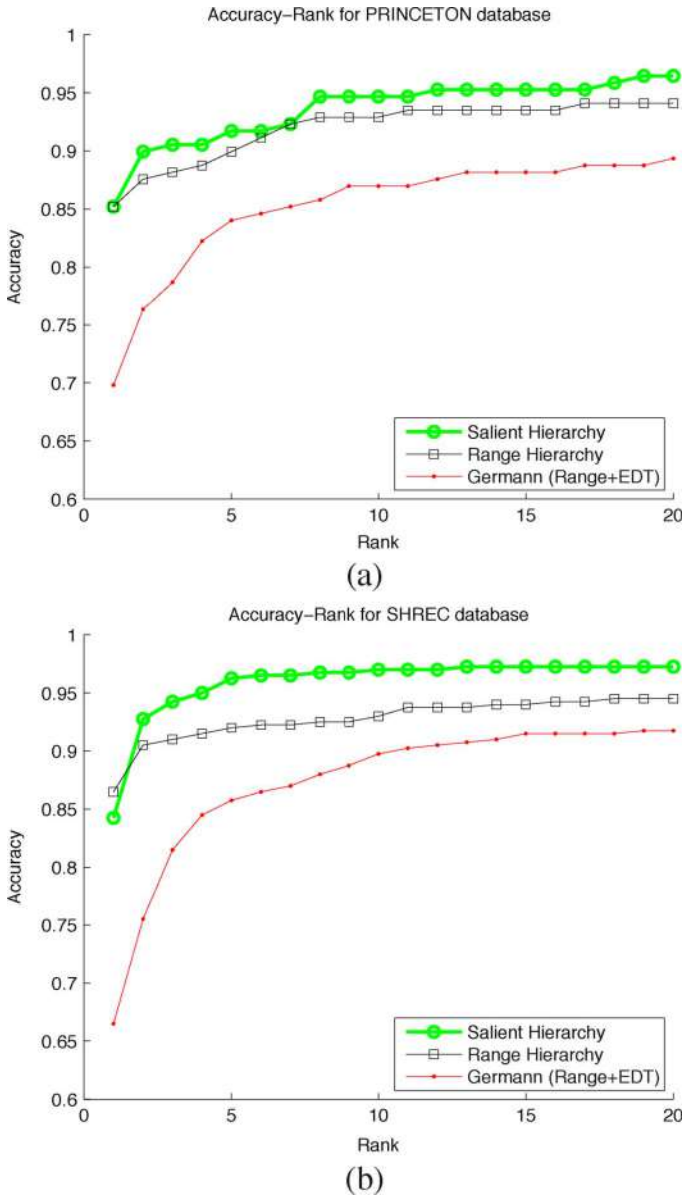


Fig. 8. Ranking curves for the proposed hierarchical matching approach using salient features and the entire range image compared with the approach of [32] for the (a) Princeton Shape Benchmark and (b) SHREC database.

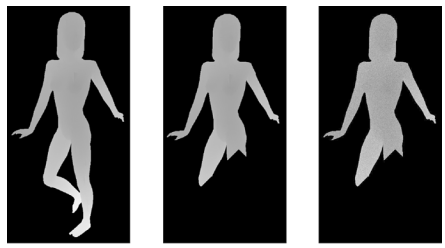


Fig. 9. Indicative range image queries. (a) Clear query. (b) Occlusion query. (c) Occlusion and noise query.

with the proposed method, and it will be further discussed in Section V-C2.

Moreover, Fig. 13 illustrates comparative "precision versus recall" diagrams of the proposed approach compared to the approach of [30], which is a representative of the view-based similarity approaches, in the PSB and SHREC databases. As can be clearly seen, the proposed approach outperforms the method in [30]. It has to be mentioned that the authors in [30] use a view-based matching approach that also utilizes depth information to perform mainly full 3-D model search. Their algorithm however also can be utilized for partial matching as described in [30]. The advantage of the proposed scheme over this approach, and existing view-based matching schemes in general, can be considered from the one side as the representation of objects using salient features that is robust to partial queries and occlusions and from the other side the efficient hierarchical matching scheme that eliminates restrictions of using only a few views to identify similarities, which is the case in [30].

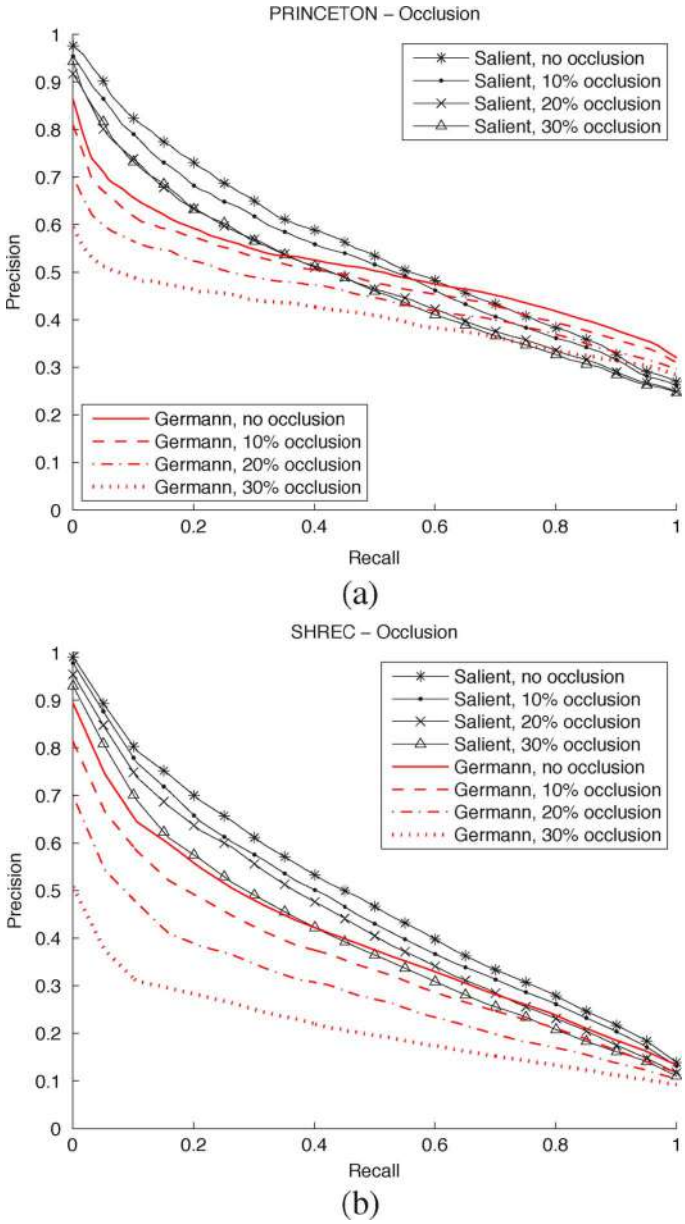


Fig. 10. Precision-recall curves for the proposed approach in the presence of occlusion compared to the approach in [32] for the (a) Princeton Shape Benchmark and (b) SHREC database.

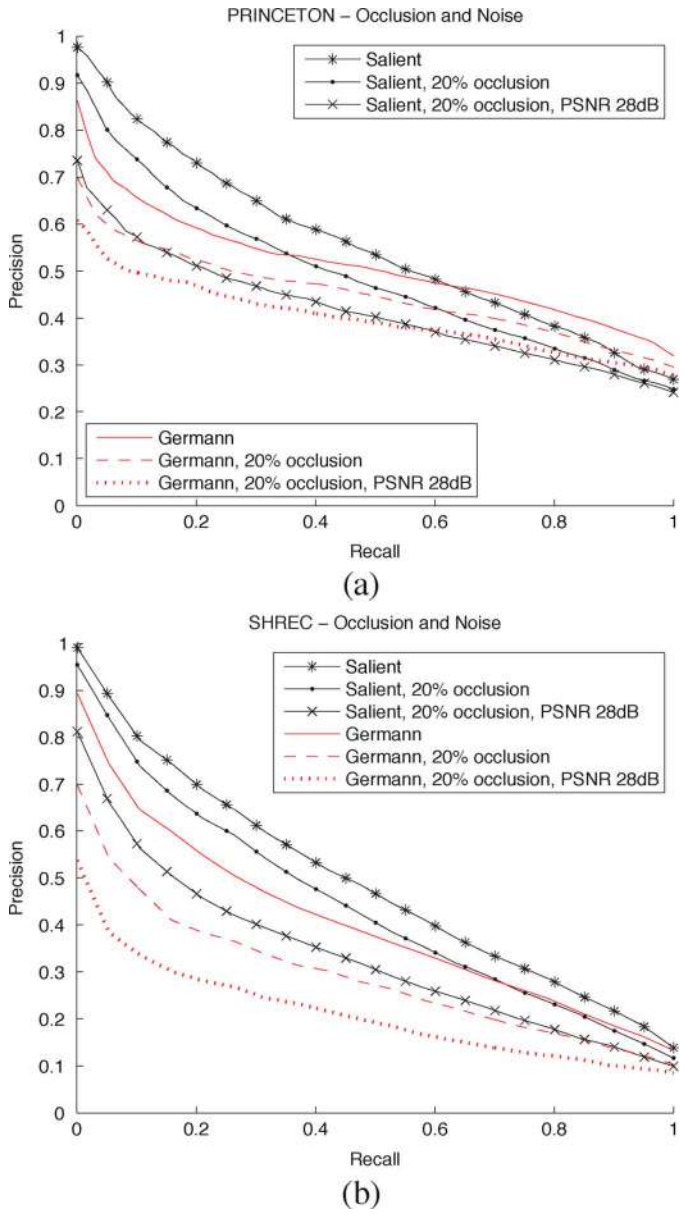


Fig. 11. Precision-recall curves for the proposed approach in the presence of occlusion and noise compared to the approach in [32] for the (a) Princeton Shape Benchmark and (b) SHREC database.

B. Evaluation on Real Range Data

The proposed method also supports matching by using as query a reconstructed 2.5-D image. These reconstructions can stem either from a 3-D scanner or using at least two images taken from different viewpoints so as to extract the relative position of the image elements [42] and to create a range image. Then, this range image is inserted as query in the matching algorithm.

To further validate the proposed approach with real queries, a range image database has been assembled, stemming from scans performed either using the MINOLTA VIVID700 3-D scanner and depth maps captured from a stereoscopic camera [45]. The database consists of the following objects: ten humans, six cups, five pliers, six mechanic objects, six bearing objects, and 17 vases. All the above categories also exist in the SHREC database. The resolution of the range images for this case is "800×800". Fig. 14 illustrates indicative images of the performed queries.

Even if the generation of these queries using the 3-D scanner and the accompanying postprocessing software is straightforward, the view-based reconstruction is achieved by the methods described in [42] that utilize two (or more) images taken from different viewpoints in order to create a depth map (range image) of the central object on the images. At first, the relative pose of the images is estimated by selecting a set of features on each image and registering them. Then with the use of the eight-point algorithm [42], the relative pose is estimated along with the depth map for one of the images.

Fig. 15 illustrates comparative precision versus recall diagrams of the proposed approach compared to the approach of

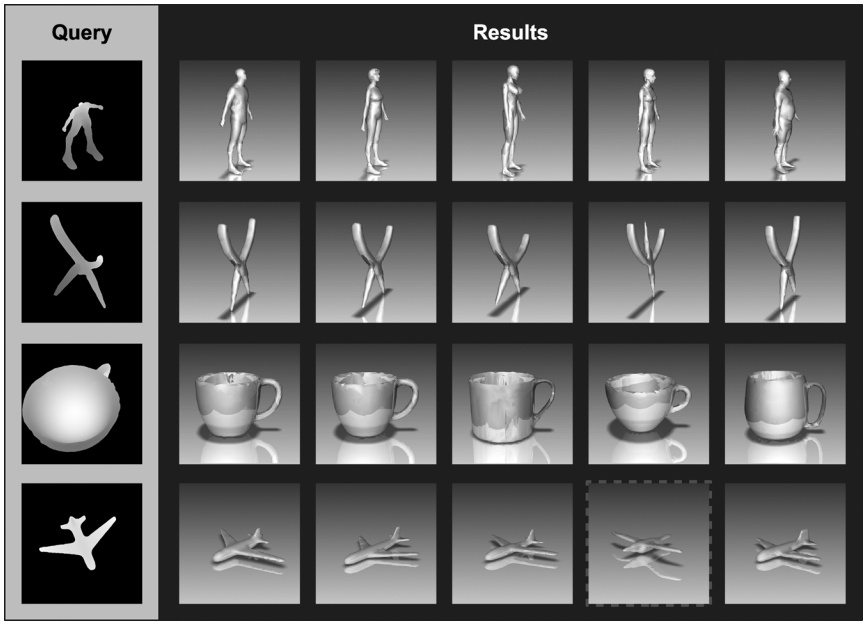


Fig. 12. Results overview: the five first retrieved models for three queries.

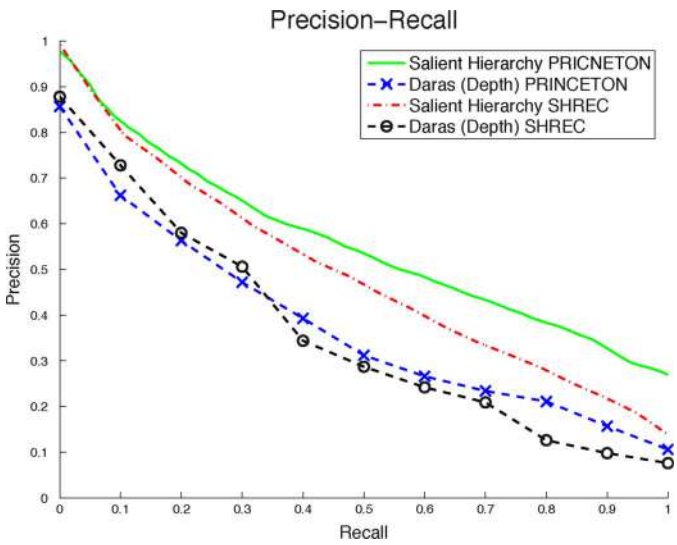
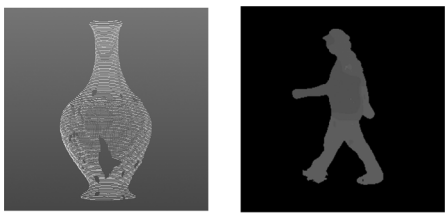


Fig. 13. Precision-recall curves for the proposed approach compared to [30] for the PSB and SHREC databases.



(a) (b)

Fig. 14. Indicative real range queries. (a) Vase captured using the MINOLTA VIVID700 3-D scanner. (b) Human captured using stereoscopic camera.

[32]. As can be clearly seen, the proposed scheme remains superior, while its performance does not drop significantly, when compared to the results of synthetic queries.

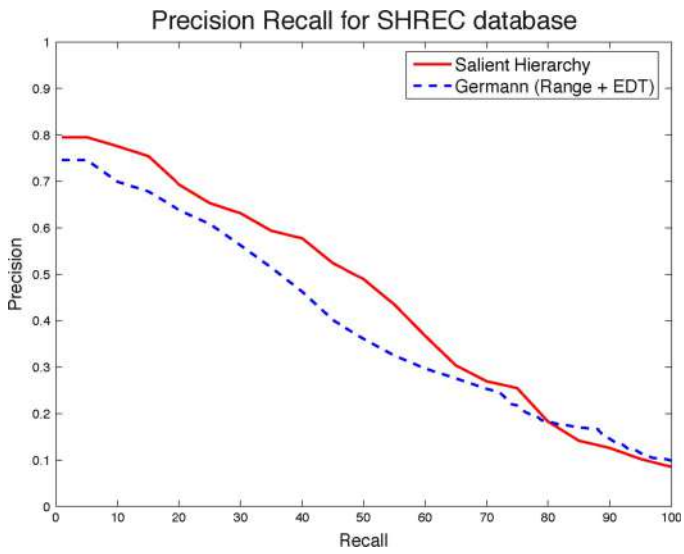


Fig. 15. Precision-recall curves based on real range image queries for the proposed approach compared to the approach in [32] for the SHREC database.

In Fig. 16, a set of results is presented by using as query a reconstructed range image. The last row corresponds to images taken with a stereoscopic camera, while the other two are from synthetic images. The matching results for the synthetic test images are very good due to the great accuracy of the reconstruction. On the other hand, the last set of test images produced less accurate results. This was due to the fact that the reconstructed surface was not very accurate (detail in lower-left part of Fig. 16).

C. Performance Analysis

1) Model Category: Fig. 17 illustrates the P-R curves for four different classes of the SHREC database. The matching performance for the class of chairs and sea animals is excellent due to

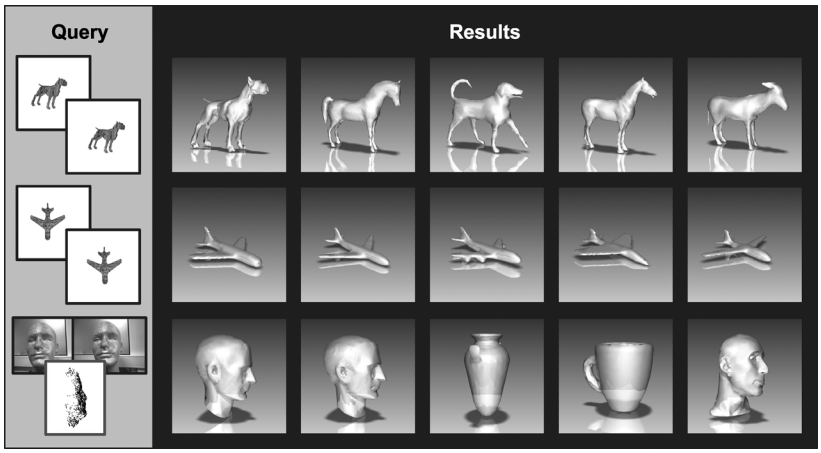


Fig. 16. 3-D reconstruction results. The range image reconstructed from the two images on the left column was entered as query in each case.

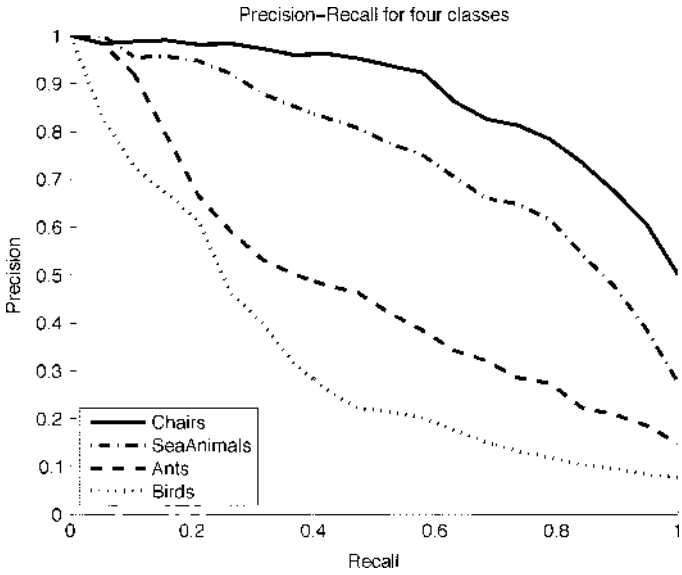


Fig. 17. Precision-recall diagram for four different classes. The variation in performance proportionately to the class of the 3-D model is shown.

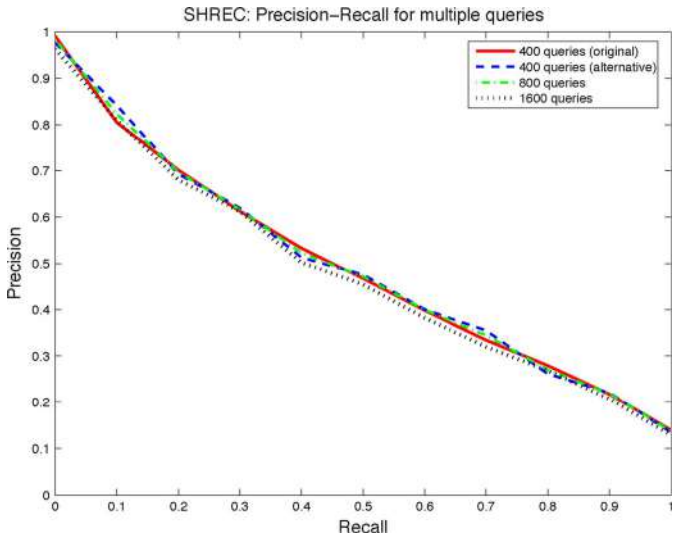


Fig. 18. Precision-recall diagrams for 400, 800, and 1600 range image queries for the SHREC database.

their unique shape that differentiates them from the other categories. For the category of the ants, the algorithm performs less accurately, while the less accurate performance is observed for the category of birds. The reason for the not very efficient performance of some categories is that the shape of their member models is similar to other models, even if they belong in other categories. For example, the birds are very often miscategorized as airplanes due to their very similar shape.

Moreover, it should be emphasized that the proposed approach is theoretically valid in the cases where topological information and characteristics lead to a protrusion degree that is significantly higher than the effect of any possible descretization and/or tessellation noise, which is also the vast majority of real objects. This means that there are object categories, like spheres, etc., where the proposed approach would not lead to optimal performance.

2) Query Range Image: The performance of the method depends mainly on the query range image. The results are much better when the query range image is as indicative as possible for the geometry of the 3-D object. For example, in the fourth row of Fig. 12, the fourth result is from a different class (birds) than the query (airplane). This was expected since the two classes have very similar characteristics. This issue is also illustrated in Fig. 17 as described in the previous section.

Moreover, for the evaluation of the proposed approach, as mentioned in Section V, the range images were randomly generated so as to assure statistical significance of the results. Additionally, and in order to assure statistical significance of the results of the SHREC database that consists of less 3-D models, "precision versus recall" curves were extracted using as queries 400, 800, and 1600 randomly generated range images that correspond to 1, 2, and 4 range images per 3-D object. As expected due to the random generation of the range image queries, the variation of the curves illustrated in Fig. 18 is statistically insignificant.

3) Symmetries: There is a chance that the method will lead to a mathematically wrong result in the case of a symmetric 3-D model. For example, in a case where the query range image is taken from a table, there is a chance that the method will return

a wrong camera position for the correct 3-D model, although most probably the error in the wrong position will be smaller than the error from another 3-D model. This happens because the extracted salient features will be symmetrical, due to the symmetry of the 3-D model. So there is a chance that due to noise in the salient features of both the 3-D model and the range image, there will be a better match from another point of view.

4) Timings: The proposed framework is divided into two separate processing phases: the preprocessing of the 3-D model databases and the online processing of the user's query. The average time to preprocess a 3-D model and generate its salient features is 1 s. Even if not computationally very expensive, the complexity of the preprocessing phase is not considered very important since it is executed only once for each object.

Regarding the online partial matching procedure, the average time to perform pairwise matching is 20 ms, which is and should be sufficiently low so as to compensate for the inherent characteristic of partial matching approaches in general that scale linearly with the number of the 3-D models of the database. This value could be further decreased with parallel processing that can be applied in the proposed hierarchical search scheme.

All experiments have been performed on an Intel Pentium IV 3.2 Ghz with 1 GB RAM.

VI. Conclusion

A novel method for partial 3-D shape search and retrieval using as query a range image was proposed in the present paper. The proposed method can retrieve objects similar to the query range data with very satisfactory accuracy. Most promising features of the method are: 1) it does not use *a priori* information for registering the partial view with the complete 3-D model, 2) it does not impose limitations to the kind of models where it can be applied, 3) it is based only on the salient features of the 3-D model, and d) multiview images can also be used for 3-D matching by creating a range image from multiple views of an object using disparity estimation and structure from motion approaches. Finally, the proposed method is seen to be robust in the presence of noise and occlusion in the query range image.

References

[1] P. Daras, D. Zarpalas, D. Tzovaras, and M. G. Strintzis, "Efficient 3-D model search and retrieval using generalized 3-D radon transforms," *IEEE Trans. Multimedia*, vol. 8, no. 1, pp. 101–114, Feb. 2006.

[2] D. Zarpalas, P. Daras, A. Axenopoulos, D. Tzovaras, and M. G. Strintzis, "3D model search and retrieval using the spherical trace transform," *EURASIP J. Appl. Signal Process.*, vol. 2007, no. 1, pp. 207–207, 2007.

[3] M. Elad, A. Tal, and S. Ar, "Content based retrieval of VRML objects—An iterative and interactive approach," in *Proc. 6th Eurographics Workshop Multimedia*, 2002, pp. 107–118.

[4] R. Osada, T. Funkhouser, B. Chazelle, and D. Dobkin, "Shape distributions," *ACM Trans. Graph.*, pp. 807–832, 2002.

[5] I. Kolonias, D. Tzovaras, S. Malassiotis, and M. G. Strintzis, "Fast content-based search of VRML models based on shape descriptors," *IEEE Trans. Multimedia*, vol. 7, no. 1, pp. 114–126, Feb. 2005.

[6] T. Funkhouser, P. Min, M. Kazhdan, J. Chen, A. Halderman, D. Dobkin, and D. Jacobs, "A search engine for 3D models," *ACM Trans. Graph.*, vol. 22, no. 1, pp. 83–105, 2003.

[7] J. W. H. Tangelder and R. C. Veltkamp, "A survey of content based 3D shape retrieval methods," in *Proc. Int. Conf. Shape Modeling and Applications 2004 (SMI'04)*, 2004, pp. 145–156.

[8] B. Bustos, D. A. Keim, D. Saupe, T. Schreck, and D. V. Vranic, "Feature-based similarity search in 3D object databases," *ACM Comput. Surv.*, vol. 37, no. 4, pp. 345–387, 2005.

[9] X. Lu, A. K. Jain, and D. Colbry, "Matching 2.5D face scans to 3D models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 1, pp. 31–43, Jan. 2006.

[10] T. S. Caetano, T. Caelli, D. Schuurmans, and D. A. C. Barone, "Graphical models and point pattern matching," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 10, pp. 1646–1663, Oct. 2006.

[11] S. Marini, S. Biasotti, and B. Falcidieno, "Partial matching by structural descriptors," in *Proc. Dagstuhl Seminar 06171*, 2006.

[12] P. Nair and A. Cavallaro, "3-D face detection, landmark localization, and registration using a point distribution model," *IEEE Trans. Multimedia*, vol. 11, no. 4, pp. 611–623, Jun. 2009.

[13] D.-Y. Chen and M. Ouhyoung, "A 3D object retrieval system based on multi-resolution Reeb graph," in *Proc. Computer Graphics Workshop*, 2002.

[14] D. Bespalov, W. C. Regli, and A. Shokoufandeh, "Reeb graph based shape retrieval for CAD," in *Proc. DETC'03*, 2003.

[15] D. Bespalov, A. Shokoufandeh, W. C. Regli, and W. Sun, "Scale-space representation of 3D models and topological matching," in *Proc. ACM Symp. Solid Modeling '03*, 2003, pp. 208–215.

[16] S. Biasotti, S. Marini, M. Mortara, G. Patane, M. Spagnuolo, and B. Falcidieno, "3D shape matching through topological structures," in *Proc. DGCI 2003*, 2003, pp. 194–203.

[17] M. Hilaga, Y. Shinagawa, and T. Kohmura, "Topology matching for fully automatic similarity estimation of 3D shapes," in *Proc. SIGGRAPH 2001*, 2001, pp. 203–212.

[18] D. Macrini, A. Shokoufandeh, S. Dickenson, K. Siddiqi, and S. Zucker, "Viewbased 3-D object recognition using shock graphs," in *Proc. Int. Conf. Pattern Recognition (ICPR'02)*, 2002.

[19] S. J. Chua and R. Jarvis, "Point signatures: A new representation for 3D object recognition," *Int. J. Comput. Vis.*, vol. 25, pp. 63–5, 1997.

[20] A. E. Johnson and M. Hebert, "Using spin images for efficient object recognition in cluttered 3D scenes," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 21, no. 5, pp. 635–651, May 1999.

[21] M. Kortgen, G.-J. Park, M. Novotni, and R. Klein, "3D shape matching with 3D shape contexts," in *Proc. 7th Central European Seminar Computer Graphics*, Apr. 2003.

[22] H.-Y. Shum, M. Hebert, and K. Ikeuchi, "On 3D shape similarity," in *Proc. IEEE Computer Vision and Pattern Recognition*, 1996, pp. 526–531.

[23] T. Zaharia and F. Preteux, "3D shape-based retrieval within the MPEG-7 framework," in *Proc. SPIE Conf. 4304*, Jan. 2001, pp. 133–145.

[24] R. Gal and D. Cohen-Or, "Salient geometric features for partial shape matching and similarity," *ACM Trans. Graph.*, vol. 25, no. 1, pp. 130–150, 2006.

[25] V. Cicirello and W. C. Regli, "Machining feature-based comparisons of mechanical parts," in *Proc. SMI 2001*, 2001, pp. 176–185.

[26] M. El-Mehalawi and R. A. Miller, "A database system of mechanical components based on geometric and topological similarity. part i: Representation," *J. Comput.-Aided Des.*, vol. 31, no. 1, pp. 83–94, Jan. 2003.

[27] A. Mademlis, P. Daras, A. Axenopoulos, D. Tzovaras, and M. G. Strintzis, "Combining topological and geometrical features for global and partial 3-D shape retrieval," *IEEE Trans. Multimedia*, vol. 10, no. 5, pp. 819–831, Aug. 2008.

[28] D. Y. Chen, X. P. Tian, Y. T. Shen, and M. Ouhyoung, "On visual similarity based 3D model retrieval," in *Proc. Computer Graphics Forum (EG 2003)*, 2003, vol. 22, no. 3.

[29] D. Vranic, "3D model retrieval," Ph.D. dissertation, University of Leipzig, Leipzig, Germany, 2004.

[30] P. Daras and A. Axenopoulos, "A 3D shape retrieval framework supporting multimodal queries," *Int. J. Comput. Vis.*, to be published.

[31] T. F. Ansary, M. Daoudi, and J.-P. Vandeborre, "A Bayesian 3-D search engine using adaptive views clustering," *IEEE Trans. Multimedia*, vol. 9, no. 1, pp. 78–88, Jan. 2004.

[32] M. Germann, M. D. Breirenstein, I. K. Park, and H. Pfister, "Automatic pose estimation for range images on the GPU," in *Proc. 3D Digital Imaging and Modeling (3DIM'07)*, 2007, pp. 81–90.

[33] A. S. Mian, M. Bennamoun, and R. A. Owens, "A novel representation and feature matching algorithm for automatic pairwise registration of range images," *Int. J. Comput. Vis.*, vol. 66, no. 1, pp. 19–40, Jan. 2006.

[34] A. S. Mian, M. Bennamoun, and R. A. Owens, "Three-dimensional model-based object recognition and segmentation in cluttered scenes," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 10, pp. 1584–1601, Oct. 2006.
[35] B. Matei, Y. Shan, H. S. Sawhney, Y. Tan, R. Kumar, D. Huber, and M. Hebert, "Rapid object indexing using locality sensitive hashing and joint 3D-signature space estimation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 7, pp. 1111–1126, Jul. 2006.
[36] X. Li and I. Guskov, "3D object recognition from range images using pyramid matching," in *Proc. Int. Conf. Computer Vision (ICCV 07)*, 2007, pp. 1–6.
[37] I. Atmosukarto and L. G. Shapiro, "A salient-point signature for 3D object retrieval," in *Proc. ACM Multimedia Information Retrieval*, 2008, pp. 208–215.
[38] R. Ohbuchi, K. Osada, T. Furuya, and T. Banno, "Salient local visual features for shape-based 3D model retrieval," in *Proc. IEEE Int. Conf. Shape Modeling and Applications 2008 (IEEE SMI'08)*, 2008, pp. 93–102.
[39] H. S. Lin, H. M. Liao, and J. Lin, "Visual salience-guided mesh decomposition," in *Proc. IEEE MMSP*, Siena, Italy, 2004.
[40] K. Moustakas, D. Tzovaras, and M. G. Strintzis, "SQ-map: Efficient layered collision detection and haptic rendering," *IEEE Trans. Visual. Comput. Graph.*, vol. 13, no. 1, pp. 80–93, Jan./Feb. 2007.
[41] D. D. Hoffman and M. Singh, "Salience of visual parts," *Cognition*, vol. 13, pp. 29–78, 1997.
[42] Y. Ma, S. Soatto, J. Kosecka, and S. Sastry, *An Invitation to 3-D Vision*. New York: Springer, 2003.
[43] D. Giorgi, S. Biasotti, and L. Paraboschi, Shape Retrieval Contest 2007: Watertight Models Track, Tech. Rep. UU-CS-2007-015, 2007, In Remco C. Veltkamp, Frank B. ter Haar: SHREC2007 3D Shape Retrieval Contest.
[44] P. Shilane, P. Min, M. Kazhdan, and T. Funkhouser, "The Princeton shape benchmark," Shape Modeling International. Genova, Italy, Jun. 2004.
[45] D. Ioannidis, D. Tzovaras, I. G. Damousis, S. Argyropoulos, and K. Moustakas, "Gait recognition using compact feature extraction transforms and depth information," *IEEE Trans. Inf. Forensics Security*, vol. 2, no. 3, pp. 623–630, Sep. 2007.

Georgios Stavropoulos received the Diploma degree in electrical and computer engineering from the Aristotle University of Thessaloniki, Thessaloniki, Greece, in 2006.

He is currently a Research Assistant in the Informatics and Telematics Institute, Centre for Research and Technology Hellas, Thessaloniki. His main research interests include virtual reality, 3-D content-based search, computer vision, stereoscopic image processing, and gait analysis.

Mr. Stavropoulos is a member of the Technical Chamber of Greece.

Panagiotis Moschonas received the Diploma degree in computer science and the M.Sc. degree in digital media from the Aristotle University of Thessaloniki, Thessaloniki, Greece, in 2005 and 2007, respectively.

He serves as a Research Associate in the Informatics and Telematics Institute, Centre for Research and Technology Hellas, Thessaloniki. His main interests include computer vision, human body modeling and tracking, biometrics, computer graphics animation, and GPU programming.

Konstantinos Moustakas (M'07) received the Diploma degree and the Ph.D. degree in electrical and computer engineering from the Aristotle University of Thessaloniki, Thessaloniki, Greece, in 2003 and 2007, respectively.

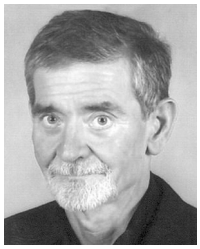
He has been a Teaching Assistant in the same department (2004–2007) and a Visiting Lecturer during 2008–2009. He served as a Research Associate in the Informatics and Telematics Institute, Centre for Research and Technology Hellas, Thessaloniki (2003–2007), where he is currently a Postdoctoral Research Fellow. His main research interests include virtual reality, 3-D content-based search, collision detection, haptics, deformable object modeling and simulation, and stereoscopic image processing. During the latest years, he has been the (co)author of more than 60 papers in refereed journals, edited books, and international conferences. He serves as a regular reviewer for several technical journals. He has been also involved in nine research projects funded by the EC and the Greek secretariat of Research and Technology.

Dr. Moustakas is a member of the IEEE Computer Society and the Technical Chamber of Greece.

Dimitrios Tzovaras (M'09) received the Diploma degree in electrical engineering and the Ph.D. degree in 2-D and 3-D image compression from Aristotle University of Thessaloniki, Thessaloniki, Greece, in 1992 and 1997, respectively.

He is a Senior Researcher in the Informatics and Telematics Institute of Thessaloniki. Prior to his current position, he was a Senior Researcher on 3-D imaging at the Aristotle University of Thessaloniki. His main research interests include virtual reality, assistive technologies, 3-D data processing, medical image communication, 3-D motion estimation, and stereo and multiview image sequence coding. His involvement with those research areas has led to the coauthoring of more than 50 papers in refereed journals and more than 100 papers in international conferences. He has served as a regular reviewer for a number of international journals and conferences. Since 1992, he has been involved in more than 40 projects in Greece, funded by the EC, and the Greek Secretariat of Research and Technology.

Dr. Tzovaras is a member of the Technical Chamber of Greece.



Michael Gerassimos Strintzis (M'70–SM'80–F'04) received the Diploma degree in electrical engineering from the National Technical University of Athens, Athens, Greece, in 1967 and the M.A. and Ph.D. degrees in electrical engineering from Princeton University, Princeton, NJ, in 1969 and 1970, respectively.

He then joined the Electrical Engineering Department at the University of Pittsburgh, Pittsburgh, PA, where he served as Assistant Professor (1970–1976) and Associate Professor (1976–1980). Since 1980, he has been a Professor of electrical and computer engineering at the University of Thessaloniki, Thessaloniki, Greece, and for the period 1999–2009, Director of the Informatics and Telematics Research Institute, Thessaloniki. His current research interests include 2-D and 3-D image coding, image processing, biomedical signal and image processing, and DVD and Internet data authentication and copy protection.

Dr. Strintzis has served as an Associate Editor for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY since 1999. In 1984, he was awarded one of the Centennial Medals of the IEEE.