Content-Based Retrieval of 3D Models

ALBERTO DEL BIMBO and PIETRO PALA
Dipartimento Sistemi e Informatica Università di Firenze, Italy

In the past few years, there has been an increasing availability of technologies for the acquisition of digital 3D models of real objects and the consequent use of these models in a variety of applications, in medicine, engineering, and cultural heritage. In this framework, content-based retrieval of 3D objects is becoming an important subject of research, and finding adequate descriptors to capture global or local characteristics of the shape has become one of the main investigation goals. In this article, we present a comparative analysis of a few different solutions for description and retrieval by similarity of 3D models that are representative of the principal classes of approaches proposed. We have developed an experimental analysis by comparing these methods according to their robustness to deformations, the ability to capture an object's structural complexity, and the resolution at which models are considered.

Categories and Subject Descriptors: H.3.1 [**Information Storage and Retrieval**]: Content Analysis and Indexing

General Terms: Theory, Design, Performance

Additional Key Words and Phrases: 3D shape description, retrieval by content of 3D models, comparative analysis

1. INTRODUCTION

In the past few years, we have observed the availability of technologies for the effective acquisition of digital 3D models of real objects, the establishment of open standards for 3D data interchange, and the increasing use of 3D models in a variety of applications in medicine, engineering, and cultural heritage, to cite only the most important fields. As a result, many collections of 3D models are being created and are available for study and usage. Representative examples of such collections are the Digital Michelangelo Project archive [The Digital Michelangelo Project], including digital models of the sculptures and architecture of Michelangelo, the Protein Data Bank archive [The Protein Data Bank], including structural data of biological macromolecules, and the National Design Repository archive [The National Design Repository], including CAD models of mechanical design, architecture, and electronic parts.

In this context, content-based retrieval of 3D models has become an important subject of research. Several researchers have investigated the possibility of performing effective retrieval of 3D models from large archives, using shape properties instead of text. Therefore, finding adequate descriptors to capture global or local characteristics of the shape, while still maintaining the representation's computational tractability, has become one of the main investigation goals.

This work is partially supported by the Information Society Technologies (IST) Program of the European Commission as part of the DELOS Network of Excellence on Digital Libraries (Contract G038-507618).
Authors' address: A. Del Bimbo, P. Pala, Dipartimento Sistemi e Informatica, Università di Firenze, via S. Marta 3, 50139 Firenze, Italy; email: {delbimbo,pala}@dsi.unifi.it.

Comparative analyses of 3D object descriptions have been developed in the past for the purpose of comparing techniques for 3D object recognition. Notable survey papers are those by Besl and Jain [1985] and Campbell and Flynn [2001]. For content-based retrieval of 3D objects, a recent survey by Tangelder and Veltkamp [2004] presents an overview of the solutions that have been proposed in the literature for the description of 3D object shapes and similarity matching. In this article, the authors have identified a number of basic characteristics under which systems for retrieval by content of 3D models are analyzed: shape representation, properties of dissimilarity measures, discrimination ability, ability to perform partial matching, robustness, and need for pose normalization.

However, for the purposes of comparison, it is also important to produce experimental evidences of advantages and limitations for each solution so as to discover and highlight elements to support the choice of the most appropriate solution depending on the target application domain. In fact, the huge variability and different complexities of 3D shapes, the different resolutions at which they can be managed, and the many contexts of applications and distinct goals of use make it impossible to identify one particular solution as superior to all the others only on the basis of a theoretical analysis. From the application viewpoint, 3D model descriptions must be analyzed according to their capability of supporting local matching, the resolution at which the 3D model is considered, the structural complexity of the object that can be managed, and the robustness to deformations applied to the model.

Local matching is important when 3D objects must be retrieved that have similar signatures to the query object in some parts. This might occur in cultural heritage applications where the artist signature can be discovered from a few characterizing elements such as, for instance, folds of dresses in a sculpture. In this case, segmentation into regions and fine local shape descriptors with topological information and adequate distance functions are required. Resolution of 3D shapes representations is directly linked to the number of mesh polygons (or vertices) that are used to describe the object. Different applications require different resolutions for 3D models. In cultural heritage applications, 3D models of 3D objects must be managed at high resolution. The 3D digital model and its description must necessarily capture even small details of the 3D shape of the original object. This strict requirement can be relaxed when 3D models are used only for presentation to the general public. Similar requirements are necessary in engineering and medical applications. For videogames and 3D models available on the web, resolution is typically much lower and, in general, the detail representation requirement can be left over. Structural complexity of 3D models is concerned with the presence/absence of symmetry, the number of extruding and intruding parts of the object, and their distribution and size. While some applications only deal with objects of a limited number of classes of complexity, other applications, like Web searching of 3D models, may deal with objects of many classes of complexity. Robustness to deformations is important when we want to retrieve copies of the same object that have been subjected to changes due to bending, relaxation, skewing, or smoothing, or other surface perturbations. In cultural heritage applications, for example, descriptions that allow controlled insensitivity to diverse degrees of deformations can be useful for 3D object retrieval.

While the capability of supporting local matching is difficult to assess in full detail, due to the high variability of cases and conditions under which extruding/intruding parts can manifest, a comparison of different solutions can be performed by checking recall and precision of global shape descriptors. In the following, we present an experimental comparison of several solutions for the description of 3D model shapes for content-based retrieval by evaluating their performance with respect to the robustness to deformations applied to the model, the structural complexity of the object that is managed, the extent to which structural elements are captured in the descriptors, and the resolution at which the 3D model is considered.

The article is organized as follows. In Section 2, we review solutions for shape description that have been proposed in the literature. Among them, five approaches have been selected as representative of

the major categories and described in detail in Section 3. In Section 4 these solutions are compared by evaluating their performance in terms of precision and recall. Considerations and conclusions are drawn in Section 5.

2. 3D MODELS DESCRIPTION

Methods addressing retrieval by global similarity of 3D models can be classified according to the principles under which shape representations are derived. Four categories can be distinguished: (i) primitive-based (ii) statistics-based, (iii) geometry-based, and (iv) view-based.

Primitive-based approaches represent 3D objects with reference to a basic set of parameterized primitive elements. Parameter values are used to control the shape of each primitive element and are determined so as to best fit each primitive element with a part of the model. An example of this class of solutions has been proposed in Kriegel [1998] where surface segments are used to model the potential docking sites of molecular structures. The proposed approach develops on the approximation error of the surface. However, assumptions on the form of the function to be approximated limit the applicability of the approach only to special contexts.

Shape descriptions based on statistical models consider the distribution of local features measured at the vertices of the 3D object mesh. The simplest approach approximates a feature distribution with its histogram. Any metric can be used to compute the similarity between the distributions of two models. In Vandeborre [2002], representation of 3D objects is captured using histograms of the curvature of mesh vertices. In Osada et al. [2002], the authors have introduced shape functions as distributions of shape properties. Each distribution is approximated through the histogram of the values of the shape function. Local features such as the distance of mesh vertices to the centroid, the distance between random pairs of vertices of the mesh, and the area of triangles between three random vertices of the mesh are considered. Ohbuchi et al. [2003b] have defined shape functions suited for objects with rotational symmetry. They have considered the principal axes of inertia of the object and used as shape functions three histograms: the moment of inertia about the axis, the average, and the variance of the distance from the surface to the axis.

A limitation of statistical approaches is that they do not take into account how local features are spatially distributed over the model surface. For this purpose, spatial map representations have been proposed to capture either the spatial location of an object or the spatial distribution of relevant features on the object surface. Map entries correspond to locations or sections of the object and are arranged so as to preserve the relative positions of the object features. In Vranić [2001], a solution is presented in which a surface is described by associating with each ray from the origin, the value of the distance to the last point of intersection of the model with the ray and then extracting spherical harmonics for this spherical extent function. In Assfalg [2003], a method is proposed for description of shapes of 3D objects whose surface is a simply connected region. The 3D object is deformed until it is a function on the sphere. Then, information about surface curvature is projected on to a 2D map which is used as the descriptor of the object shape. Recently, curvature correlograms have been proposed [Antini 2005] to capture the spatial distribution of curvature values on the object surface.

Geometry-based solutions use geometric properties of the 3D object and their measures as global shape descriptors. Many approaches have been proposed. The system developed within the Nefertiti project supports retrieval of 3D models based on both shape geometry and appearance (i.e., color and texture) [Paquet 1999]. Kolonias et al. [2001] have used dimensions of the object bounding box (i.e., its aspect ratios), a binary voxel-based representation of geometry and "set of paths", that outline the shape (*model routes*). In Mokhtarian [2001], points where Gaussian and median curvature are maxima and torsion is maximum have been considered as representative of the object shape.

Elad et al. [2001] have used moments (up to the 7th order) of surface points, exploiting the fact that, different from the case of 2D images, 3D models computation of moments is not affected by self-occlusions. In Zhang [2001], a representation based on moment invariants and Fourier transform coefficients has been combined with active learning to take into account user relevance feedback and improve the effectiveness of retrieval. In Novotni [2003], a method has been presented to compute 3D Zernike descriptors from voxelized models. 3D Zernike descriptors capture object coherence in the radial direction and in the direction along a sphere. However, the effectiveness of the approach is strongly dependendent on the quality of the voxelization process.

View-based descriptions use a set of 2D views of the model and appropriate descriptors of their content to represent the 3D object shape. One problem with this approach is concerns the need for representations that are computationally tractable. In Mahmoudi [2002] and Ohbuchi [2003], a number of views of the 3D object is taken and, for each view, the 2D profile is considered. Hence, PCA has been used to reduce all object views to a limited set of representative views that are used to represent the whole 3D object shape. In Assfalg [2004], signatures of spin images have been proposed. In their original formulation [Johnson 1999], spin images are 2D histograms of the surface locations around a point. Each mesh vertex defines a family of cylindrical coordinate systems, with the origin in $\boldsymbol{p}$, and the axis along $\boldsymbol{n}$. The spin image is obtained by projecting all the other vertices over the tangent plane, retaining for each vertex the radial distance and the elevation and discarding the polar angle. The 2D information of the spin image is reduced to a one-dimensional feature vector, partitioning the image into a finite number of regions and considering the point density in each region. Signatures are hence derived by clustering all spin image vectors and taking the centers of the clusters as their representatives. In Chen [2003], 2D views (*Light fields*) of the object are taken from observation points uniformly distributed on the surface of a sphere centered in the object's centroid. For each of these views, Zernike moments and Fourier descriptors are computed so as to reduce the two-dimensional information to a one-dimensional feature vector. Computational complexity of retrieval is reduced by a multistep approach supporting early rejection of nonrelevant models.

3. SELECTED SOLUTIONS FOR COMPARATIVE ANALYSIS

For the purpose of comparison of 3D shape descriptions, we have taken into consideration five solution as representatives of the three major categories of approaches that have been proposed in the literature. In particular, we have selected:

For Statistics-Based Solutions

—Curvature histograms ([Vandeborre 2002]),
—Shape functions ([Osada 2002]);

For Geometry-Based Solutions

—Geometric moments ([Elad 2001]);

For View-Based Solutions

—Spin Image signatures ([Assfalg 2004]),
—Light fields ([Chen 2003]).

Principles of operation, characterizing elements, and details of implementation complete the rest of the section.

3.1 Curvature Histograms

Surface curvatures are estimated at a generic vertex v_i of the 3D object mesh by considering the variations of the surface normal over the *platelet* V^{v_i} of vertex v_i, that is, the set of all 1-connected mesh vertices around v_i. Given a generic vertex of the platelet $v_j \in V^{v_i}$, $v_j^{\perp}$ is defined as the normal to $\mathcal{M}$ at point v_j. The curvature $\bar{k}_{v_i}$ at vertex v_i is then estimated as:

$$\bar{k}_{v_i} = \frac{1}{2} \frac{\sum_{v_j \in V^{v_i}} |v_i^{\perp} - v_j^{\perp}|}{|V^{v_i}|}. \tag{1}$$

According to this definition, the value of $\bar{k}_{v_i}$ is always in $[0, 1]$.

Curvature values are quantized into $2N + 1$ classes of discrete values according to:

$$\mathcal{Q}(\bar{k}) = \begin{cases} N\Delta & \text{if } \bar{k} > N\Delta \\ i\Delta & \text{if } \bar{k} \in [i\Delta, (i+1)\Delta) \\ -i\Delta & \text{if } \bar{k} \in [-i\Delta, -(i+1)\Delta) \\ -N\Delta & \text{if } \bar{k} < -N\Delta \end{cases} \tag{2}$$

with $i \in \{0, \ldots, N-1\}$ and Δ a suitable quantization parameter. In the experiments, $N = 100$ and $\Delta = 0.15$ have been used.

The histogram of curvature $h_{c_i}(\mathcal{M})$ of the mesh $\mathcal{M}$ is defined as:

$$h_{c_i}(\mathcal{M}) = N_V \cdot \Pr_{v_i \in \mathcal{M}} [v_i \in \mathcal{M}_i],$$

where N_V is the number of mesh vertices.

To evaluate the dissimilarity between two curvature histograms, the Kullback-Leibler divergence has been used:

$$d_{KL} = \sum_{k=0}^{N} h_k(\mathcal{M}_1) log \frac{h_k(\mathcal{M}_1)}{h_k(\mathcal{M}_2)}. \tag{3}$$

3.2 Shape Functions

Object shapes are described by the probability distribution of a shape function that measures geometric properties of the 3D object. In our implementation, the shape function measures the Euclidean distance between pairs of vertices of the 3D object mesh $\mathcal{M}$. The probability distribution has been approximated through the 128-bin histogram that quantizes values of the shape function $H(\mathcal{M}) = (h_0(\mathcal{M}), \ldots, h_{127}(\mathcal{M}))$, where

$$h_k(\mathcal{M}) = \Pr_{v_i, v_j \in \mathcal{M}} (\delta(v_i, v_j) = k),$$

and $\delta(v_1, v_2)$ is the normalized and quantized value of the distance between two vertices $v_1 = (x_1, y_1, z_1)$ and $v_2 = (x_2, y_2, z_2)$ of the object mesh:

$$\delta(v_1, v_2) = 128 \left\lfloor \frac{\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}}{L} \right\rfloor.$$

All objects must be uniformly scaled so as to fit within a 3D box of predefined size so that values of the shape function are normalized with respect to the box diagonal.

Dissimilarity is evaluated according to the Kullback-Leibler divergence (Equation (3)).

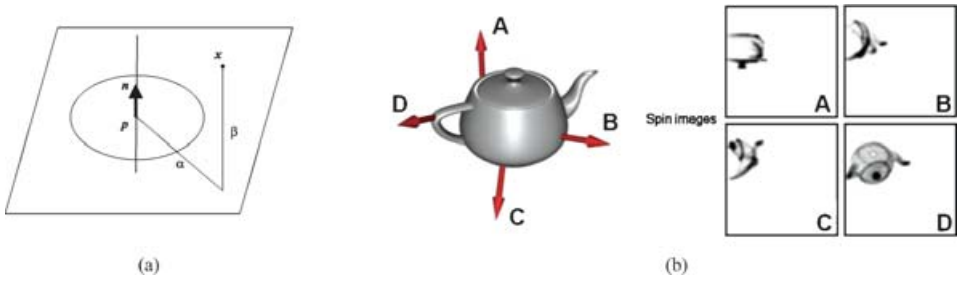

Fig. 1. (a) The object centred 3D coordinates system for the Spin Image representation. (b) A sample 3D model and spin images computed from four distinct points on the object surface.

3.3 Geometric Moments

For a 3D object polygonal mesh of N vertices $\{P_i = (x_i, y_i, z_i)\}_{i=1}^{N}$, the (prq)-th moment is defined according to the following expression:

$$m_{pqr} = \frac{1}{N} \sum_{i=1}^{N} w_i x_i^p y_i^q z_i^r,$$

where w_i is a weight proportional to the area of the polygonal faces adjacent to vertex P_i. Weights w_i make the moment-based representation invariant with the number of vertices of the mesh. Scale invariance is obtained by taking the coordinate system centered at the centroid $m_{100}, m_{010}, m_{001}$ of the model and scaling all models uniformly so as to fit within a regular 3D box of predefined size.

For our experiments, moments up to the 3-th order have been used. A generic object is therefore represented through a feature vector $\mathbf{f}_{mom} \in \mathbb{R}^{19}$. Dissimilarity between two models has been measured by the Euclidean distance between their feature vectors.

3.4 Spin Image Signatures

For each mesh vertex $\boldsymbol{p}$, a *spin image* is built mapping any other mesh vertex $\boldsymbol{x}$ onto a two-dimensional space (see Figure 1(a)) according to:

$$S_{\boldsymbol{p}}(\boldsymbol{x}) \rightarrow [\alpha, \beta] = [\sqrt{\|\boldsymbol{x} - \boldsymbol{p}\|^2 - (\boldsymbol{n} \cdot (\boldsymbol{x} - \boldsymbol{p}))^2}, \boldsymbol{n} \cdot (\boldsymbol{x} - \boldsymbol{p})],$$

where $\boldsymbol{n}$ is the normal direction to the mesh at vertex $\boldsymbol{p}$.

We have derived grey-level spin images by considering the density of mesh vertices that map on the same point of the spin image and evaluating the influence of each vertex over the neighboring pixels of its projection, according to a bilinear interpolation scheme. In Figure 1(b), a sample model of a teapot is shown with spin images computed from four different vertices of the model.

Following an approach similar to that expounded in Goshtasby [1985], spin images have been partitioned into sectors of circular crowns for the upper $(\beta > 0)$ and lower $(\beta < 0)$ half-planes and circular sectors centered in the origin (See Figure 2.) For each of them, we have considered the number of vertex projections that fall in the regions $C^p = \langle cp_1, \ldots, cp_{np} \rangle$, $C^n = \langle cn_1, \ldots, cn_{nn} \rangle$, and $S = \langle s_1, \ldots, s_{ns} \rangle$, respectively. Experimentally, $np = nn = ns = 6$ has been found to be a satisfactory trade-off between representation compactness and selectivity. This leads to compress on of the spin image informative content into a 18-dimensional vector $D = \langle C^p, C^n, S \rangle$.

Description vectors have been clustered using fuzzy clustering [Bezdek 1999] so as to take the centers of the clusters as signatures of the spin image representation. The optimal number of clusters c has been derived according to Kim [2001]: given two functions that express a measure of under- and

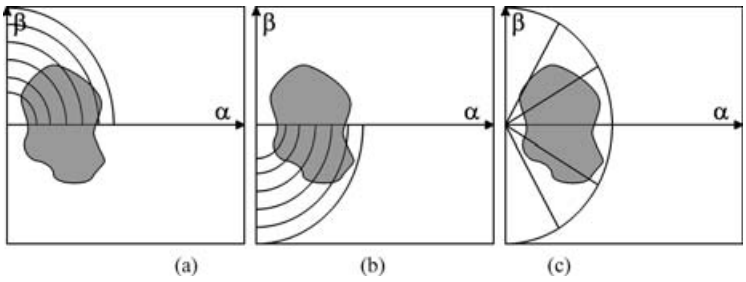

Fig. 2. Circular crowns and sectors used to extract a signature from a spin image.

Fig. 3. Example of views used for the Light Field representation of a Mercur statue model.

over-partitioning, respectively, the optimal number of clusters is the number that minimizes the sum of the two functions— representing the trade-off between under- and over-partitioning.

Similarity between Spin Image signatures of 3D objects is obtained considering the permutation $p : \{1, \ldots, l\} \rightarrow \{1, \ldots, k\}$ that minimizes the sum of distances between the corresponding cluster centers.

3.5 Light Fields

Camera positions for the Light Fields representation are chosen at the vertices of a regular dodecahedron so that 20 views of the object are extracted. Each view is obtained as a binary image (color or gray-level information of the object surface is disregarded). Since views taken from opposite vertices of the dodecahedron are identical, only 10 views are sufficient to represent a 3D object with enough accuracy (an example is shown in Figure 3).

For each view, image content is represented using a combination of 35 Zernike moments and 10 Fourier descriptors, all included in a 45-dimensional feature vector that represents the Light Field descriptor of the view. The L1 distance between two feature vectors $f_1 = \{f_{1,1}, f_{1,2}, \ldots, f_{1,45}\}$ and $f_2 = \{f_{2,1}, f_{2,2}, \ldots, f_{2,45}\}$ is used to measure their dissimilarity:

$$d(f_1, f_2) = \sum_{i=1}^{45} |f_{1,i} - f_{2,i}|.$$

The dissimilarity of two objects is hence evaluated as the minimal dissimilarity obtained by rotating the viewing sphere of the Light Field descriptor of the first object with respect to the other. In this way, invariance to object rotation is preserved.

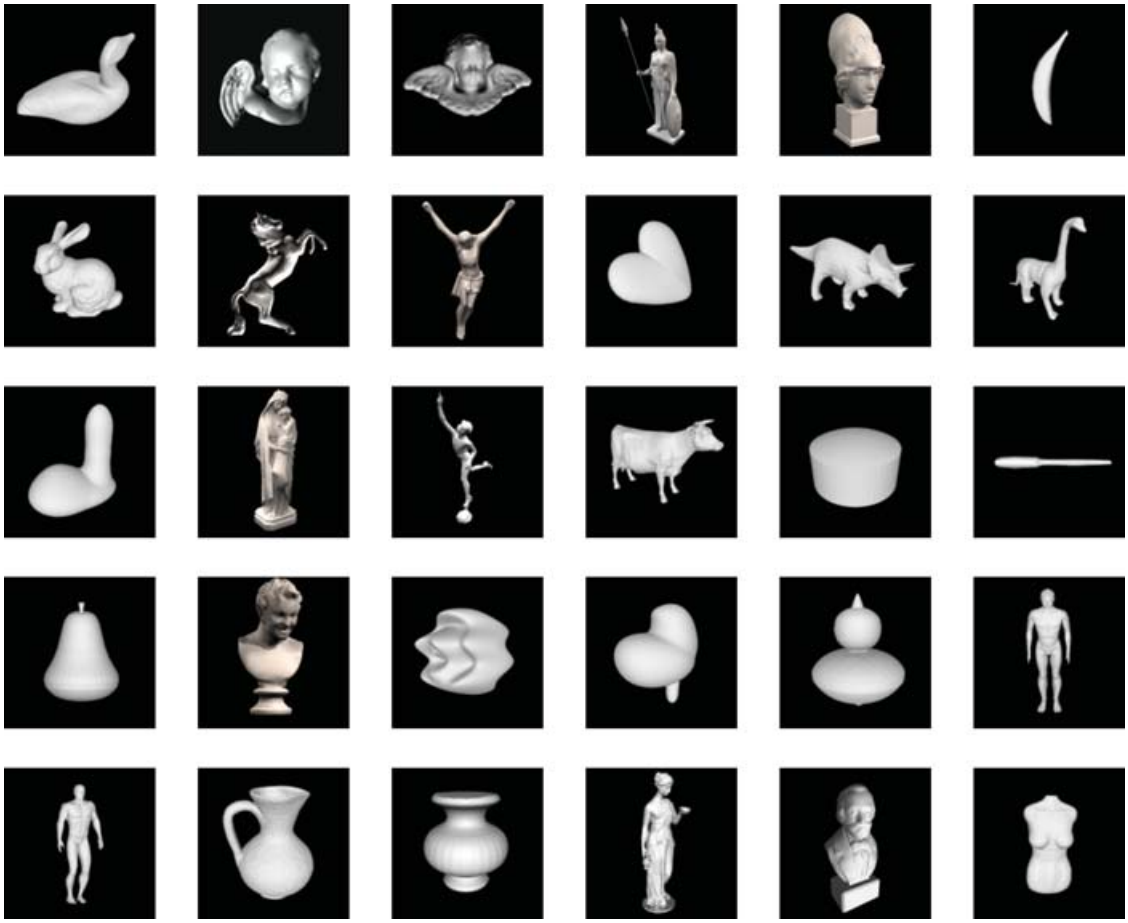

Fig. 4. Sample 3D models from the Art-Models database.

4. COMPARATIVE ANALYSIS AND EXPERIMENTAL RESULTS

The effectiveness of the shape descriptions selected is evaluated from two distinct perspectures: considering the robustness of the description with respect to geometric deformations applied to the models (Section 4.1) and considering the capability of the description to represent the structure peculiarities of the model for models of different structural complexity (Section 4.2). Two different databases were used.

—The *Art-Model* database, including about 300 high-resolution models from miscellaneous sources (many are high quality scans of 3D models from the De Espona 3D Models Encyclopedia[1]), was used to test the robustness to geometric deformations. The models have mesh resolution ranging from 3180 to 5150 mesh vertices. Models were subjected to special deformations, using the editing functions of 3DStudio Max authoring software. A modified version of the objects was also obtained adding Gaussian noise to surface points (so that they are moved from their original locations). Sample models from the *Art-Model* database database are shown in Figure 4.

[1]De Espona 3D Models Encyclopedia is available at `http://www.deespona.com`.

—The *Princeton-Model* database, including 500 low and high resolution models of the Princeton Shape Benchmark[PSB] archive, was used to check the dependency of recall and precision on object structural complexity. Each model in this database is annotated with the type of the object represented. Object classification is hierarchical and takes into account both the function and the structure of the object. A variety of types, including humans (standing and walking humans), animals, and buildings (homes, temples, castles, etc.) are included. Resolutions change considerably from model to model, ranging from 22 to 10120 mesh vertices.

4.1 Robustness to Geometric Deformations

To test robustness to model deformations, we have applied different types of deformations at different degrees of intensity to some models of the Art-Models database (in Figure 5, effects of deformations on a sample Model are shown). Modified versions of each object model were used as queries against the Art-Model database. Precision and recall were computed separately for each class of modifiers. In Figure 6, deformed versions of bust models are shown that are used for ground truth in our experiments.

Results in terms of interpolated precision vs recall curves are shown in Figures 7–11. It can be observed that Geometric Moments and Curvature Histograms almost always present poor performance. In contrast, Spin Image signatures, Light Field, and Shape Functions have different performance for different types of deformations. Performance also changes according to the degree of deformation applied. In particular, they have similar performance with respect to changes induced by skew and spherify modifiers (Figures 9, 10). In both cases, precision values decrease rapidly as the degree of deformation increases. Light Field has lower precision than Spin Image signatures at high degrees of skewness and spherify deformations.

For deformations induced by bend and relax modifiers (Figures 7, 8), precision and recall of Light Field decrease rapidly. Spin Image signature and Shape Functions, however, have a more robust response.

The highest insensitivity to Gaussian noise is provided by the Light Fields representation (Figure 11). Shape Functions have similar, although lower performance. Performance of Spin Image signatures degrades at high noise values.

4.2 Capability to Represent Structure Peculiarities

To investigate the capability of each representation to capture the structural peculiarities of the models, precision and recall have been derived separately for a set of 3D models of the Princeton Shape Benchmark database having different structural complexity. The object type information provided for each model has been used as groundtruth to compute precision and recall curves. We have selected five representative classes of models: Gazebos, Ants, Humans, Hands, and Faces. Gazebos and Ants are fairly complex models with many protruding elements. Gazebos have structural symmetry. Humans and Hands have less complex structure, but still have prominent protruding elements. Faces are simpler models, with concave and open surface and minor protrusions. For each class, the first three models in the class have been used as query objects, and precision and recall values have been computed averaging the results obtained for the three queries. Figure 12 shows the performance of the different representations for the five representative classes of models. Spin Image signatures and Light Field always have the highest performance for classes of objects with high and medium structural complexity (Gazebos, Ants, Humans, and Hands). Spin Image signatures ranks best for Ants and Hands, while Light Fields ranks best for Gazebos and Humans classes. Figures 13–17 show, for each considered class of objects, retrieval results for the top-performing solutions. For each query, the top six ranked models are displayed. Retrieved models not belonging to the same class of the query model are highlighted.

In summary, in most of the cases analyzed, Light Fields and Spin Image signatures have superior performance. Results presented in Figure 13 show that Light Field descriptors are better than

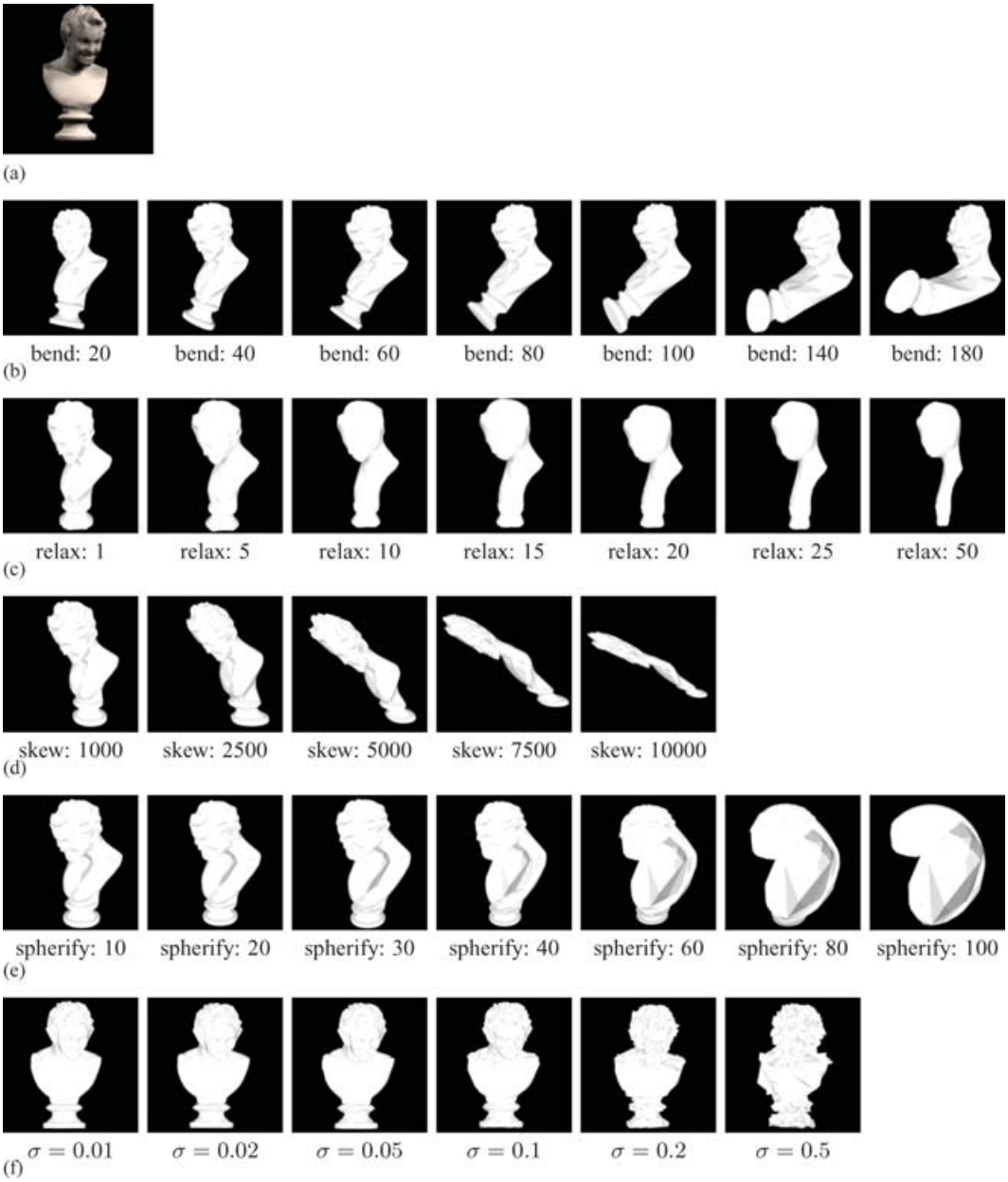

Fig. 5. Effects of bend (b), skew (c), relax (d) spherify (e) and Gaussian noise (f) modifiers on the sample model(a).

Spin Images signatures for capturing characterizing elements of models with many intruding extruding parts. For models in class 118 (Figure 14), we notice that retrieval results are almost equivalent for the two solutions from the perceptual viewpoint. Indeed the different performance figures observed in Figure 12 are due to the fact that, in the PSB database, models of humans are in a different class from walking humans. Shape Functions, Geometric Moments, and Curvature

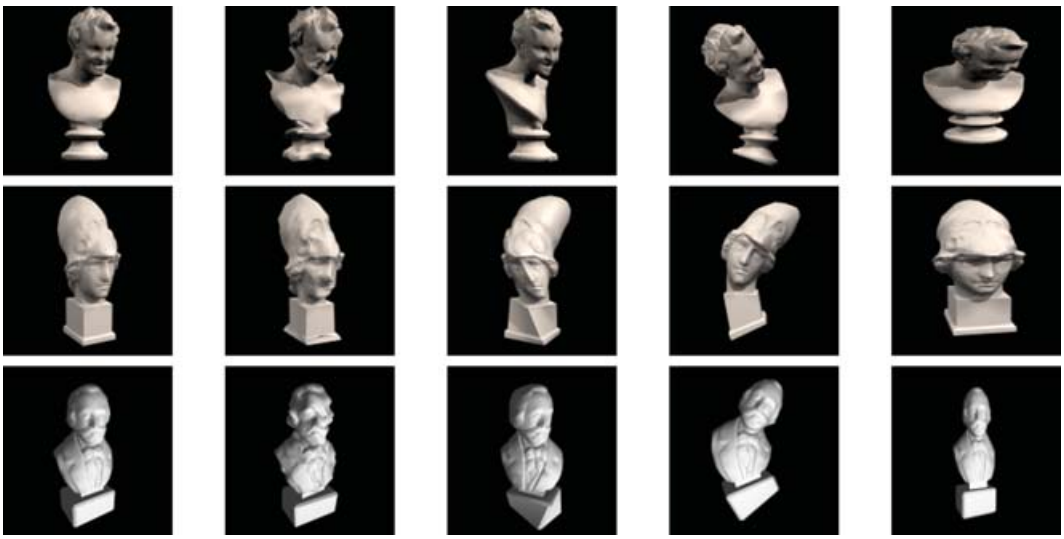

Fig. 6. Models of busts included in the Art-Models database.

Histograms almost always have lower performance, although differences are attenuated for simpler models.

Comparing these conclusions with those derived for robustness to object deformations where Spin Image signatures and Shape Functions showed superior performance with respect to Light Fields, we notice that performance measures are affected by the resolution at which models are considered. Curvature Histograms, Shape Functions, Geometric Moments, and Spin Image signatures representations are obtained by cumulating measures of object properties at the vertices of the mesh. In particular, they take into account the mesh curvature for Curvature Histograms, the distance between vertex pairs for Shape Functions, the position of the vertex for Geometric Moments, and the projection of vertex positions for Spin Image signatures. As a result, the performance of these representations may change drastically if the number of mesh vertices changes substantially. While Spin Image signatures provide effective representations for models captured at medium to high resolution (as those in the Art-Models archive), they don't perform very well with models with a small number of mesh vertices (as in the models from the PSB achive considered in the experiments).

On the other hand, Light Fields are derived as measurements of rendered views of the object. This makes the description less sensitive to the number of mesh vertices (the smoothness of the object silhouette is affected at most).

5. CONCLUSIONS

In this article, we have presented a comparative evaluation of 3D object representations for the purpose of retrieval by content from digital archives. Five different approaches have been considered, Curvature histograms and Shape functions for Statistics-based solutions, Geometric moments for Geometry-based solutions, and Spin Image signatures and Light Fields for View-based solutions.

Based on results of the comparative evaluation, the following considerations have been derived. Geometric moments do dot appear to be able to capture salient and discriminating features of 3D objects. Small and local deformations of the object shape can result in significant changes of moments' values. Statistics-based solutions never provide better performance than the other solutions probably because they do not take into account the spatial distribution of shape features. View-based approaches

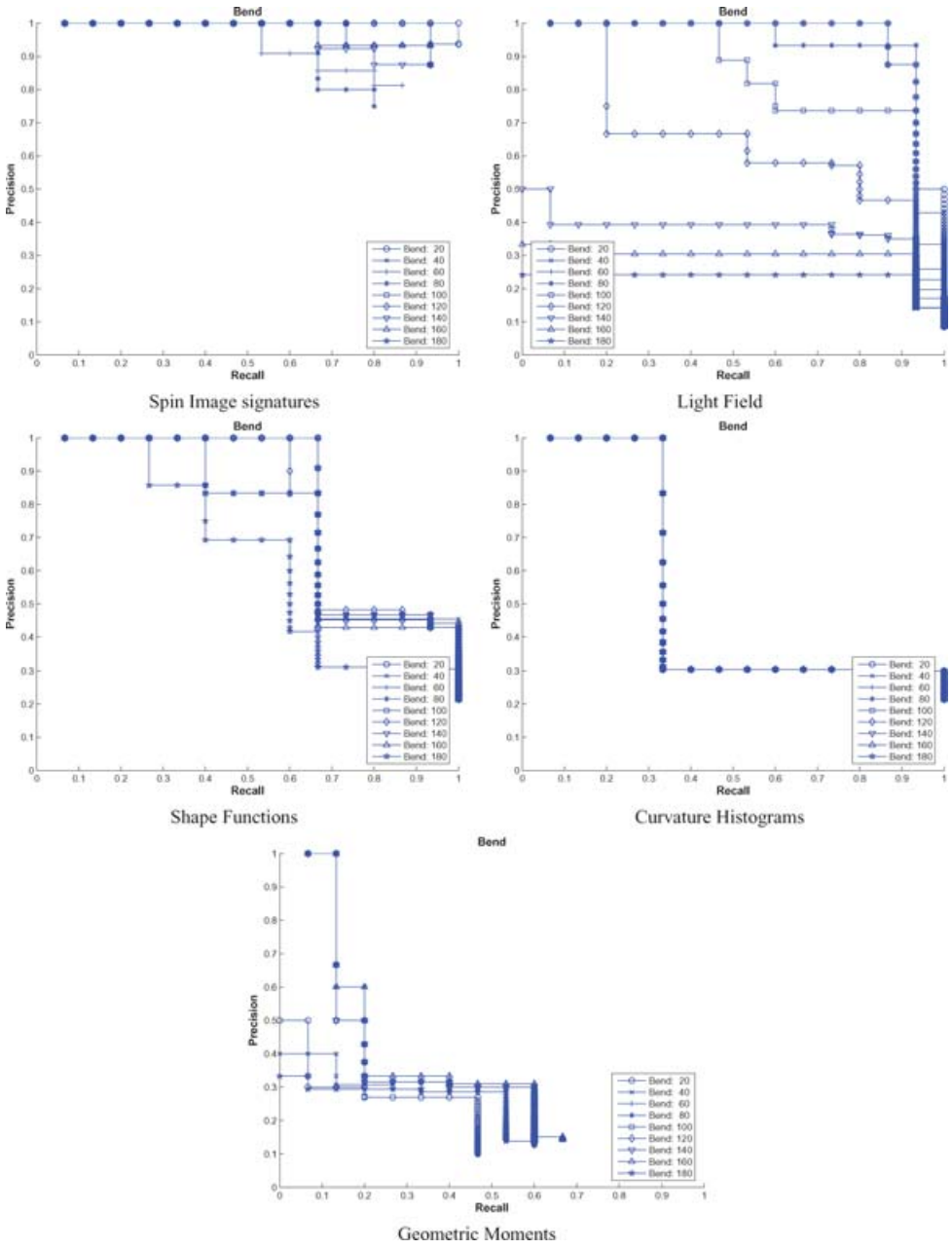

Fig. 7. Robustness to Bend modifier for Spin Image signatures, Light Field, Shape Functions, Curvature Histograms, and Geometric Moments representations.

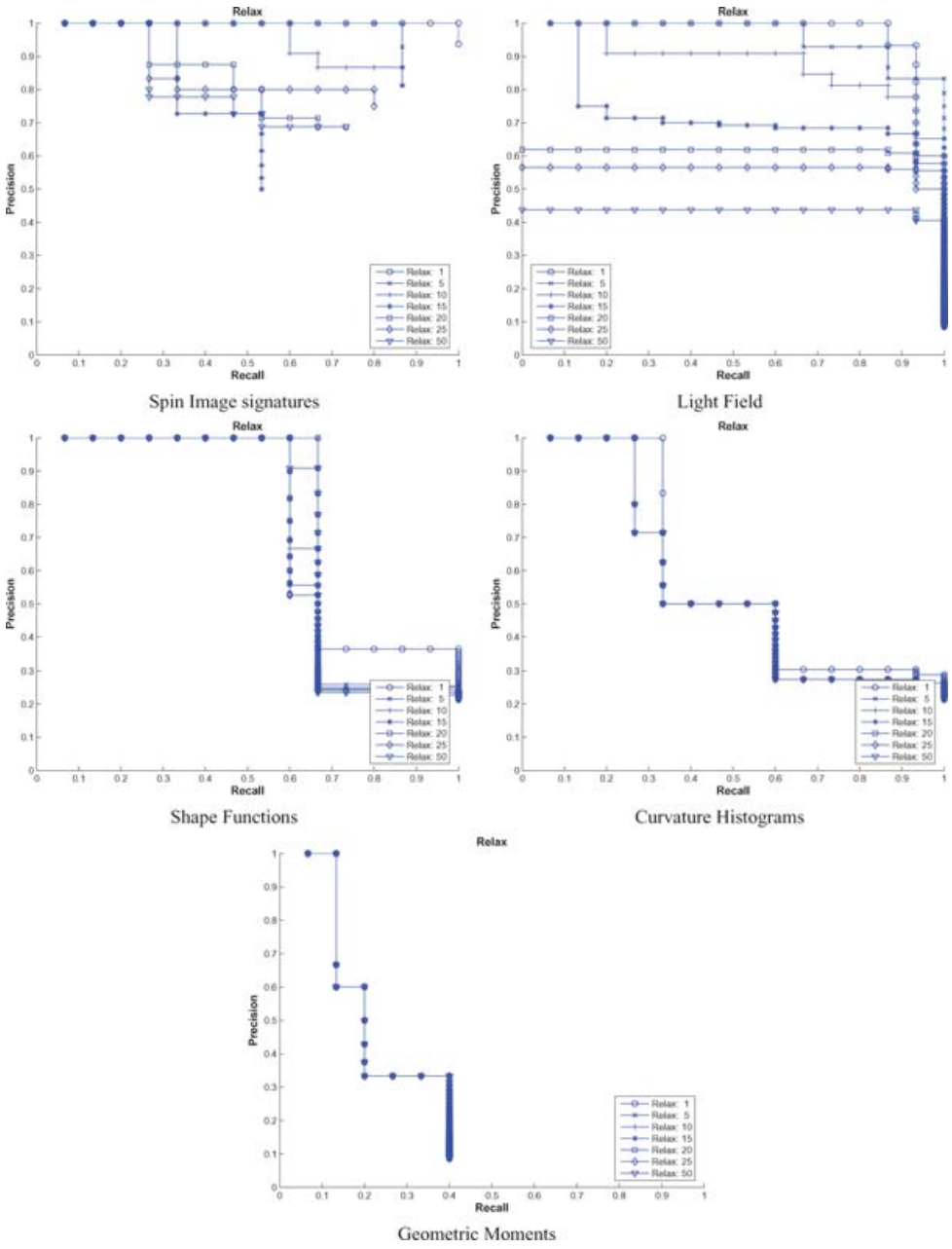

Fig. 8. Robustness to Relax modifier for Spin Image signatures, Light Field, Shape Functions, Curvature Histograms, and Geometric Moments representations.

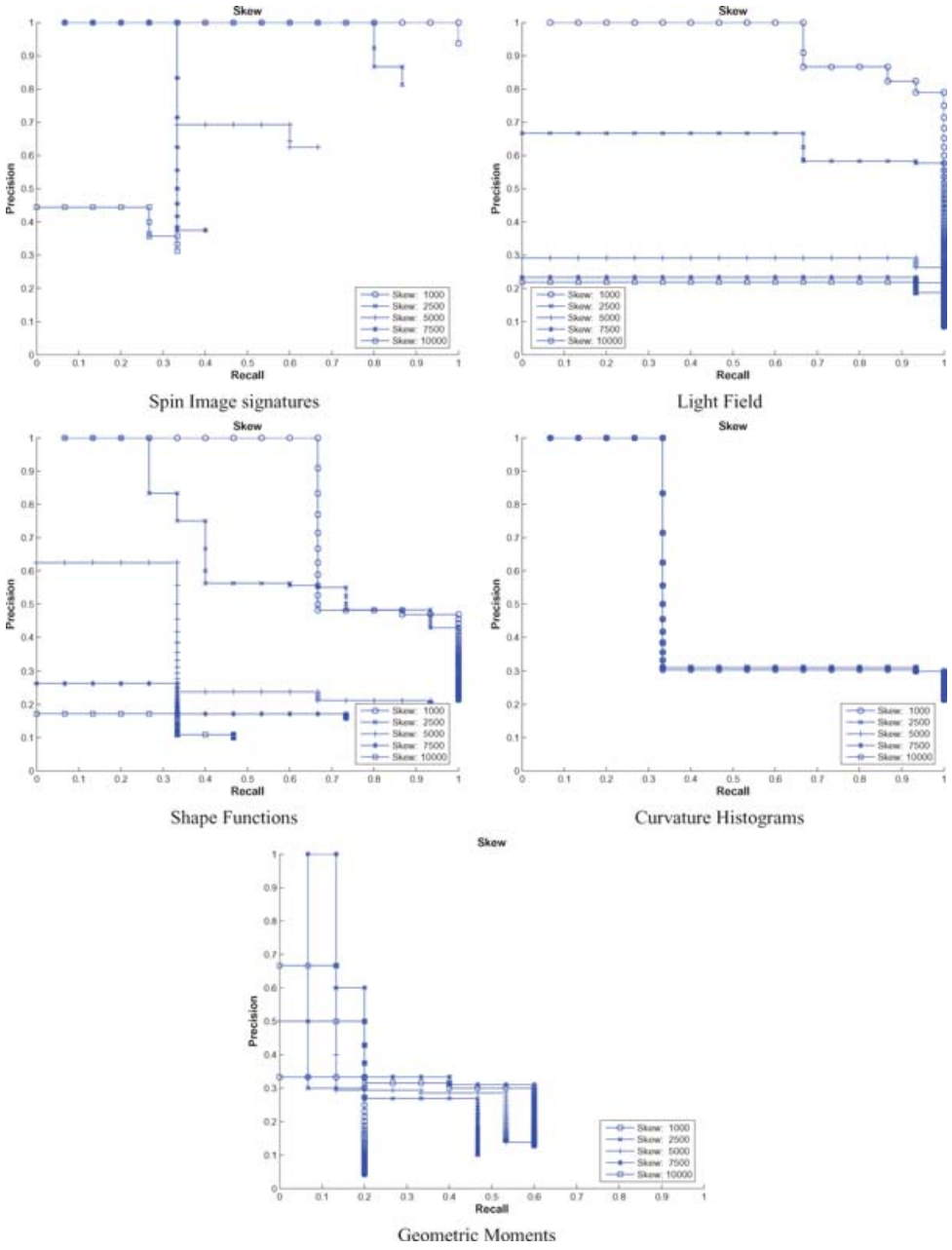

Fig. 9. Robustness to Skew modifier for Spin Image signatures, Light Field, Shape Functions, Curvature Histograms, and Geometric Moments representations.

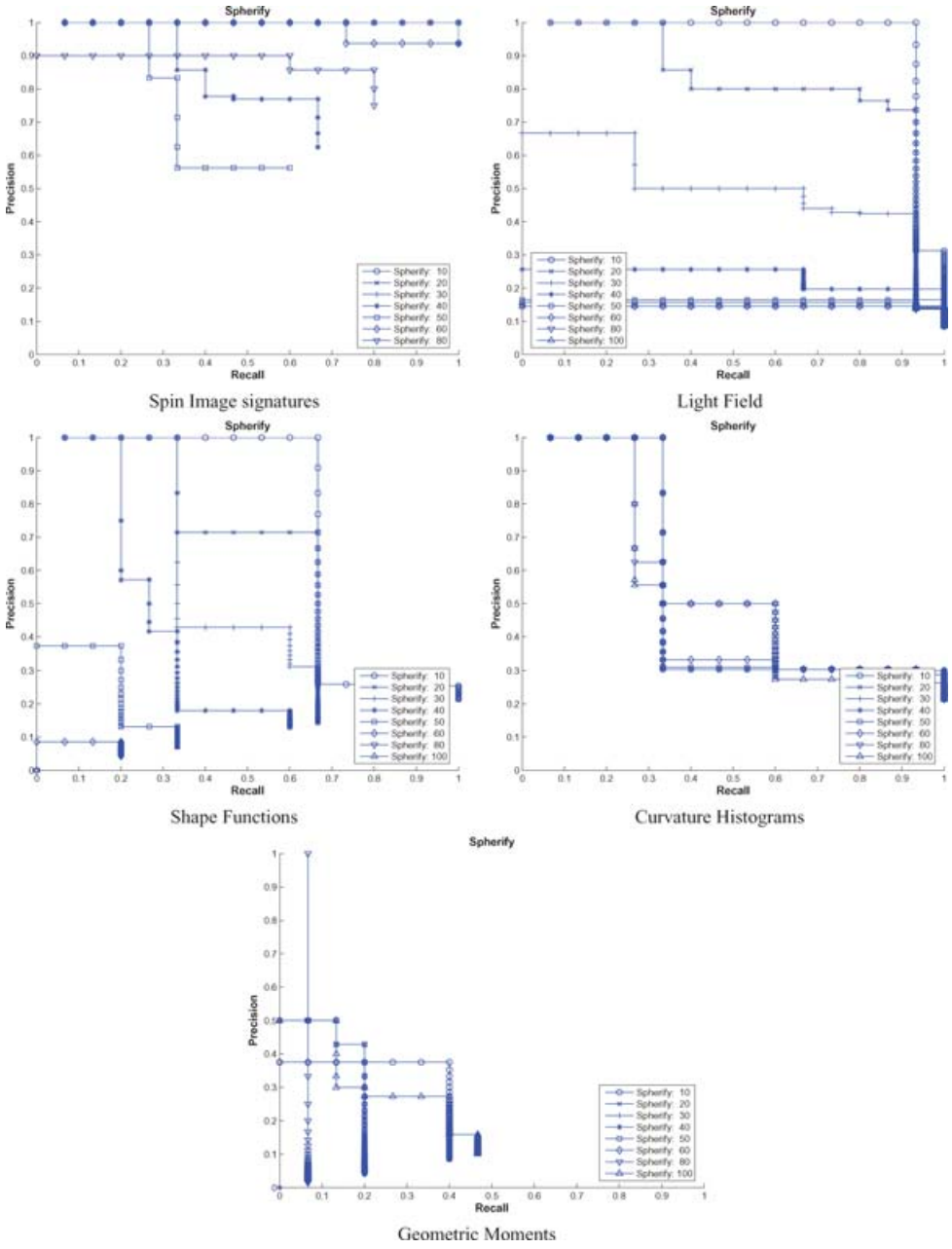

Fig. 10. Robustness to Spherify modifier for Spin Image signatures, Light Field, Shape Functions, Curvature Histograms, and Geometric Moments representations.

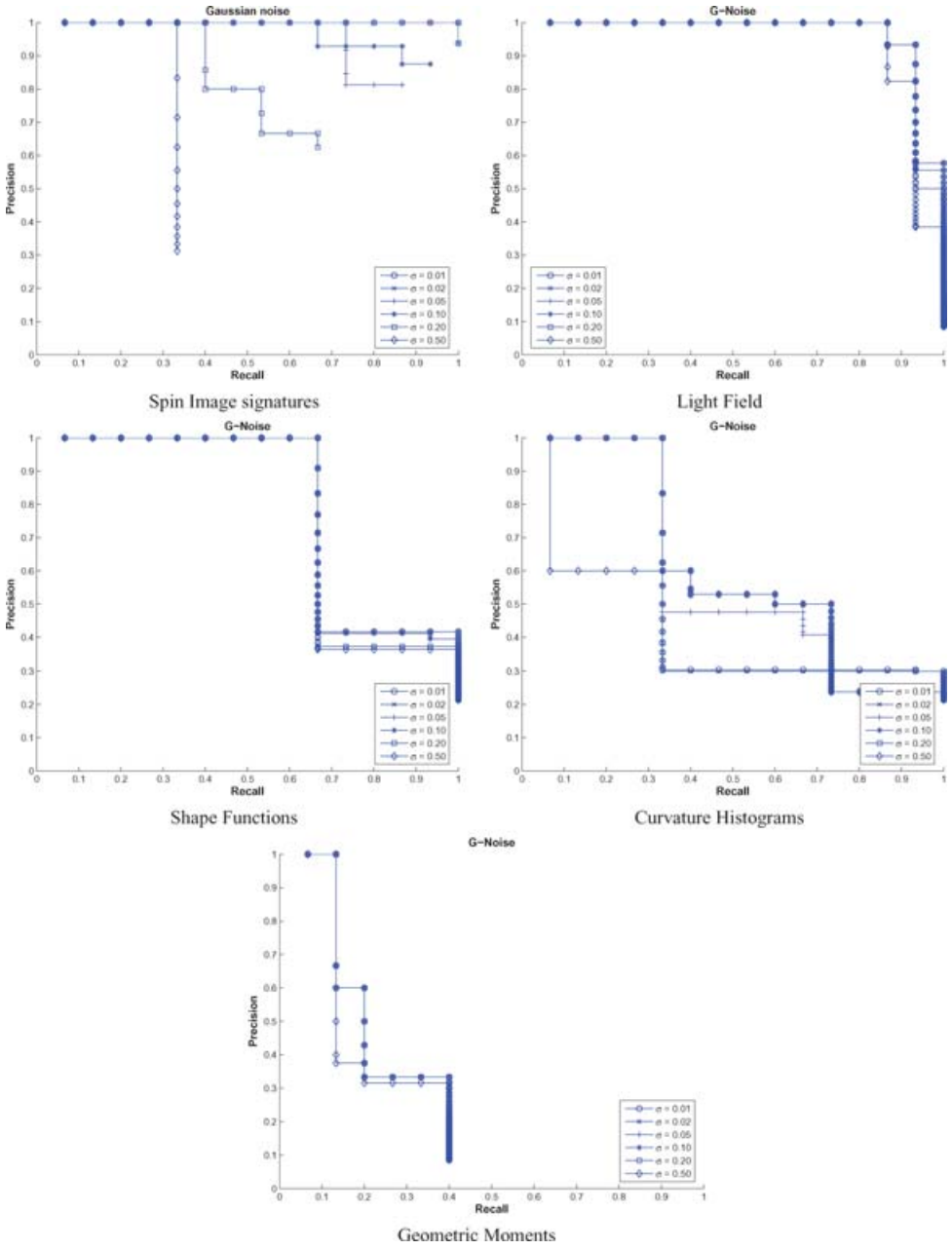

Fig. 11. Robustness to Gaussian noise modifier for Spin Image signatures, Light Field, Shape Functions, Curvature Histograms, and Geometric Moments representations.

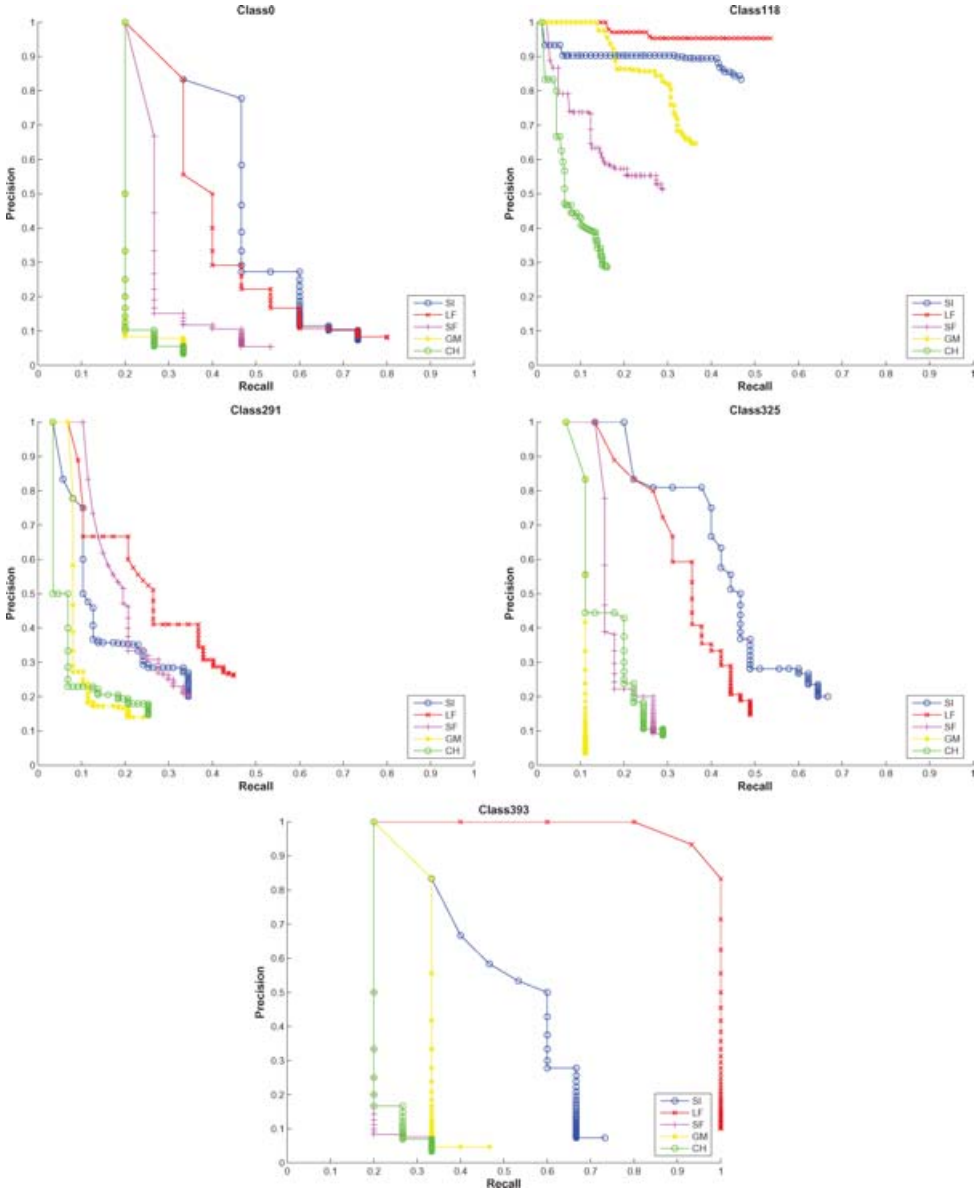

Fig. 12. Capability of each representation to capture structural peculiarities for different classes of models. Representations: SI = Spin Image signatures, LF = Light Field, SF = Shape Functions—D1, GM = Geometric Moments, CH = Curvature Histograms; Object classes: Class0 = Ant, Class118 = Human, Class291 = Face, Class325 = Hand, Class393 = Gazebo.

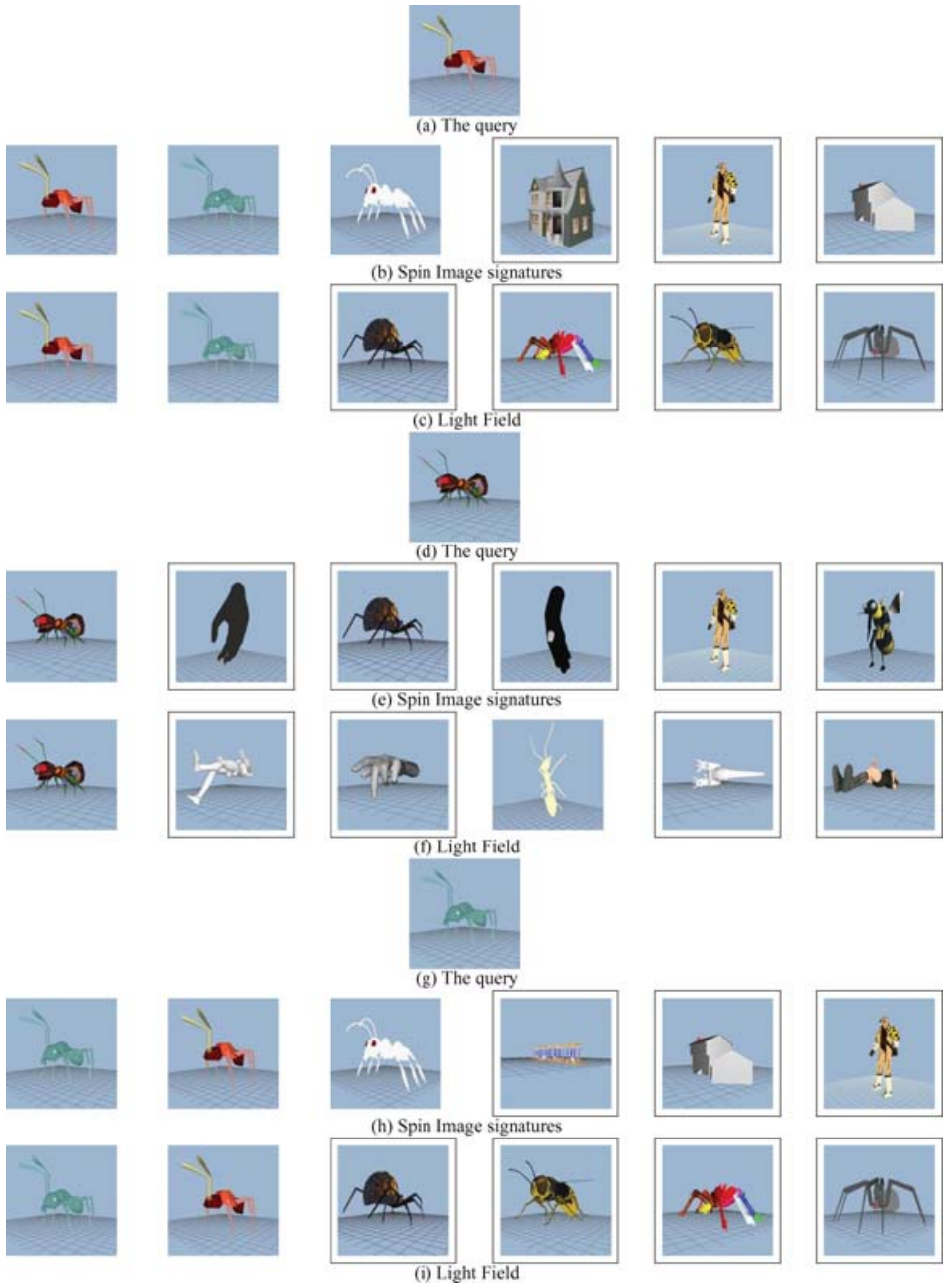

Fig. 13. Examples of retrieval by similarity for Class0 (Ant) models for the top performing representations (Spin Image signatures and Light Field). The first retrieved model is the query model. Nonrelevant models retrieved are highlighted.

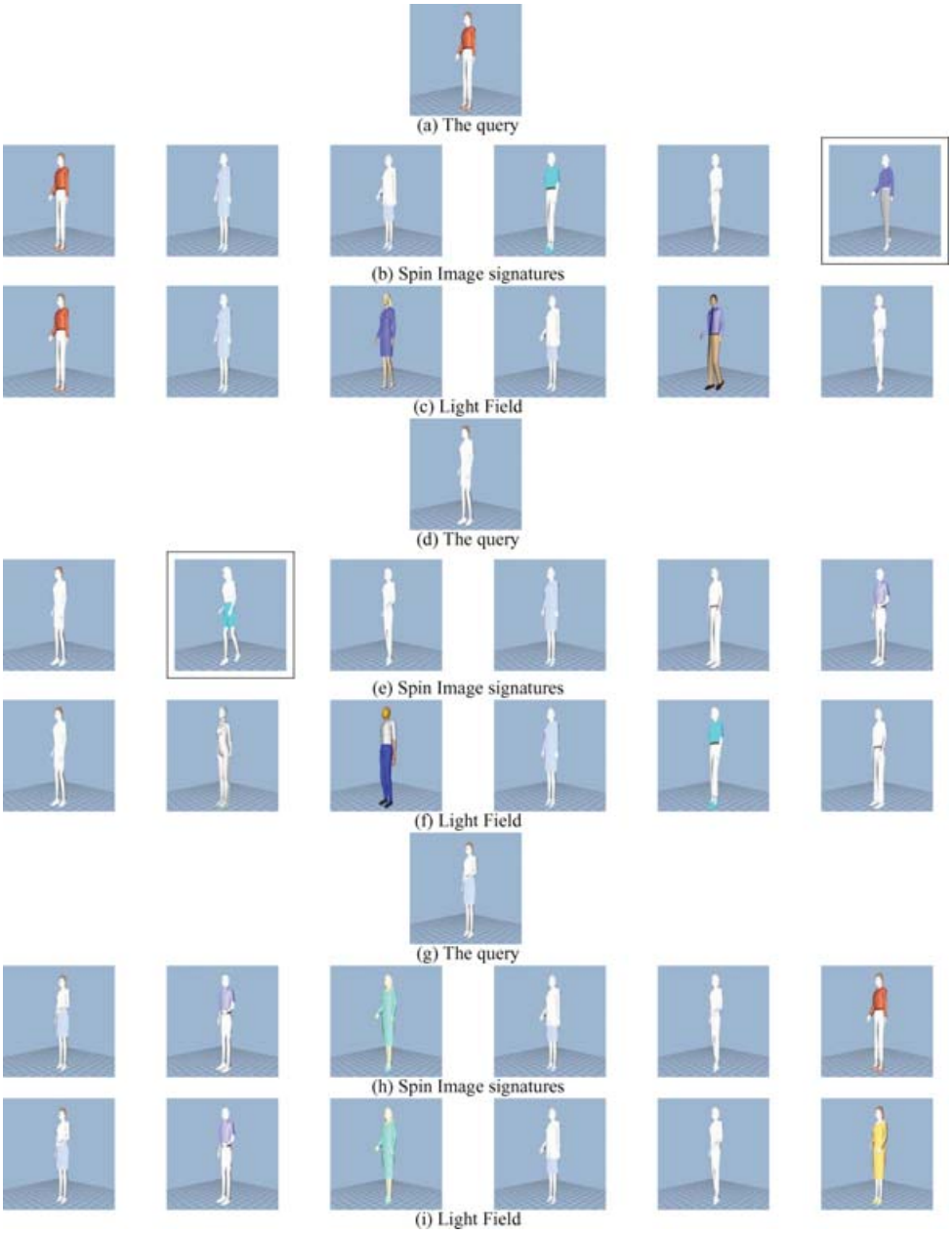

Fig. 14. Examples of retrieval by similarity for Class118 (Human) models for the top performing representations (Spin Image signatures and Light Field). The first retrieved model is the query model. Nonrelevant models retrieved are highlighted.

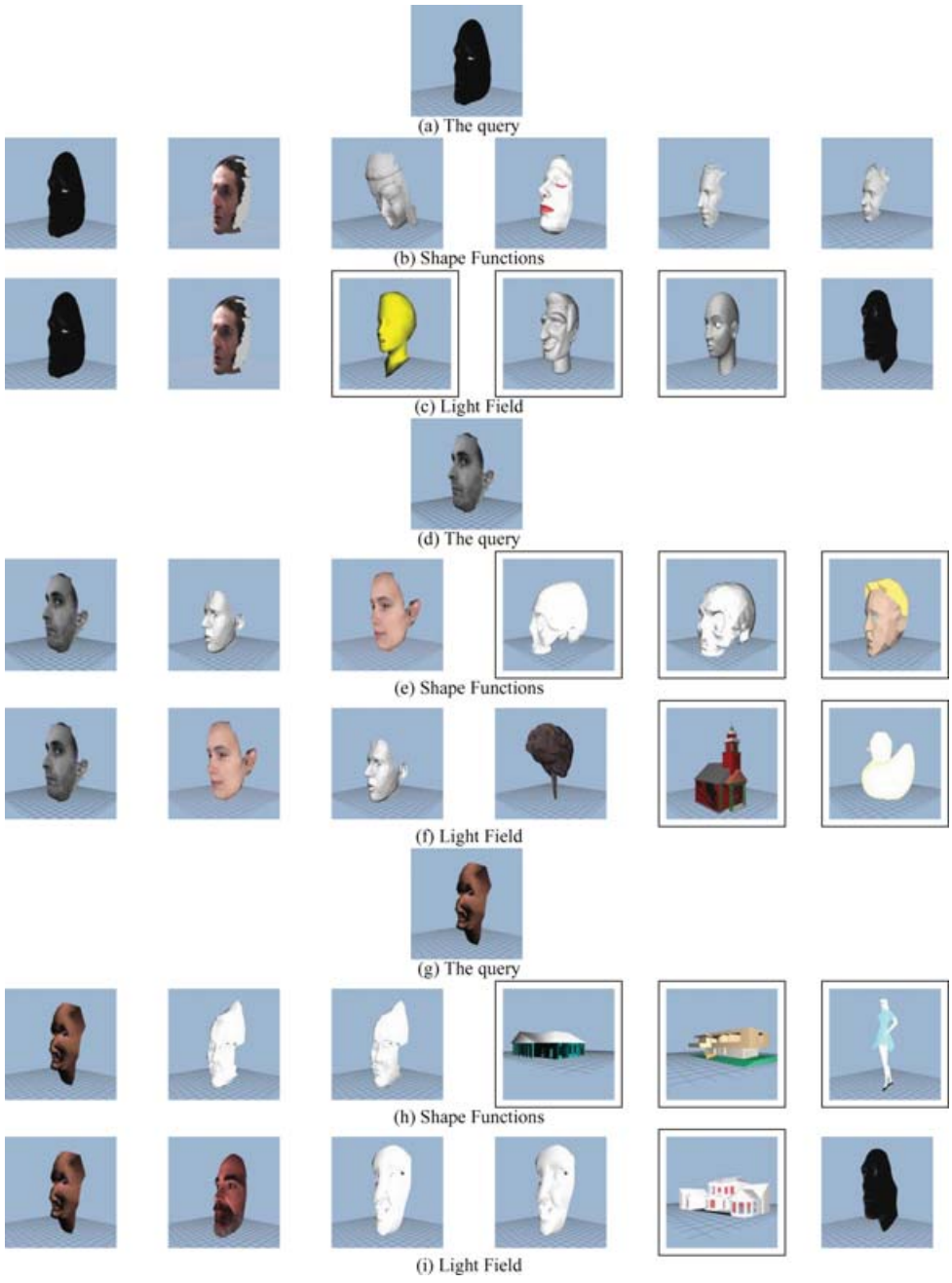

Fig. 15. Examples of retrieval by similarity for Class291 (Face)models for the top performing representations (Light Fields and Shape Functions). The first retrieved model is the query model. Nonrelevant models retrieved are highlighted.

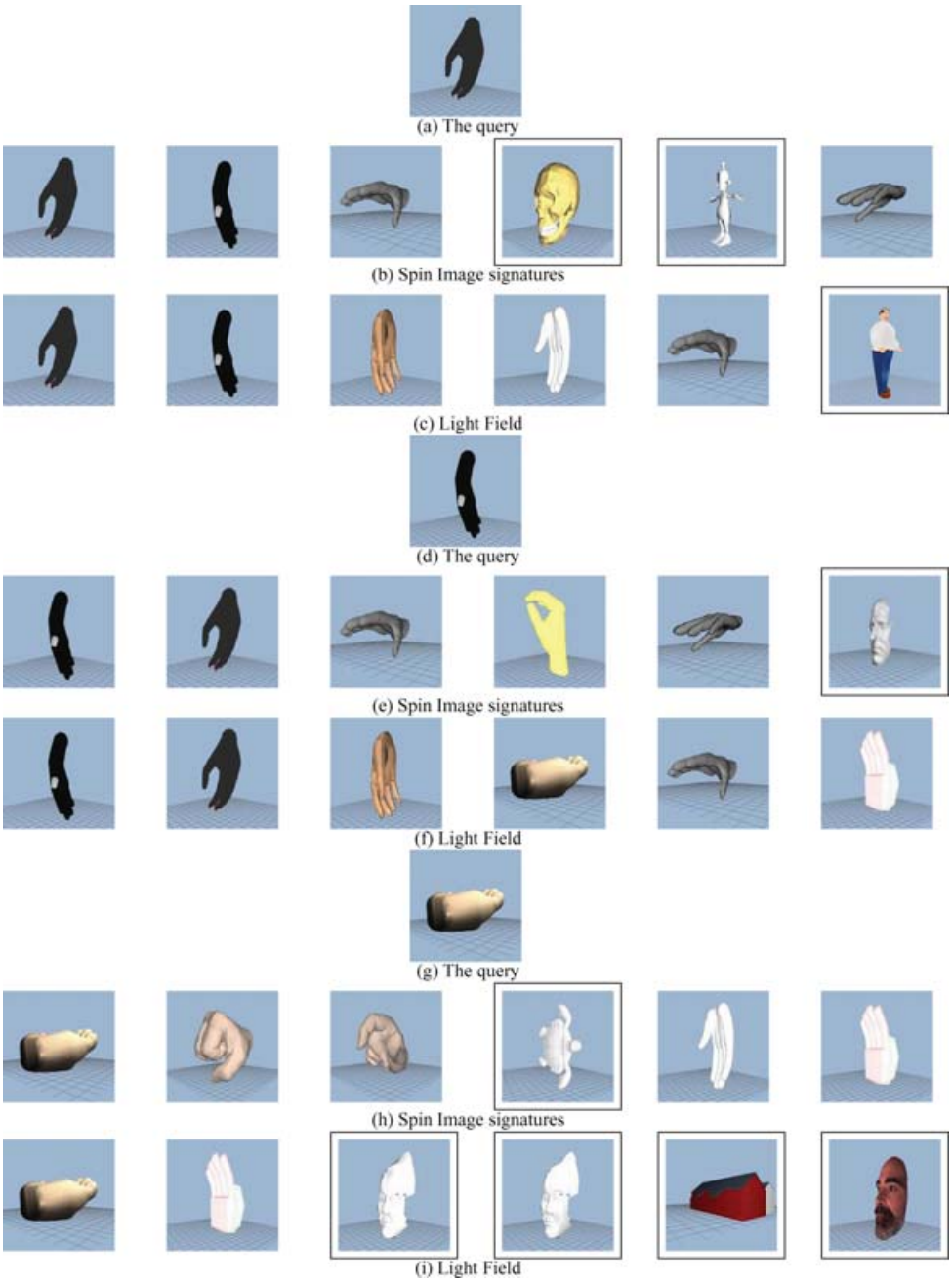

Fig. 16. Examples of retrieval by similarity for Class325 (Hand) models for the top performing representations (Spin Image signatures and Shape Functions). The first retrieved model is the query model. Nonrelevant models retrieved are highlighted.

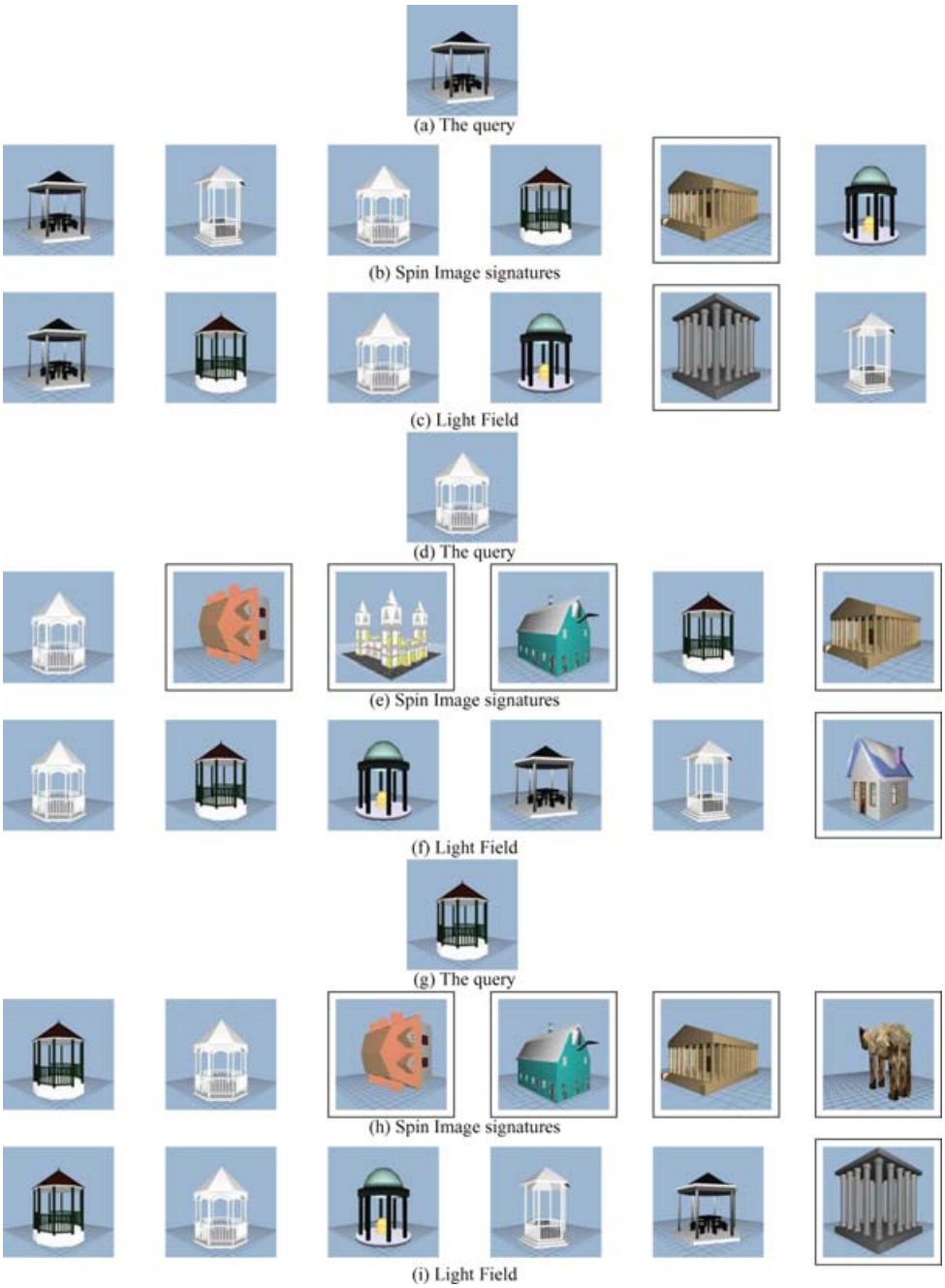

Fig. 17. Examples of retrieval by similarity for Class393 (Gazebo). models for the top performing representations (Spin Image signatures and Light Field). The first retrieved model is the query model. Nonrelevant models retrieved are highlighted.

appear to be superior to the other solutions. However, Spin Image signatures and Light Field have very different performance for models with different resolutions, the former showing superior performance with high-resolution 3D models, the latter performing better for low-resolution models.

REFERENCES

ANTINI, G., BERRETTI, S., DEL BIMBO, A., AND PALA, P. 2005. Retrieval of 3D objects using curvature correlograms. In *Proceedings of the International Conference on Multimedia and Expo (ICME'05)*. (July), Amsterdam, The Netherlands.

ASSFALG, J., DEL BIMBO, A., AND PALA, P. 2003. Curvature maps for 3D CBR. In *Proceedings of the International Conference on Multimedia and Expo (ICME'03)*. (July), Baltimore, Maniland.

ASSFALG, J., DEL BIMBO, A., AND PALA, P. 2004. Spin images for retrieval of 3D objects by local and global similarity. In *Proceedings of the 17th International Conference on Pattern Recognition (ICPR-04)*. (Aug.), Cambridge, UK. 23–26.

BELYAEV, A. G., BOGAEVSKI, I. A., AND OHTAKE, Y. 2000. Polyhedral surface smoothing with simultaneous mesh regularization. In *Proceedings of Geometric Modeling and Processing (Theory and Applications)*. (Apr.), Hong Kong, China. 229–237.

BERRETTI, S., DEL BIMBO, A., AND VICARIO, E. 2001. Efficient matching and indexing of graph models in content-based retrieval. *IEEE Trans. Patt. Analy. Machine Intelli.*, *23*, 10 (Oct.), 1089–1105.

BESL, P. J. AND JAIN, R. C. 1985. Three-dimensional object recognition. *Comput. Surv.*, *17*, 1, 75–145.

BEZDEK, J. C., KELLER, J., KRISHNAPURAM, R., AND PAL, N. R. 1999. *Fuzzy Models and Algorithms for Pattern Recognition and Image Processing*. Kluwer Academic Publisher, Boston, MA.

CAMPBELL, R. J. AND FLYNN, J. 2001. A survey of free form object representation and recognition techniques. *Comput. Vision Image Understand. 81*, 2, 166–210.

CHEN, D. Y., TIAN, X. P., SHEN, Y. T., AND OUHYOUNG, M. 2003. On visual similarity based 3D model retrieval. In *Proceedings of Eurographics Computer Graphics Forum* (EG'03) *22*, 3.

COLOMBO, C., DEL BIMBO, A., AND PERNICI, F. 2005. Metric 3D reconstruction and texture acquisition of surfaces of revolution from a single uncalibrated view. *IEEE Trans. Patt. Analy. Machine Intell. 27*, 1, 99–114.

DEL BIMBO, A. AND VICARIO, E. 1998. Using weighted spatial relationships in retrieval by visual contents. In *Proceedings of the IEEE Workshop on Content-Based Access of Image and Video Libraries (CBAIVL'98)* (June). Santa Barbara, CA. 35–39.

DESBRUN, M., MEYER, M., SCHRODER, P., AND BARR, A. H. 2000. Discrete differential-geometry operators in *n*D. Caltech.

ELAD, M., TAL, A., AND AR, S. 2001. Content based retrieval of VRML objects—An iterative and interactive approach. *EG Multimedia* (Sept.). 97–108.

ESHERA, M. A. AND FU, K.-S. 1984. A graph measure for image analysis. *IEEE Trans. Syst. Man Cybern.*, *14*, 3, (May/June), 398–0407.

GAREY, M. R. AND JOHNSON, D. 1979. *Computer and Intractability: A Guide to the Theory of NP-Completeness*. Freeman, San Francisco.

GARLAND, M. 1999. Multiresolution modeling: Survey & future opportunities. In *Proceedings of Eurographics'99* (Sept.).

GOSHTASBY, A. 1985. Description and descrimination of planar shapes using shape matrices. *IEEE Trans. PAMI*. *7*, 738–743.

JOHNSON, A. E. AND HEBERT, M. 1999. Using spin-images for efficient multiple model recognition in cluttered 3-D scenes. *IEEE Trans. Patt. Analy. Machine Intelli*. *21*, 5, 433–449.

KIM, D.-J., PARK, Y.-W., AND PARK, D.-J. 2001. A novel validity index for determination of the optimal number of clusters. *IEICE Trans. Inform. Syst.*, *E84-D*, 2, (Feb.), 281–285.

KOLONIAS, I., TZOVARAS, D., MALASSIOTIS, S., AND STRINTZIS, M. G. 2001. Content-based similarity search of VRML models using shape descriptors. In *Proceedings of the International Workshop on Content-Based Multimedia Indexing*, (Sept.), Brescia, Italy, 19–21.

KRIEGEL, H. P. AND SEIDL, T. 1998. Approximation-based similarity search for 3D surface segments. *GeoInformatica J. 2*, 2, Kluwer Academic Publisher, 113–147.

MAHMOUDI, S. AND DAOUDI, M. 2002. 3D models retrieval by using characteristic views. In *Proceedings of 16th International Conference on Pattern Recognition 2*, (Aug), 457–460.

MOKHTARIAN, F., KHALILI, N., AND YEUN, P. 2001. Multi-scale free-form 3D object recognition using 3D models. *Image Vision Comput. 19*, 5, 271–281.

NOVOTNI, M. AND KLEIN, R. 2003. 3D Zernike descriptors for content based shape retrieval. In *Solid Modeling* 2003.

OHBUCHI, R., MINAMITANI, T., AND TAKEI, T. 2003a. Shape similarity search of 3D models by using enhanced shape functions. In *Theory and practice of Computer Graphics*, *2003*. 97–104.

OHBUCHI, R., NAKAZAWA, M., AND TAKEI, T. 2003b. Retrieving 3D shapes based on their appearance. In *Proceedings of MIR'03* (Nov.). Berkeley, CA. 39–46.

OSADA, R., FUNKHOUSER, T., CHAZELLE, B., AND DOBKIN, D. 2002. Shape distributions. *ACM Transa. Graph.*, *21*, 4 (Oct.), 807–832.

PAQUET, E. AND RIOUX, M. 1999. Nefertiti: A query by content system for three-dimensional model and image database management. *Image Vision Comput.*, *17*, 2, 157–166.

SHILANE, P., MIN, P., KAZHDAN, M., AND FUNKHOUSER, T. 2004. The Princeton shape benchmark. In *Proceedings of Shape Modeling International*. (June), Genova, Italy.

RÖSSL, C., KOBBELT, L., AND SEIDEL, H. P. 2000. Extraction of feature lines on triangulated surfaces using morphological operators. In *Smart Graphics*, *Proceedings of the 2000 AAAI Symposium*.

SNYDER, J. P. AND BUGAYEVSKI, L. M. 1995. *Map Projections—A Reference Manual*. Taylor & Francis.

TANGELDER, J. W. H. AND VELTKAMP, R. C. 2004. A survey of content based 3D shape retrieval methods. In *Proceedings of Shape Modeling International*. 145–156.

TAUBIN, G. 1995. A signal processing approach to fair surface design. *Computer Graphics (Annual Conference Series)*, *29*, 351–358.

THE DIGITAL MICHELANGELO PROJECT. http://graphics.stanford.edu/data/mich/.

THE NATIONAL DESIGN REPOSITORY. http://edge.mcs.drexel.edu/repository/frameset.html.

THE PROTEIN DATA BANK. http://www.rcsb.org/pdb.

VANDEBORRE, J.-PH., COUILLET, V., AND DAOUDI, M. 2002. A practical approach for 3D model indexing by combining local and global invariants. In *Proceedings of the 1st International Symposium on 3D Data Processing, Visualization, and Transmission (3DPVT'02)*.

VRANIĆ, D. V., SAUPE, D., AND RICHTER, J. 2001. Tools for 3D-object retrieval: Karhunen-Loeve-transform and spherical harmomics. In *Proceedings of the IEEE Workshop on Multimedia Signal Processing*.

ZHANG, C. AND CHEN, T. 2001. Indexing and retrieval of 3D models aided by active learning. In *ACM Multimedia*, 615–616.

Received December 2005; accepted December 2005