

Comparative analysis of shape descriptors for 3D objects

Graciela Lara López¹ · Adriana Peña Pérez Negrón¹ ·
Angélica De Antonio Jiménez² ·
Jaime Ramírez Rodríguez² · Ricardo Imbert Paredes²

Received: 17 July 2015 / Revised: 7 December 2015 / Accepted: 2 February 2016 /
Published online: 25 February 2016
© Springer Science+Business Media New York 2016

Abstract One of the basic characteristics of an object is its shape. Several research areas in mathematics and computer science have taken an interest in object representation in both 2D images and 3D models, where shape descriptors are a powerful mechanism enabling the processes of classification, retrieval and comparison for object matching. In this paper, we present a literature survey of this broad field, including a comparative analysis based on the above shape descriptor processes. In view of their significance, we identified the shape descriptors implemented using the concept of visual salience. This paper gives an overview of this topic.

Keywords Shape descriptors · Matching and similarity · Voxelization · Pose normalization and visual salience

✉ Graciela Lara López
graciela.lara@red.cucei.udg.mx

Adriana Peña Pérez Negrón
adriana.pena@cucei.udg.mx

Angélica De Antonio Jiménez
angelica@fi.upm.es

Jaime Ramírez Rodríguez
jramirez@fi.upm.es

Ricardo Imbert Paredes
rimbert@fi.upm.es

¹ Módulo “O” División de Electrónica y Computación, CUCEI, Universidad de Guadalajara, Av. Revolución 1500, Col. Olímpica C.P. 44430, Guadalajara, Jalisco, Mexico

² Escuela Técnica Superior de Ingenieros Informático, Universidad Politécnica de Madrid, Madrid, Spain

1 Introduction

The shape of an object expresses its appearance. Shape can communicate ideas, as well as attracting the viewers' attention. Therefore, it can also represent a salient feature. Humans are capable of identifying an object from its shape alone. Technological interest in implementing this human capability has enabled the extraction of shape-semantic information, usually through a process of segmentation. However, the description of the shape of an object is still a difficult task, with a number of limitations.

A shape descriptor can be defined as a mapping from the 3D object space to some high-dimensional vector space [1]. The main goal of shape description is to get the maximum amount of shape information with the least possible dimensionality [21] in feature vectors or data structures, a suitable numerical representation [46], and the extracted numerical characteristics that describe the shape of a multimedia object. For Guo et al. [29], a good descriptor should be descriptive, compact, and robust. Vranić et al. [98] defined four important criteria for a 3D object shape descriptor:

1. Invariance with respect to translation, rotation, scaling, and reflection of the 3D object;
2. Robustness with respect to level of detail;
3. Efficient feature extraction and search; and
4. Multiresolution feature representation.

These criteria are further discussed in Section 4 as part of the comparative analysis of the different methods proposed for shape descriptor.

Shape descriptors of 3D model are a helpful mechanism for classifying, retrieving, clustering, matching and establishing similarities between objects. They play an important role in different areas such as computer-aided design (CAD), computer-aided manufacturing (CAM), virtual reality, entertainment, medicine, molecular biology, physical simulation, and e-commerce [1, 15, 21, 22, 46, 79, 107]:

- In CAD/CAM, shape descriptors are especially applicable in physical anthropology; which plays an important role in industrial design. For example, clothing design or ergonomics [73], as well as in the matching of solid models of 3D mechanical parts [21]. Likewise, local features can significantly improve the manufacturing costs, manufacturing process selection, production and functional parameters of 3D objects in the CAD tool field [11], as well as finding similarities, useful in furniture design [86] and image reconstruction [60].
- For virtual reality and entertainment, the use of 3D models improves the realism in film and video game production. In this industry, 3D objects can be reused and adapted based on their similarity to reduce production costs [15].
- In medicine, object similarities are useful to detect organs deformations. For example, they have been used in a specific part of the brain called the hippocampus to help diagnose diseases like epilepsy [43].
- In molecular biology, shape descriptors have been applied to analyze molecular surfaces [4, 95] and molecular stability [19].
- In physical simulation, Novotni et al. [63] applied shape descriptors to find the best-fitting shoe for a given 3D foot scan.

- In e-commerce, a customer can start with a few typical style options and then use the search engine to retrieve similar styles, in furniture shopping [73].

Shape descriptor analysis has also been used for computer vision and texture analysis [49, 75] in order to represent articulated objects [58] or to compute the similarity between object deformations [87]. In aerial images to distinguish and categorize areas such as parking, residential or schools [106, 109]; and fine-grained image recognition for example for insects [105].

This paper is structured as follows. Section 2 contains a review of shape representation methods, and a general description of six different taxonomies. Section 3 describes nine shape descriptor categories, extracted from the aforementioned taxonomies and outlines 58 shape descriptors. Section 4 compares these shape descriptors, analyzes their frequencies and the percentages of their main features. Finally, in Section 5 conclusions are presented.

2 Methods and taxonomies

This overview covers related work on shape descriptor classification for 2D images and 3D models. First it addresses the most common object shape representation methods, followed by the best-known taxonomies for 3D shape descriptors based on the characteristics of the respective shape.

2.1 Methods for object shape representation

Zhang D. et al. [108] divided the methods for representing the shape of an object into two categories: contour-based methods, which represent objects/shapes as a whole, and region-based methods, which represent segments/sections. Both categories have two subdivisions: structural and global methods. They include a large set of shape description techniques which are described later: chain code, polygon, B-spline, invariants, perimeter, compactness, eccentricity, shape signature, Hausdorff distance, Fourier descriptors, wavelet descriptors, scale space, autoregressive, elastic matching, area, Euler number, geometric moments, Zernike moments, pseudo-Zernike moments, Legendre moments, grid method, shape matrix, convex hull, media axis, and core. Ling et al. [52] pointed out two methods for object shape representation: one based on the extraction of local features and the other based on the extraction of global features (see Sections 3.5.1 and 3.5.2). Tangelder et al. [89] also introduced a set of shape representation methods based on the volume and surface of the 3D models, identified as implicit surfaces, constructive solid geometry (CSG), binary space partitioning (BSP) trees, octrees, boundary representation (B-rep) and free-form surfaces.

These methods were organized into different categories of shape descriptors, which were then integrated into the different taxonomies proposed. There follows a description of the six different taxonomies proposed for different purposes.

2.2 Taxonomies for shape descriptors

We concluded from the assortment of approaches that there is no universally accepted method for building a shape descriptor taxonomy. Furthermore, these taxonomies serve different purposes.

Zhang L. et al. [107] proposed a classification divided into three categories: (1) feature-based shape descriptors, (2) graph-based descriptors, and (3) other methods. Their taxonomy is based on the most popular shape descriptors for 3D object classification and retrieval. In particular, this taxonomy considers the spatial partition and the representation of the features of the 3D models, where the 3D shape can be discriminated by its geometric features and topological properties. Zhang L. et al. [107] discriminated the shape by measuring and comparing its features. Furthermore, their taxonomy targets the design of 3D object space methods to keep all possible information of an object in a low-dimensional vector. Five sub-categories were compared based on the following criteria: original shape features, spatial partition methods, pose normalization (see Section 4.6 for a detailed explanation), transformation invariance, and advantages or disadvantages.

Bustos et al. [15] divided the shape descriptors into five categories: (1) statistics, (2) extension-based, (3) volume-based, (4) surface geometry, and (5) image-based methods. This taxonomy targets the retrieval of shapes across 3D objects. They also conducted a qualitative comparison of some of the proposed shape descriptors based on the technical description published in the literature using the following criteria: dimension, invariance, object representation, object consistency and metric (the measure of similarity).

In his doctoral thesis, Akgül [1] presented a taxonomy for shape descriptors divided into five categories: (1) histogram-based, (2) transform-based, (3) graph-based, (4) 2D image-based and (5) other methods. He focused on a general approach based on the geometric or topological information contained in the 3D object. He also considered similarity for object retrieval, and compared the retrieval performance resulting from the fusion of two descriptors against other well-known 3D shape descriptors.

Also in his doctoral thesis, Dos Santos [21] presented a five-category shape descriptor classification: (1) histogram-based, (2) transform-based, (3) graph-based, (4) 2D image-based and (5) other methods. This taxonomy is very similar to the proposed by Bustos et al. [15], although Dos Santos compared the behavior of some shape descriptors in order to identify the most suitable options for his thesis. He developed a prototype to compute 3D shape descriptors and evaluate shape-matching performance based on these descriptors and some of their combinations, rating the accuracy and general performance of the 3D models retrieved in queries.

Tangelder et al.'s [89] taxonomy is organized into three main groups: (1) feature-based, (2) graph-based, and (3) geometry-based methods. This taxonomy focuses on the use of matching methods for content retrieval based on the 3D shape considering the surface and volume of the 3D models. They compared the matching methods according to the following criteria: shape model, triangle inequality, efficiency, discriminative power, partial matching, robustness (see Section 4.5), and pose normalization requirement (see Section 4.6).

EINAghy et al. [23] proposed a taxonomy again with five categories, as follows: (1) view-based, (2) graph-based, (3) geometry-based, (4) statistics-based and (5) general methods. Like the above four, the purpose of this taxonomy is to recover 3D objects based on similarities. Their comparison was based on 3D object representation requirements, efficiency, discriminative power, partial matching, robustness and sensitivity, and pose normalization. This comparison is very similar to the one proposed by Tangelder et al. [89].

Table 1 presents the nine descriptor categories found in these six taxonomies [1, 15, 21, 23, 89, 107]. The categories are: histogram-based, transform-based, graph-based, 2D image-based, feature-based, geometry-based, extension-based, volume-based, and other methods. The histogram-based, 2D image-based, geometry-based, and other methods categories are similar to

Table 1 Categories of shape descriptors

Category	References
(**) Histogram-Based	[1, 15, 21, 23]
Transform-Based	[1, 21]
Graph-Based	[1, 21, 23, 89, 107]
(**) 2D Image-Based	[1, 15, 21, 23]
Feature-Based	[89, 107]
(**) Geometry-Based	[23, 89]
Extension-Based	[15]
Volume-Based	[15]
(**) Other Methods	[1, 21, 23, 107]

the statistics-based, view-based, surface geometry and general methods, respectively. Therefore, they have been grouped into just four categories and are highlighted in Tables 1 and 2 with two asterisks between parentheses (**).

Note that some descriptors included in these taxonomies use the transformation of a 3D object into a set of 2D images.

Table 2 groups the shape descriptors by category and author. The totals for shape descriptors by author and year of publication are as follows: Zhang L. et al. [107] with 10 descriptors, Bustos et al. [15] with 26, Akgül [1] with 21, Dos Santos [21] with 30, Tangelder et al. [89] with 10, and EINaghy et al. [23] with 23 descriptors.

We found some shape descriptors that were placed into different categories by different authors. They are labeled with an asterisk between parentheses (*). For example, some researchers classified the shape spectrum descriptor in the histogram-based category and others in the geometry-based category. Although this descriptor might be placed in either category, we put it in the histogram-based category because the descriptors of this category use feature counters, and the shape spectrum descriptor represents the shape of an object as a set of quantitative characteristics based on the geometric information of the object. We placed the descriptors marked with (*) in the category that we think best fits their characteristics.

Section 3 gives a general description of these categories and their associated descriptors.

3 Descriptor categories and shape descriptors

Descriptors were classified in categories according to their computer structure or computer techniques such as histogram-based, graph-based or image-based techniques.

3.1 Histogram-based descriptors

This category includes all the shape descriptors that adopt a histogram, even if they are not used in a rigorous statistical sense. With regard to shape descriptors, the histogram is typically an accumulator or container that collects the numerical values of certain features calculated from the shape representation [1, 21]; it maintains the neighboring points or their properties [9]. In this category, descriptors are partitions of certain spaces in a 3D model, where the complete space is decomposed into disjoint cells that correspond to the histogram bins [36].

Table 2 Descriptors grouped by category

Taxonomy	Descriptor Name	Zhang L. et al. [107]	Bustos et al. [15]	Akgül [1]	Dos Santos [21]	Tangelder et al. [89]	EINaghy et al. [23]
3.1 (**)	Cord and Angle	x	√	√	√	x	x
Histogram-Based	Histograms						
	Color Distribution	x	x	x	√	x	x
	Curvature Histogram	x	x	x	√	x	x
	Shape Distributions	x	√	√	√	x	x
	Modified Shape Distributions	x	x	x	√	x	x
	(*) Shape Histograms	x	x	√	√	x	√
	3D Shape Contexts	x	x	x	√	x	√
	(*) Shape Spectrum	x	x	√	√	x	x
	(*) Complex Extended Gaussian Images	x	x	√	√	x	x
	Density-Based 3D Shape	x	x	x	√	x	x
	3D Hough Transform	x	x	√	√	x	x
	(*) Generalized Shape Distributions	x	x	√	x	x	√
	Geometric Hashing	x	x	x	x	x	√
	(*) Spatial Maps	x	x	x	x	x	√
	Simple Statistics	x	√	x	x	x	x
	Parameterized Statistics	x	√	x	x	x	x
	(*) Geometric 3D moments	x	√	x	x	x	x
3.2	(3D) Angular Radial	x	x	√	√	x	x
Transform-Based	Transform						
	(*)Voxel-3D Fourier Transform	x	x	√	√	x	x
	Spherical Harmonics Transform	x	x	√	√	x	x
	Planar-Reflective Symmetry Transform	x	x	x	√	x	x
	(*) Rotation Invariant Spherical Harmonics	x	x	√	√	x	x
	Distance Transform and Radial Cosine Transform	x	x	√	√	x	x
	Spherical Wavelet Transform	x	x	√	x	x	x
3.3 Graph-Based	Reeb Graphs	√	x	x	√	√	√
	Multiresolution Reeb Graphs	x	x	√	√	x	x
	Size Graphs	x	x	x	√	x	
	Skeleton(al) Graphs	√	x	√	√	√	√
	Spectral Graph	x	x	x	x	x	√

Table 2 (continued)

Taxonomy	Descriptor Name	Zhang L. et al. [107]	Bustos et al. [15]	Akgül [1]	Dos Santos [21]	Tangelder et al. [89]	EINaghy et al. [23]
3.4(**) 2D Image-Based	B-Rep Graph	x	x	x	x	x	√
	Model Graph	√	x	x	x	√	x
	Silhouette Descriptor	x	√	√	√	x	√
	Depth Buffer Descriptor	x	√	√	√	x	√
	Lightfield Descriptor	x	√	√	√	x	√
	Elevation Descriptor	x	x	x	√	x	x
	(*) Spin Images	x	x	x	√	x	√
3.5 Feature Based	SIFT	x	x	x	x	x	√
	Global Features	√	x	x	x	√	x
	(*) Spatial Maps	√	x	x	x	√	x
	Local Features	√	x	x	x	√	x
3.6 (**) Geometry Based	(*) Distribution based	√	x	x	x	x	x
	Volumetric Error	x	x	x	x	√	√
	(*) Weighted Point Sets	x	x	x	x	√	√
	(*) Geometrical 3D Moments	x	x	x	x	x	√
	(*) Shape Spectrum	x	√	x	x	x	√
	(*) Extended Gaussian Image	x	√	x	x	x	√
	Canonical 3D Morphing	x	x	x	x	x	√
	Canonical 3D Hough Transform	x	√	x	x	x	x
	Heat Kernel Signatures	x	x	x	x	x	√
	View Based	x	x	x	x	√	x
Deformation Based	x	x	x	x	√	x	
3.7 Extension Based	Surface Normal Directions	x	√	x	x	x	x
	Ray Based	x	√	x	x	x	x
	(Complex) (Shading)	x	√	x	x	x	x
3.8 Volume Based	Discretized model volume	x	√	x	x	x	x
	(*) Shape Histograms	x	√	x	x	x	x
	Rotation Invariant Point Cloud	x	√	x	x	x	x
	(*) Model Voxelization	x	√	x	x	x	x
	Voxelized Volume	x	√	x	x	x	x
	(*) Rotation Invariant Spherical Harmonics	x	√	x	x	x	x
(*) Reflective Symmetry	x	√	x	x	x	x	

Table 2 (continued)

Taxonomy	Descriptor Name	Zhang L. et al. [107]	Bustos et al. [15]	Akgül [1]	Dos Santos [21]	Tangelder et al. [89]	EINaghy et al. [23]
3.9 (**) Other Methods	(*) Weighted Point Sets	x	√	x	x	x	x
	(*) Spin Images	x	√	√	x	x	x
	3D Zernike Moments	√	x	√	√	x	x
	Spherical Moments	x	x	x	√	x	x
	(*) Reflective Symmetry	x	x	√	√	x	x
	(*) Generalized Shape Distributions	x	x	x	√	x	x
	Relevance Feedback	x	x	x	x	x	√
	Bag of Features	x	x	x	x	x	√
	(*) EGI (Extended Gaussian Image)	√	x	x	x	x	x
	(*) CEGI (Complex Extended Gaussian Image)	√	x	x	x	x	x
	(*) Skeleton Based	x	√	√	x	x	x
Topological matching	x	√	x	x	x	x	
Total descriptors by author		10	26	21	30	10	23

These shape descriptors describe the distribution of points on the model across all rays from the origin [42]. Histogram-based descriptors have been widely used in computer vision tasks, such as matching, image retrieval [57, 59, 62], and texture analysis [49, 75].

3.1.1 Cord and angle histograms

These descriptors use information about the spatial extension and orientation of the 3D object. The descriptor is calculated from three histograms: the first histogram represents the distribution of the angles between the cords and the first reference axis; the second histogram denotes the distribution of the angles between the cords and the second reference axis; and the third histogram offers the distribution of the radius. In this type of descriptor, the global features are used to characterize the general shape of the objects. This type of descriptor is simple to apply because it is not very “picky” about the details of the objects [70, 73].

3.1.2 Color distribution

This shape descriptor uses a voxelized (see Section 4.7) representation of the 3D object, where each voxel contains a value associated with a color. This color value is calculated with the information of the texture map of the 3D object, the properties of its materials and the extracted color of the vertices. The use and location of color are important to get the histogram that describes the object based on its colors, the red-green-blue triplet (R, G, B). The dominant color is determined for each triplet, and then the angle between the corresponding normal to that point and the first eigenvector are calculated. The statistical distribution of these angles is

represented as a set of three histograms according to the dominant color. Three color-dependent distributions of the normals are obtained: one for “red”, one for “green” and one for “blue”. Some researchers have proposed the use of a wavelet based on a model with six dimensions: x, y, z, R, G, B ; the six-dimensional wavelet transform is calculated to construct a histogram [70, 71].

3.1.3 Curvature histogram

This descriptor uses a curvature index as a function of the two principal curves of a surface. This offers the possibility of describing the shape of an object from a given point. However, some information on the amplitude of the shape on the surface is lost and therefore this method is noise sensitive. The surface curvatures are computed at a generic vertex v of a three-dimensional polygonal mesh [44, 92]. The curvatures can be estimated by calculating the curvature at each face of the mesh by fitting a quadric to the neighborhood of that face using the least squares method and then determining the principal curvatures k_1 and k_2 using the eigenvalues of a Weingarten endomorphism. The principal curvatures are usually defined as the eigenvalues of the Weingarten map [92].

3.1.4 Shape distributions

This shape descriptor uses the global geometric characteristics of the shape of the object: distance, angle, area and volume measurements across random points on the surface. Five functions measure the properties of an object quickly and easily: A_3, D_1, D_2, D_3 and D_4 . A_3 is the distribution function that measures the angle formed by three random points on the surface of the object. D_1 is the distribution function that measures the distance between a fixed point and one random point on the surface. D_2 is the distribution function with the distance between two random points on the surface. D_3 is the distribution function of the square root of the area of the triangle defined by three random points on surface. And D_4 is the distribution function that measures the cube root of the volume of the tetrahedron defined by four random points on the surface. These distribution functions are invariant to rigid motions, rotations and translations [68].

3.1.5 Modified shape distributions

This descriptor, proposed by Osada et al. [68], is a modified version of the D_2 descriptor. The modifications are in the angle-distance histogram (AD) and in the absolute-angle distance histogram (AAD), where a quasi-random number sequence is used instead of the pseudo-number sequence to select points. To calculate the AD histogram, the distance of a pair of points and the angles formed by their surfaces s are measured. The AAD is computed in the same way, but it does not consider the sign of the inner product; in contrast, the AD histogram respects the angle sign [66].

3.1.6 Shape histograms

This descriptor was proposed to describe 3D solid models using three techniques for the decomposition of the space: a shell model, a sector model and a spider web model, which is a combination of the first two. In the shell model, the 3D space is decomposed into concentric

shells around the center point. In the sector model, the 3D model is decomposed into sectors that emerge from the center point. The combined model presents more detailed information and has a higher dimensionality than the pure shell models and the pure sector models. The simple points on the surface of the 3D model and the extracted features are organized in histograms or frequency distributions to represent their occurrence. This descriptor approach is intuitive and discrete for complex spatial objects; although it has the shortcoming of using Euclidean distance to compare two shape histograms that require a Mahalanobis measure. Moreover, this descriptor requires planning and pose normalization (for rotation or translation, see Section 4.6) in each pre-processing stage. As part of the pre-processing, the 3D models are invariant to translation by moving the origin to the centroid [4].

3.1.7 3D shape contexts

This descriptor calculates the 3D shape using N points of reference, sampled from the shape boundary. The vectors of the shape are defined with respect to each of these points, that is, each point captures the distribution of the remaining points through relative positions to this point. The distribution of this calculation is represented as a histogram with the relative coordinates of the remaining points ($N-1$). The relative coordinates are based on the last specified point and are used when the location of a point relative to the previous point is known. This method is applied to all the sampled points, generating a descriptor that is a set of N histograms [21, 45]. The bins of the histograms are defined from the overlay of concentric shells around the centroid of the model and sectors emerging from the centroid [23].

3.1.8 Shape spectrum

This descriptor emerged from the idea of making a view based on the representation of the 3D free-form object that uses an index function [20]; the shape index is a function of two principal curvatures of a point on a regular 3D surface [44]. Although it has appealing characteristics —invariance with respect to rotation, translation and scale—, the unreliability of the curvature estimation leads to a lack of robustness. Likewise, the shape spectrum quantitatively characterizes the object shape summarizing the area on the surface of an object at each index value. The shape index shows information about the local geometrical attributes of the 3D surface expressed as the angular coordinate of the polar representation of the principal curvature vector. The shape index also provides a scale for representing salient elementary shapes such as convex, concave, rut, ridge and saddle, and is invariant with respect to scale and Euclidean transforms.

3.1.9 Complex extended gaussian images (CEGI)

Complex extended Gaussian images (CEGI) is an extended concept of the extended Gaussian images (EGI) approach (see Section 3.6.3). CEGI represents a 3D object using the information associated with its surface, which can also be used to establish an object's pose. The representation of this descriptor consists of a spatial orientation histogram in which each weight associated with a normal distance is represented as a complex number. The normal distance of the face from the selected origin is encoded as the weight phase, whereas the weight magnitude is the visible area of the face. The CEGI representation can

estimate both the orientation and translation of a detected object with respect to a stored model or a prototype [40].

3.1.10 Density-based 3D shape

This descriptor is defined as a sampled probability density function. The function extracts the shape of the 3D object from its geometric features and its local surface features. The density-based shape descriptor considers three types of local geometric multidimensional features of a point p , which are represented in the following three functions: (1) the radial shape function S_r , which is a magnitude component measuring the distance of point p to the origin and a direction component pointing to the location of point p ; (2) the tangent plane function S_t , which is a magnitude component which stands for the distance of the tangent plane at p to the origin and a direction component of the unit surface normal vector of the tangent plane; (3) the cross-product function S_c , which is decoupled into a magnitude component, which is the same as S_r , and a direction component of the cross-product between the vector representation of point p and the unit surface normal vector at p . The cross-product function is the part that encodes the interaction between the first two features [2].

3.1.11 3D Hough transform (3DHT)

This descriptor is based on accumulating points within a set of planes. Each plane is defined by a triplet, that contains the distance from the origin of the coordinate system to the plane, and the two angles of the azimuthal (an angular measure in a spherical coordinate system) and the elevation, which are associated with the spherical representation of the plane's unit length normal vector [101].

3.1.12 Generalized shape distributions (GSD)

This descriptor is a generalization of the D2-shape distribution in the form of a 3D histogram that considers local and global shape signatures. A *signature* is a simple representation of an object or a process in the form of a mathematical function, a feature vector, a geometric shape, or others, intended to uniquely capture the significant characteristics of an object [6]. The GSD method counts the number of specific local shape pairs at certain distances: two dimensions of the histogram account for pairs of shape signatures, which are simply k-means quantized versions of spin images computed on the mesh, and the third dimension records the Euclidean distances of the local shape pairs. The GSD descriptor can detect similar parts in two shapes, whereas most global shape descriptors fail to achieve this goal. The descriptor is represented as an indexing data structure to reduce space complexity [54].

3.1.13 Geometric hashing

This descriptor is characterized because it uses the basic geometric features of a 3D object, starting from a set of points or lines and their geometric relations that are encoded using minimal transformations with standard methods of analytic geometry. A set of basis points is chosen from a set of points and the coordinates of the remaining points are calculated with respect to that basis. They are then stored in a histogram for each coordinate set. The process is repeated for all possible combinations of the basis, and the final resultant histogram is stored in

a hash table which indexes all the objects [23]. This descriptor method is based not on measurements from the extracted 3D models but on the distributions of such measurements [47].

3.1.14 Spatial maps

The distinctive feature of this descriptor is that it has a set of representations that capture the spatial location of an object, that is, a set of relative positions of the object location [89]. Spatial maps are methods that are based on object space partitioning, and thus the transformation and matching methods of these descriptors are more or less selected or designed in terms of the spatial partitions. The processing methods and correlation of this descriptor are roughly selected and designed in terms of the spatial partitioning [99].

3.1.15 Simple statistics

This descriptor considers the basic characteristics of an object, including scale, the dimensions of its bounding box, volume, number of vertices, polygons, statistical moments, statistical data about distances of pairs of points randomly selected from the 3D object, and the composition information about a visual object, like its position and orientation. The position information is encoded with the measured normalized coordinates of the geometric center of the object's bounding box. The orientation information is encoded with the information of the angle(s) formed by the major axis with the vertical and horizontal axis [73]. The information of this type of descriptor is commonly summarized as histograms.

3.1.16 Parameterized statistics

This descriptor proposes a feature vector based on three units of statistical measures obtained from nine axes. The method normalizes the location and orientation of the model using the model's center of mass and principal axes of inertia. The statistical measures are previously parameterized. The three statistical measures are represented with three histograms that integrate the model's shape descriptor. The measures are calculated considering: (1) the moment of inertia about the axis; (2) the average distance of the surfaces from the axis; and (3) the variance of the distance to the surfaces from the axis. The Euclidean distance and the elastic-matching distance are used to measure the distance between pairs of feature vectors [67].

3.1.17 Geometric 3D moments

Moments are used as descriptors to represent the shape of the 3D objects. They are defined as feature vectors that are especially useful in the processes of object retrieval and classification [15, 23]. The moments provide a statistical representation of an object with respect to its scalar values. However, statistical moments are not invariant with respect to translation, rotation, and scale. The number of moments influences the completeness of the object representation. This descriptor method considers two types of moments by order: (1) low-order moments that describe the most salient and basic characteristics; and (2) high-order moments that tend to describe finer structures. The order is equivalent to a low- or high-resolution representation [73]. Different forms can be used to calculate statistical moments depending on the center of

mass for all triangles with respect to the mass of the triangle [72], or from object points sampled uniformly with a ray-based scheme [78], and from the centers of mass (centroids) of all object faces [72].

3.2 Transform-based descriptors

These descriptors capture the surface points on a 3D voxel or spherical grid by means of a scalar-valued function which is processed by transformation tools such as the 3D Fourier transform, angular radial transform, spherical trace transform, spherical harmonics or wavelets [3, 110]. An important advantage of the transform-based methods is descriptor compaction due to the fact that the feature vector retains only few transform coefficients. Furthermore, these descriptors consider the invariance that can be obtained by discarding the phase of the transform coefficients at the expense of some extra shape information [110]. In other words, the shapes are described in a transformation invariant manner, so that any transformation of a shape will be described in the same way. This yields the best measure of similarity in any transformation [42].

3.2.1 (3D) angular radial transform (3D ART)

The angular radial transform (3D ART) descriptor is defined as a moment-based image description method that represents pixel distribution within a 2D region. Based on this idea, the 3D ART transform descriptor is described as a complex unitary transform defined on a unit sphere. In this sense, the objects have to be represented within a spherical coordinate system which can be set up during the preprocessing stage. It is claimed that the unspecified rotations of the descriptor cannot be expressed as the sum of constant values on the angular components that can modify the values of the descriptor. Accordingly, if the descriptor considers the rotation around the z -axis, the values of the 3D ART coefficient do not change. The 3D ART descriptor shares the properties of the original 2D descriptor, such as robustness to rotation, noise and scaling. With the 3D ART descriptor, it is possible to generate compact descriptors in shorter times. This approach outperforms the spherical harmonics descriptor in terms of speed and comes close on accuracy [77].

3.2.2 Voxel-3D fourier transform

Also referred to by Bustos et al. [2005] as model voxelization, this descriptor is known as 3D DFT. This descriptor was implemented considering a process of pose normalization, followed by a process of voxelization to an object using the so-called bounding cube3 (BC). The process of pose normalization allows the BC to be subdivided into N equal-sized cubes. The voxelization process consists on subdividing the BC into N^3 equal-sized cubes (cells) and calculating the proportion of the overall surface area of the objects inside cube. Subsequently, a 3D DFT is applied to the voxelized model to represent the feature in the frequency domain [97]. The actual voxel data can be used as a 3D shape descriptor [1, 107].

3.2.3 Spherical harmonics transform

Huang et al. [36] specified that spherical harmonics represent the volume of an object for a set of spherical basis functions. The descriptor is obtained by measuring the energy contained in

different frequency bands according to the following procedure: first take the volume of an object and divide it into a set of concentric layers; second compute the frequency decomposition in each shell directly from the mesh surface; and third concatenate the norm for each frequency component at each radius into a 2D histogram indexed by radius and frequency. Inspired by an idea of Saupe et al. [78], Dos Santos [21] considered that a 3D model can be represented with a function on the sphere. To do this, rays are emitted from object's center of mass, and, for each ray defined, the value equal to distance from the origin to the last point of intersection with the object surface is estimated. These values provide a sample of function, called spherical extent function, for a shape [21, 78, 98]. To simplify the above, a dense sample is taken and the spherical harmonics are calculated to describe the shape. The spherical harmonics are a Fourier basis on a sphere equivalent to the sine and cosine on a line or circle [21].

3.2.4 Planar-reflective symmetry transform

The planar-reflective symmetry transform (PRST) descriptor is characterized by a mapping from a scalar-valued function defined over a d -dimensional space of points to a scalar-valued function. PRST is defined over the d -dimensional space of the plane, such that the scalar value associated with every plane reflection is a measure of symmetry with respect to that plane. On the other hand, the descriptor is centered on the transformation from the space of points to the space of planes that captures a continuous measure of the symmetry of a shape with respect to all planes through its bounding volume. The use of the planar reflective symmetry transforms two geometric properties, the center of symmetry and the main symmetry axes. These properties can be used to align objects in a canonical coordinate system, sets of coordinates used to describe a physical system at any given point in time [76].

3.2.5 Rotation-invariant spherical harmonics

This descriptor is a function in terms of the amount of energy that different frequencies contain. These values do not change according to rotation, and therefore the resulting descriptor is rotation invariant. The function can be viewed as a generalization of the Fourier descriptor method to the case of a spherical function. The observation of fixed rotations and the distance of a point from the origin are used to get a rotation-invariant representation of a voxel grid. The object is intersected with a set of concentric spheres to construct a spherical function from voxel values for each sphere. Finally, the frequency decomposition of each spherical function is calculated, and the calculated norms of each frequency component at each radius are obtained. The result is a rotation-invariant representation on a 2D grid indexed by radiuses and frequency [42].

3.2.6 Distance transform and radial cosine transform

Dutađci et al. [22] compared two 3D indexation methods, supported by a 3D discrete Fourier transform (DFT) and the radial cosine transform (RCT) for the retrieval of object categories such as cats, tables, airplanes, etc. Through two (binary and continuous) representations of 3D object voxelization, they estimated a 3D discrete Fourier transform and radial cosine transform. The binary representations of the values of the voxel on the surface of the object are represented by 1 and 0. In the continuous representation, the space is filled with the help of an

inverse distance function (IDF), that is, a function of the 3D distance transform. In this last function, the object surface is filled with 1, but this value decreases when one moves away from the object surface. The normalized spectral energy (NSE), which has the property of being rotation invariant, is used to calculate the 3D DFT descriptors. Thanks to this property, the 3D DFT descriptor can build a multiresolution representation of each object. Finally, the RCT coefficients constitute a set of rotation-invariant shape descriptors. The descriptor is represented as an easily obtainable feature vector.

3.2.7 Spherical wavelet transform

This descriptor is a natural extension of the spherical harmonics and the 3D Zernike moments methods that represent the image features, avoiding redundancy and information overlap between the moments. Laga et al. [46] represented a 3D object as a function sampled on the unit sphere using the azimuthal and the polar angles, respectively, in order to calculate the coefficients and construct discriminative descriptors. They apply first the spherical wavelet transform (SWT) to the spherical shape function. The shape function is implemented considering two parts—an approximation and the details—by applying low-pass and high-pass filters in the horizontal (along the azimuthal) and vertical (polar) directions. The values of approximation and the spherical wavelet coefficients are used to build the shape descriptor.

3.3 Graph-based descriptors

Graph-based descriptors aim to get a geometric meaning of the shape of a 3D object using a graph to project how the components of the shape are interconnected [89]. These descriptors are considered to be more complex and sophisticated than those based on feature vectors. Also, they have the advantage of more accurately encoding the properties of the geometric shape of the object. The descriptors use spectral graph theory tools. The information contained in a graph can be represented as numerical descriptions [1, 3]. The graph-based descriptors have the advantage of reducing the problem of shape dissimilarity through graph comparison. These descriptors are used especially for retrieving articulated objects [46].

3.3.1 Reeb graphs

Hilaga et al. [33] described a Reeb graph as a topological skeleton used under a scalar function of a 3D object and used a series of cross-sections of the object to determine nodes and arcs of the graph to represent 3D shapes. Tangelder et al. [89] defined the Reeb graph mathematically as a quotient space of a shape and a quotient function. At the same time, Reeb graphs are defined by geodesic distances, that is, the shortest distance between two points. For the implementation of Reeb graphs, each node in each Reeb graph corresponds to a connected component of the object in such a way that μ -values in components that fall within the same interval are determined by the resolution at which the graph is constructed. The relationships between parent–child nodes represent adjacent intervals of these μ -values for the contained object parts. In this case, the resulting descriptor provides invariance properties [1, 23, 33, 89, 90, 107]. The method of this shape descriptor is good for matching articulated objects; however it is sensitive to topological changes [89].

3.3.2 Multiresolution reeb graph

This descriptor is based on Reeb graphs at multiple levels of resolution using the integral geodesic distance as a function which is invariant to rotation and translation. The function is defined as the height of a point on the surface of the object or the value of the curvature of that point. This descriptor is robust to changes caused by the mesh simplification or subdivision [33, 90]. The method computes the scale-space of a shape composition represented as a rooted tree. It uses geometric features, such as volume, cord and curvature of the surface part associated with a node of the Reeb graph [90]. A disadvantage of this method is that it requires a detailed geometry of the models [21].

3.3.3 Size graphs

This descriptor, also based on the Reeb graph, builds a central skeleton of a 3D model. The method applies a size function to create a graph by size. A 3D object is associated with a size graph (\mathcal{G}, ϕ) , where \mathcal{G} is a centerline skeleton representing \mathcal{S} (\mathcal{S} is a topological space), f is a real continuous function driving the centerline extraction, and ϕ is a measuring function labeling each node of the graph with local geometrical properties of the model [12]. Mortara et al. [61] specified that the function (f) of a Reeb graph is obtained in four different ways: the extreme curvature, high curvature regions, the centroid and the topological point of view. Using these functions a centerline skeleton is retrieved from the original model; then the value of the function of each node of the skeleton is calculated to get the size graph.

3.3.4 Skeleton (al) or skeleton-based graphs

A skeletal graph is obtained from a voxelized solid object and is represented and connected as a directed acyclic graph (DAG) by applying the minimum spanning tree algorithm. This means that a set of 3D model voxels is reduced to a set of the most representative voxels. Each node of a DAG is associated with a geometric feature vector and a signature vector that encodes part of the topological information of subtrees rooted at this node. Each node of the graph contains a topological signature vector (TSV), which is used for indexing. The TSV is defined recursively over the subgraphs of the node using eigenvalues of their adjacency matrices [85]. The skeletal descriptor has the advantage that the graphs are smaller-sized topologies than B-rep graphs. Therefore, these descriptors can be used for subgraph isomorphism (graphs with the same structure) at a very low computational cost. Additionally, the features of local parts can be stored for a more precise comparison [23]. Figure 1 illustrates skeletal graphs matching.

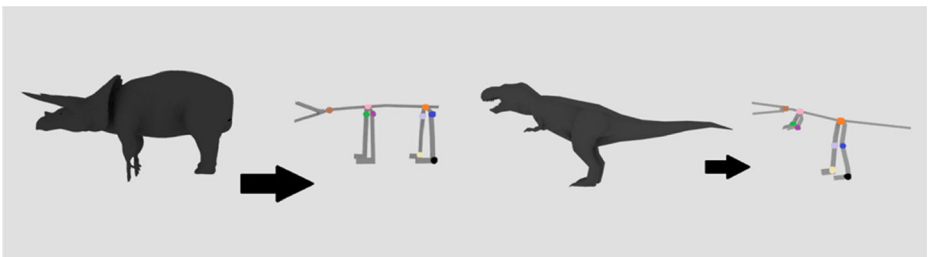


Fig. 1 Example of skeletal graphs of a pair of matching objects

accomplished using node-to-node correspondence based upon the topology, and the radial distance for the edge [85].

3.3.5 Spectral graph

A graph is a mathematical object defined by a set of vertices or nodes and a collection of edges connecting these nodes [58]. In accordance with research conducted by Chung [19], a graph spectral is based on the Laplacian matrix of the graph and correlates with the graph invariants better than the spectra of the original adjacency matrix. All matrices are built from the graph's original adjacency matrix, but include the information on each node's degree which measures the distribution of the connections in the graph [58]. The spectral graph information is encoded in the form of numerical descriptions [1].

3.3.6 B-Rep graph

The boundary representation descriptor (B-rep) is a model graph from the engineering field. B-rep represents a model with respect to its vertices, edges and faces, where the model is represented as a graph according to bounded B-spline surfaces, that is, a function expressed as a linear combination. B-splines are used as a curve fitting a numerical differentiation of a set of experimental data [25]. EINaghy et al. [23] mentioned that the set of bounding surfaces of the model is represented by the nodes of the graph, whereas the edges represent the curves intersecting the corresponding surfaces. All representations are large and complex even for simple shapes. Therefore, work is required on developing approximate algorithms based on heuristics and randomization that are often employed to determine the best match between graphs.

3.3.7 Model graph

A model graph can represent the geometry of the shape of a 3D solid object. With the help of a graph it is possible to capture the structure and the connected components of the shape of a 3D model. This representation with the information contained in each node of the graph shows up the correspondences between models. The principal advantage of this descriptor is that it can represent a 3D model with various levels of detail, which can provide local geometric information. Like the B-rep graph descriptor, this descriptor is not easy to build and has been evaluated as computationally inefficient. Particularly, the implementation of these descriptors is not easy to use with human and animal models [89].

3.4 2D image-based descriptors

This category of descriptors represents and compares the shape of a 3D object as a collection of its 2D projections taken from different viewpoints. A standard descriptor for 2D images, like Fourier descriptors and Zernike moments, is considered in order to describe each projection [46]. The descriptors of this category are designed for methods of similarity. Multiple images of a 3D object are captured from several positions with a camera and are stored in a database. The images are processed to find the similarity between the views of the query object and the models in the database [23, 36]. A particular characteristic of these shape descriptors is that they are a summary of the values of the pixels of a digital image, containing information of the

silhouette of an object. Therefore, the shape descriptor is represented with a vector that contains the number of parameters derived in this manner [48].

3.4.1 Silhouette descriptor

This descriptor represents the contour of an object or scene by its limits, and the interior silhouette of the object usually in a dark color. The descriptor gets all possible parallel projections of an object in three planes for each one of the principal axes (see Fig. 2). The shape is then represented as 2D views, with the projection of the 3D object [31, 96]. For the implementation of this descriptor, the objects first have to be PCA (principal component analysis) normalized (i.e., pose normalization, see Section 4.6) and scaled into a canonical unit cube invariant to rotation that is axis-parallel to the principal axes (canonical unit cube is defined as a cube in the canonical coordinate frame defined by a set of vertices) [96]. Each view is represented as a feature vector of the Fourier coefficients [15, 23, 31, 36]; the discrete Fourier transform is used to represent the shape features in the spectral domain. Therefore, the absolute values of the resulting coefficients are used as the vector of silhouette-based features [21, 89].

3.4.2 Depth buffer descriptor

This descriptor calculates the similarity between the depth buffer of the feature vectors of two 3D models [31]. The depth buffer descriptor is based on the same representation as the silhouette descriptor, which, likewise, uses significant images. Each model is represented and scaled into a canonical unit cube. The method considers for each of the principal axes, two grayscale images that are generated using parallel projections. This generates six images instead of the three silhouettes. Each gray pixel of the image is coded with a value of eight bits and represents the distance between the 3D model and the sides of the unit cube, the viewing plane [96].

3.4.3 Lightfield descriptor

Chen et al. [18] defined this descriptor as the basis representation of a 3D model invariant to translation and scaling through a set of 10 rendered images of different parallel projections (viewing directions, silhouettes or views) that can be put on 20 vertices of a regular dodecahedron. The camera-generated projections are distributed uniformly over a 3D model representing the shape of the model. Each projection is a binary color or grayscale image, containing information about the represented object surface [23]. The images are encoded using a combination of 35 coefficients for the Zernike moments descriptor, and 10 coefficients for the Fourier descriptor [18].



Fig. 2 Example of silhouettes of a 3D model

3.4.4 Elevation descriptor

The principal idea of this descriptor is to treat a 3D model on six 2D projections from different views: front, left, right, rear, top and bottom. Each projection is one elevation that is represented by a grayscale image, in which the gray values represent the altitude information of the 3D model. Subsequently, each grayscale image is decomposed into several concentric circles. The elevation descriptor is calculated taking the difference between the altitude sums of two successive concentric circles. This descriptor has the feature of being invariant to translation and scaling and is robust for rotation [21, 81].

3.4.5 Spin images

A set of 3D points with their associated directions are used to create a spin image. The spin images make up a general representation of the shape [39]. The orientation of the vertices in the surface mesh is considered for each spin image. The information associated with these images is represented as a 2D histogram or a 2D indexed accumulator, which encodes the vertex density of an object in a space [36]. According to Dos Santos [21], this descriptor proposes an object recognition system based on matching surfaces, using the spin image representation.

3.4.6 Scale-invariant feature transform (SIFT)

This descriptor method involves transforming an image into a collection of local feature vectors, which are possibly invariant to image translation, scaling and rotation, and partially invariant to changes in illumination and affine 3D projection. SIFT is based on four stages: (1) scale-space extrema selection, (2) keypoint localization, (3) orientation assignment and (4) keypoint descriptor. The first stage consists of identifying the entire scales and images space by looking for the space locations using a difference of the Gaussian function to identify potential points of interest. In the second stage, the interest points are selected by a degree of stability. In the third stage, an orientation is assigned to each interest point location based on the local gradient directions of the image. Finally, during the fourth stage, a local image descriptor is created for each interest point based on the gradients of the images [56, 57].

3.5 Feature-based descriptors

This category of descriptors was proposed by Zhang L. et al. [107] and Tangelder et al. [89]. This category refers primarily to descriptors of global and local features, which were the point of reference for the development of many other descriptors grouped in other categories. The descriptors in this category express the geometric and topological properties of the shape of each 3D model. The shape of an object is discriminated by measuring and comparing its features. These descriptor methods aim to represent the shape of a 3D object with the implementation of a compact vector. A simple way to do this is by using functions defined on the unit sphere [46]. Feature-based descriptors extract the features of the 3D model in a fast and simple way [15].

3.5.1 Global features

Some of these global features are: invariant moments, Fourier transform descriptors, volume and surface area, or the shape boundary associated with the geometry ratios of an object. Global features are intuitive for people and can be easily obtained, although they are of no use for discrimination at the local level, that is, of object details. There are a lot of efficient ways to calculate these features from the mesh representation of an object [36, 89, 104, 107], and they are easy to implement. The descriptor is often used as an active filter in object retrieval for the purpose of comparison or can be used in combination with other methods to improve the shape descriptor [21].

3.5.2 Local features

Value vectors for this descriptor are calculated from a number of points on the surface of the object [21]. The local features are inherent to the shape of the 3D object and play an important role in the discrimination of complex objects [107]; they can speed up searches by reducing the model to a small number of features more easy to manage. The vectors are compared to measure if two points are similar, or used in a classifier. Here the descriptor can pick out smaller or larger features. The method uses a metric, which reference to the value of curvature, calculated at every point on the surface [32]. Local features provide a very distinctive measure of similarity for each object.

3.6 Geometry-based descriptors

Geometry is always specified in 3D models in contrast with other application-dependent features [15]. The geometry features usually used to describe the 3D model are volume, surface area or curvature, and ratios, like the surface area to volume ratio; compactness, that is, the non-dimensional ratio of the volume squared over the cube of the surface area; crinkliness, that is, the surface area divided by a model made up of the surface area of a sphere with the same volume as the 3D object; convex hull features; the bounding box aspect ratio; or Euler numbers [23].

3.6.1 Volumetric error

This method is based on the observation that different objects occupy the volume in different ways. Volumetric error descriptors require a pose-normalized process. Also, by definition, they use voxelized shapes. This could result in high computational costs if the model has need of a lot of voxels [63, 89]. Because two objects may not be similar even if they have an equal total volume [63], a simple volume difference technique is not good enough to compare 3D models, and different approaches have been proposed for this comparison measure (see, Tangelder et al. [89]; EINaghy et al. [23]). Novotni et al. [63] proposed, in their extraction step, to calculate the volumetric error between one object and a sequence of offset hulls of the other object and vice versa, producing two histograms to measure similarities.

3.6.2 Weighted point set

In this method, a set of 3D salient points is selected from 3D objects. These points are weighted in different ways according to different approaches; the weighted point set of two objects is then used to compare the 3D objects (see, Tangelder et al. [89]; EINaghy et al. [23]). Some examples of weighted point set are: the vertex with the highest Gaussian curvature; the choice of one vertex using the center of mass of all the vertices in the cell as a weight [88]; or a hierarchy of weighted point sets, representing spherical shape approximations [80].

3.6.3 Extended gaussian image (EGI)

The shape of an object's surface can be used for object recognition. One way to do this is using parallel rays on a regularly spaced grid creating its depth map. Unfortunately, depth maps are not easily transformable when the object rotates. The EGI solves the problem of converting a local representation in the viewer-centered coordinate system into a global description based on the object-centered coordinate system [35]. An EGI model of a 3D object is derived from the collection of its surface normal on each surface patch in the 3D world. A spike model is created by collecting normals from surface patches of the 3D model and is moved to a common point of the application; the end points lie on the surface of a unit sphere. This mapping is called the Gauss map, and the unit sphere is called the Gaussian sphere. By giving a unit mass to each end point of the normal vector, we can observe the distribution of mass on the Gaussian sphere. This distribution of mass will be normalized. The resulting distribution of mass on the Gaussian sphere is called the extended Gaussian image (EGI) of the object [38]. The extended Gaussian image has several important properties that make it useful for shape analysis and matching: it is invariant to translation; the EGI scales and rotates with the model in the 3D space; the EGI is an invertible representation for convex models; and it is very good at distinguishing between man-made objects and natural objects [23, 82].

3.6.4 Canonical 3D morphing

Object morphing is a seamless transition from one object to another, generally producing a sequence of intermediate objects [21]. In this method, the similarity of two 3D objects is measured by the energy required to transform them into a canonical space or shape, for example a sphere. Different approaches have been used to measure this energy (see EINaghy et al. [23]), which requires a pose normalization technique (see Section 4.6).

3.6.5 Canonical 3D hough transform (C3DHTD)

The basic idea of this descriptor is to accumulate points of the 3D objects within a set of planes. The planes are determined by parameterizing the space using spherical coordinates (e.g., distances from the origin, azimuth angles and elevation angles to get three different planes). Each triangle of the object contributes to each plane with a

weight equal to the projection area of the triangle on the plane, but only if the scalar product between the normal of the triangle and the plane is above a given threshold. The rotation invariance for this descriptor is approximated by PCA pose normalization of the 3D object, along with the determination of the principal axes, and the use of their center of gravity as the origin of the coordinate system for the Hough transform. This descriptor, also called C3DHTD and first proposed by Zaharia et al. [103], is intrinsically topologically stable and not invariant to geometric transformations [15].

3.6.6 Heat kernel signature (HKS)

Diffusion geometry depends on the heat diffusion equation, which governs the conductions of heat on a surface. The heat diffusion process describes the evolution of a function on the surface over time and is governed by the heat kernel (HK): which is uniquely defined for any two points x, y on the surface, and a time parameter t [93]. Sun et al. [84] proposed the use of the diagonal of the heat kernel as a local descriptor, referred to as the heat kernel signature (HKS), by capturing the information about the neighborhood of a point on a shape and recording the dissipation of the heat from the point onto the rest of the shape. Detailed, highly local shape features are observed through the heat diffusion behavior over a short time, whereas the shape summaries in large neighborhoods are observed from the heat diffusion behavior over a longer time. Thanks to this heat diffusion property, multi-scale matching can be performed between points by comparing their signatures at different time intervals. This model presents a number of advantages, including the fact that the HKS is deformation invariant, captures differential information in a small neighborhood and global information about the shape for large values, and is clearly analogous to the multi-scale feature descriptors used in the computer vision community. For small scales, this descriptor takes into account local information, creating topological noise. It can be built across different shape representations [23, 84].

3.6.7 View based

The underlying idea of this descriptor is that if two 3D models are similar, they also look similar from all viewing angles. This is how humans recognize similarity in objects. Using this method the problem of 3D retrieval is reduced to a 2D projection of the spatial objects, and image retrieval techniques can be used. The main challenge of the view-based descriptors is to get enough views to describe all possible model features, which is related to an enormous storage space cost [23].

3.6.8 Deformation based

This method compares a pair of 2D shapes by measuring the amount of deformation required to exactly register the shapes. This depends on the natural arc length parameterization of the contours of the objects [89]. In the shape space of 3D models, a deformation sequence is shown by a curve. The geodesic distance between two points yields the similarity between two shapes according to the preserved shape space property. The geodesic distance illustrates the similarity between two shapes, used to compute the similarity between the deformations [87].

3.6.9 Surface normal directions

A normal vector shows the entire shape of a 3D model. Histograms constructed by the angle between the first two principal axes and the surface normal vectors of all polygons, generating a lot of information with a low level of detail, used to compare 3D objects [70].

3.7 Extension-based descriptors

These descriptors are created from samples of features taken along certain spatial directions with a starting point in the center of the object [15]. The 3D object is usually treated as functions defined on spheres and described in terms of samples taken from these functions [7].

3.7.1 Ray based

A ray-based feature vector is determined by the measures of the extent of the object from its origin to the furthest unit vector intersection point. The extent is zero if the mesh is not intersected. The samples of numbers recovered from the PCA-normalized 3D object are the components of its feature vector in the spatial domain [98, 99]. This descriptor has also been called sphere projection and has performed well in virtual reality modeling language (VRML) models available over the Internet. A distinctive feature of this model is its large dimensionality. One of its main attributes is that it is able to characterize many samples of a function on a sphere with few parameters. Spherical harmonics were proposed as a suitable tool for use in this method. The magnitudes of the complex coefficients are obtained using the fast Fourier transform on the sphere (SFFT) of samples as components of the feature vector. Apart from the ray extent, a rendered perspective projection of the object on an enclosing sphere was considered. This information can be regarded as information about shading on the enclosing sphere. The complex feature vector uses both ray-based and shading-based feature vectors [30].

3.8 Volume based

This category includes all descriptors that represent the shape of a solid object through the volumetric representation obtained from the surface of a voxelized object [15] (see Fig. 3). This representation is computationally expensive, and its accuracy depends on the size of the voxel (see Section 4.7).

3.8.1 Discretized model volume

The basic idea of this descriptor is to divide the space occupied by the model. Then the content of the fragments of the divided model are added as part of the information on feature vector. As in other cases, this descriptor uses a pose normalization process and PCA [15].

3.8.2 Rotation invariant point cloud

This descriptor takes into account the density of the point clouds as a feature vector. A 3D model placed in a unit cube, which is divided into coarse grids, is required to

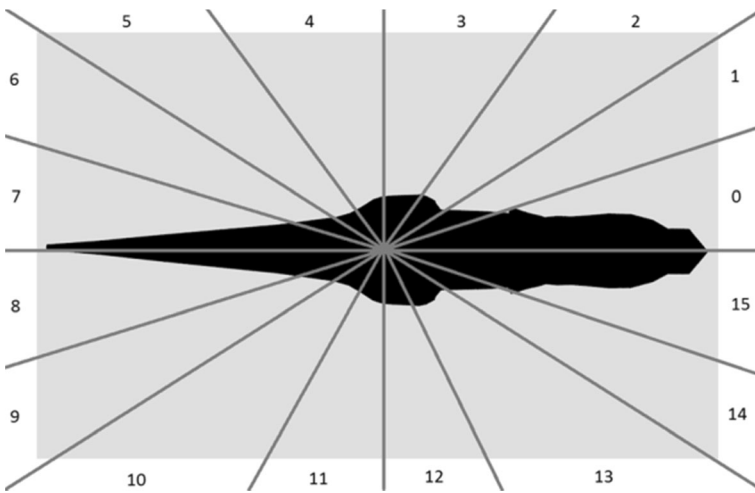


Fig. 3 Example of volume-based feature vector

implement this descriptor. The points of each cell are counted to determine the density of point clouds. The density of the resulting point cloud represents the shape descriptor providing information on the curve, the height and the position of the 3D model [86].

3.8.3 Voxelized volume

This descriptor was proposed as a method for estimating the similarity between 3D voxelized models taking the volume of the models in the space and the internal structure of this volume as a reference. The method of this descriptor is to represent a voxelized model in a set of statistical moments, which are calculated from different levels of detail of the voxel grid by applying the Daubechies D4 wavelet transform on a 3D grid. Each resulting moment is represented with a histogram of the distribution of the built moment. The bins are associated with the level of detail and the amplitude that corresponds to the normalized moments. Each histogram is an object descriptor [73].

3.8.4 Reflective symmetry

This descriptor uses the shape of a voxelized 3D object represented in terms of its symmetrical global features through a continuous measure and 2D planes. Its method consists of a spherical function to get the invariance measure of a 3D model with respect to the reflection of each plane through its center of mass. The points captured in the reflective symmetry descriptor correspond to a measure of global shape, where the peaks are the planes of near reflective symmetry, and the valleys correspond to the planes of near anti-symmetry. This descriptor has the feature of being stable against high frequency noise, scale invariant and robust with respect to the object detail level [41].

3.9 Other methods

These descriptors usually serve the purpose of improving the retrieval process by being integrated to other retrieval-oriented 3D object descriptors. The descriptors and employed techniques are described below [23].

3.9.1 3D Zernike moments

The 3D Zernike moments descriptor is a natural extension of the spherical harmonics descriptors (see Section 3.2.3). It is based on the polynomials of the same name, which are developed by means of a process of orthogonalization (key process for obtaining a set of vectors generating the same vector subspace from a set of linearly independent vectors in a vector space) defined within a unit circle of the complex plane. The moments are computed as a projection of the function defining the object onto a set of orthonormal functions within this unit circle. Zernike moments are able to extract the features of a 3D voxelized model, capturing global information about the 3D shape, which, unlike boundary-based methods, does not require closed boundaries. Zernike moments are invariant to translation, scaling and rotation. At the same time, they are robust to noise and easy to build. In recent years, these descriptors have become a popular tool for digital image reconstruction, pattern recognition and shape analysis [64].

3.9.2 Spherical moments

Liu W. et al. [53] proposed this method that extracts a feature vector of a 3D model based on an analysis of multilevel spherical moments. A process of pose normalization is required for its implementation in order to align each 3D model into a canonical coordinate system. Subsequently the surface of the model is rasterized into a voxels grid, and each voxelized model is aligned from the center of its mass towards the center of the grid. Finally, a set of homocentric spheres centered on the center of the 3D voxelized model is used to produce N spherical images. This can be used to verify if one of these voxels intersects a trigonal pixel on the surface of the spheres and the voxels of the object. For each sphere, there are M moments with which it is possible to build a feature vector that has N dimensions using the moments of all the spheres.

3.9.3 Relevance feedback

According to Leifman et al. [50], this descriptor facilitates the interactive retrieval of 3D objects including the user perceptual information. The descriptor is a technique based on the representation of low-level data by means of a three-step iterative process. In the first step, the system retrieves similar 3D models which it presents to the user in descending order of similarity. In the second step, the user provides information concerning the relevance of some of the results of the retrievals. Finally, the system uses these models to learn and improve its performance in the next iteration for 3D object retrieval. In this type of descriptor, the valuable judgment of the users can identify the multilevel relevance of the retrieved objects [5], that is, the user can assign importance scores to a series of results. Then, the query metric may automatically be adjusted to a new classification more in line with the provided relevance scores [15].

3.9.4 Bag of features

This descriptor seeks to represent a feature vector of an object located in an image. The method makes use of the local features because it does not obtain the global geometric shape of the model. Each local feature of an object is encoded in the descriptor with one of several words within a vocabulary. The vocabulary can be generated using k-means clustering varying the value of k . Finally, regardless of the position of its features, each object is represented with a histogram that contains word frequencies. The histogram is now the feature vector extracted from the image [23, 54, 65]. The use of this descriptor focuses on partial matching retrieval of 3D models, and is recommended for geometrically detailed and highly articulated objects. El Wardani et al. [24] made an improvement of this method; the new solution consists on integrating a set of features extracted from 2D views of the 3D object, and employing the SIFT algorithm into histograms. Here the method uses a vector quantization based on a global visual codebook.

3.9.5 Topological matching

The principal idea of this descriptor is that it describes the shape of a 3D object by using its graph structure in a Reeb graph. The graphs capture and interpret information about the structure of the skeleton of an object. The object is divided into connected parts through a continua function, which is defined over the surface of the object. A geodesic distance between points over the surface is used to define the function. The geodesic distance provides an invariance rotation and avoids problems caused by noise or small undulations [33].

4 Comparative analysis of shape descriptors

The shape descriptors described in this article have particular characteristics according to their context, purpose, or method. Method of similarities are used to categorize the descriptors; most of these methods used a similarity measure based on the Euclidean distance. These descriptors are efficient for the purpose for which they were created.

This analysis was based on the characteristics evaluated by Tangelder et al. [89] and ElNaghy et al. [23], that is: a) shape model, b) matching, c) robustness, d) efficiency and e) pose normalization. The characteristics that they evaluated were oriented toward the content-based retrieval of objects. These same characteristics were evaluated in the same way in the search of shape similarity in 3D objects. Therefore, we incorporated elements that provide useful information for the shape representation of 3D models and the comparison of the similarity between objects: for this, the selected characteristics were: a) shape model, b) 3D objects with 2D image projections, c) efficiency, d) matching or similarity, e) robustness, f) pose normalization, g) voxelization required, and h) visual salience. Also we included i) databases for testing, and j) references (some cited references). Next, these characteristics are detailed.

4.1 Shape model

The descriptors analyzed in this paper use different representation models, also called types, formats and methods for 3D shape extraction. These models are mesh=triangulation, point cloud, volume and solid.

Zhang L. et al. [107] and EINaghy et al. [23] considered that the most popular format for the shape representation of a 3D object is a polygonal mesh, also known as triangulation, on the grounds of its simplicity. This format builds a representation of the boundary surfaces of the inside and outside of an object. When the entire boundary surface is represented by a union of 3D polygons and these polygons are triangles, then we have a representation in the format of triangle mesh [96].

The point cloud format is characterized by a set of data points in a coordinate system. For 3D models, this method represents the exterior surface of an object. And, for 2D images, the representation of points corresponds only to a set of pixels.

According to EINaghy et al. [23], an object can be represented in terms of the volume that it occupies. Hoffmann [34] described the format of solid model representations that use CSG, that is, a solid is represented as a set of theoretic Boolean expressions of primitive solid objects with a simpler structure. In this case, the surface and the interior are defined implicitly. At the same time, a B-rep describes the oriented surface of a solid as a data structure composed of vertices, edges, and faces. Icke [37] stated that the volume-based format is known as volumetric (solid) representations, which, for 2D shapes, is associated with the area. CSG and octree are examples of volumetric (solid) representations.

4.2 3D objects with 2D image projections

Although we live in a 3D world, humans only see in 2D, that is, the human visual system only receives centered projections on 2D plane images. For this reason, different frameworks are used for the recognition of 3D objects. So, an object can be viewed as a set of, sometimes grayscale, 2D images [47]. Several shape descriptors analyze the 3D object from multiple views or 2D images. Each projection is rasterized to an image that can be described using a 2D visual descriptor, e.g., contour-shape, color or texture. 3D objects are matched by comparing their 2D images [96]. This characteristic is typical of all the descriptors included in the 2D image-based category.

4.3 Efficiency

As a database can store a large number of 3D objects for retrieval, clustering, classification, alignment and registration, and approximation and simplification with respect to their shape, shape descriptors should have an efficient feature extraction structure. Likewise, the efficiency of a descriptor can be measured with respect to the query response time (computational time) [23] in terms of storage requirements [29, 69] or descriptor dimensionality reduction in order to make its representation more compact [42, 51].

Table 3 Comparison between 3D object shape descriptors

Section	Descriptor Name	Shape Model	3D object with 2D projections	Efficiency	Matching / Similarity	Robustness	Pose Normalization required	Voxelization required	Visual Salience	Tested in the Database	Some references
3.1.1	Cord and Angle Histograms	Mesh = Triangulation	No	High	Yes	Medium	Yes	Yes	No	- NRC three-dimensional database, with 400 range images and 1,000 VRML models	[1, 15, 21, 70, 73]
3.1.2	Color Distribution	Mesh = Triangulation	No	High	Yes	Medium	Yes	Yes	No	- NRC three-dimensional database, with 400 range images and 1,000 VRML models	[21, 70, 73]
3.1.3	Curvature Histogram	Mesh = Triangulation	No	Medium	Yes	Unknown	Unknown	Unknown	No	- Not identified	[21, 44, 93]
3.1.4	Shape Distributions	All models	No	High	Yes	High	Yes	No	No	- Database of 3D polygonal models downloaded from the World Wide Web	[1, 15, 21, 36, 68, 82]
3.1.5	Modified Shape Distributions	Mesh = Triangulation	No	Medium	Yes	High	Yes	No	No	- Not identified	[2, 66]
3.1.6	Shape Histograms	Point cloud	No	High	Yes	High	Yes	No	No	- Patrick Min and Prof. Funkhouser at Princeton University (215 VRML models)	[1, 4, 15, 21, 23, 82]
3.1.7	3D Shape Contexts	Mesh = Triangulation	No	High	Yes	High	Yes	No	No	- Private with 1,213 VRML	[10, 21, 23, 45]
3.1.8	Shape Spectrum	Mesh = Triangulation	No	Medium	Yes	High	Yes	No	No	- Princeton Shape Benchmark (PSB), with 1,814 models	[1, 20, 21, 23, 102, 103]
3.1.9	Complex Extended Gaussian Images	Mesh = Triangulation	No	Low	Yes	Medium	Yes	No	No	- FSSP, database of proteins	[1, 21, 40, 82, 107]
3.1.10	Density-Based 3D Shape	Mesh = Triangulation	No	High	Yes	High	No	No	No	- Database available at http://www.3dtef.com	[2, 21, 107]
3.1.11	3D Hough Transform	Mesh = Triangulation	No	High	Yes	Unknown	No	No	No	- MPEG-7 3D model database, with 1300 mesh in VRML 2.0 format	[1, 21, 101]
										- PSB databases, with 1814 models	
										- PSB databases, with 1814 models	
										- Sculpteur Database (SCUdb), with over 800 models	
										- MPEG-7 3D model database, with	

Table 3 (continued)

Section	Descriptor Name	Shape Model	3D object with 2D projections	Efficiency	Matching / Similarity	Robustness	Pose Normalization required	Voxelization required	Visual Saliency	Tested in the Database	Some references
3.1.12	Generalized Shape Distributions	Mesh = Triangulation	Yes	High	Yes	High	No	No	No	1300 mesh in VRML 2.0 format - PSB databases, with two sets of 907 models each (training and testing)	[1, 21, 23, 54]
3.1.13	Geometric Hashing	Point Cloud	Yes	High	Yes	High	No	No	No	- CAD-CAM objects models	[23, 47]
3.1.14	Spatial Maps	All models	Unknown	High	Yes	High	Yes	No	No	- Not identified, but includes 1,900 models	[23, 89, 99, 107]
3.1.15	Simple Statistics	Mesh = Triangulation	Unknown	Unknown	Yes	Unknown	Unknown	Unknown	No	- Not identified	[15, 73]
3.1.16	Parameterized Statistics	Mesh = Triangulation	Unknown	Medium	Yes	Unknown	Yes	No	No	- Not identified, but includes 261 models	[15, 67, 89]
3.1.17	Geometric 3D moments	Mesh = Triangulation	No	High	Yes	High	Yes	Yes	No	- CAESAR anthropometric database NRC's NH and CLEOPATRA applications available at http://cook.iitg.nrc.ernet:8800/Neterrit	[15, 23, 73, 78]
3.2.1	(3D) Angular Radial Transform	Mesh = Triangulation	No	High	Yes	High	Yes	Yes	No	MPEG-7 3D model database - MPEG-7 3D model database base	[1, 13, 21, 77]
3.2.2	Voxel-3D Fourier Transform	Mesh = Triangulation	No	High	Yes	Medium	Yes	Yes	No	- PSB databases, with 1814 models - Renault database, with 5000 models	[1, 3, 15, 82, 97, 107]
3.2.3	Spherical Harmonics Transform	Mesh = Triangulation	No	High	Yes	High	Yes	Unknown	No	- Database available at http://www.3dcafe.com , with 1,830 3D-models - Database including 1,839 objects, mostly collected from the Internet (www.3dcafe.com and www.vierpoint.com)	[1, 21, 36, 82, 98]
3.2.4	Planar-Reflective Symmetry Transform	All models	Yes	Medium	Yes	Medium	No	Yes	Yes	- MPEG-7 3D model database base	[21, 41, 76]
3.2.5		Mesh = Triangulation	No	High	Yes	High	Yes	Yes	No	- PSB databases, with a set of 907 models for testing	[1, 15, 21, 42]

Table 3 (continued)

Section	Descriptor Name	Shape Model	3D object with 2D projections	Efficiency	Matching / Similarity	Robustness	Pose Normalization required	Voxelization required	Visual Saliency	Tested in the Database	Some references
	Rotation Invariant Spherical Harmonics									- Database of 1,890 "household" objects provided by Viewpoint	
3.2.6	Distance Transform and Radial Cosine Transform	Mesh = Triangulation	No	Low	Yes	Medium	No	Yes	No	- PSB databases, with two sets of 907 models each (training and testing)	[1, 21, 22]
3.2.7	Spherical Wavelet Transform	Volume	Unknown	High	Yes	High	Yes	No	No	- PSB databases	[1, 46]
3.3.1	Reeb Graphs	Volume	No	Medium	Yes	Medium	No	No	No	- Experiments on bones of human cadavers (medical applications)	[1, 21, 23, 33, 83, 89, 90, 107]
3.3.2	Multiresolution Reeb Graphs	Volume	No	Medium	Yes	Medium	No	No	No	- 230 different polyhedral meshes selected from the Viewpoint models. http://www.viewpoint.com	[1, 21, 33, 89, 90, 107]
3.3.3	Size Graphs	Mesh = Triangulation	No	Unknown	Yes	Unknown	Yes	No	No	- 3DCAFE free stuff; http://www.3dcafe.com	[12, 16, 21]
3.3.4	Skeleton(al) Graphs	Volume	No	Medium	Yes	Medium	No	Yes	No	- Stanford University dataset and our original data, http://www-graphics.stanford.edu/~data/3Dscanrep	[1, 15, 21, 23, 85, 89, 107]
3.3.5	Spectral Graph	Solid	No	Medium	Yes	Medium	Unknown	Unknown	No	- Not identified	[1, 23]
3.3.6	B-Rep Graph	Solid	No	Medium	Yes	Medium	No	No	No	- Experiments with CAD models like cylinders, cube dihedrals, etc.	[23, 25, 89]
3.3.7	Model Graph	Solid	No	Medium	Yes	Medium	No	No	No	- Experiments with CAD models.	[23, 89, 107]
3.4.1	Silhouette Descriptor	Mesh = Triangulation	Yes	Medium	Yes	Medium	Yes	No	No	- Private database, with 1,841 mesh models,	[1, 15, 21, 23, 36, 89, 96]

Table 3 (continued)

Section	Descriptor Name	Shape Model	3D object with 2D projections	Efficiency	Matching / Similarity	Robustness	Pose Normalization required	Voxelization required	Visual Saliency	Tested in the Database	Some references
3.4.2	Depth Buffer Descriptor	Mesh = Triangulation	Yes	High	Yes	High	Yes	No	No	available at http://www.3dcatf.com , www.viewpoint.com	[1, 15, 21, 23, 36, 96]
3.4.3	Lightfield Descriptor	Mesh = Triangulation	Yes	Low	Yes	High	Yes	No	No	- Official MPEG-7 test set (the MPEG-7 collection (M1) consists of 227 models in VRML 2.0 format) and the Princeton Shape Analysis and Retrieval training collection	[1, 15, 18, 21, 23, 36]
3.4.4	Elevation Descriptor	Mesh = Triangulation	Yes	High	Yes	High	Yes	Yes	No	- Experiments with the MPEG-7 collection - Database available at http://3d.csie.nyu.edu.tw with 10,911 3D models	[21, 81]
3.4.5	Spin Images	Point Cloud	Yes	Medium	Yes	Medium	No	No	No	- MPEG-7 3D SSD - PSB database, with 1,814 models	[1, 21, 23, 36, 39]
3.4.6	SIFT	Solid	Yes	High	Yes	High	No	No	No	- Library of 20 models of complex objects - Experiments with models photographed against a black background, and outlines of objects extracted by segmenting the background region (image size: 384 × 512 pixels)	[23, 56]
3.5.1	Global Features	All models	No	High	Yes	High	Yes	Yes	No	- Not identified	[21, 36, 89, 104, 107]
3.5.2	Local Features	Mesh = Triangulation	No	Medium	Yes	Medium	No	No	No	- Database with 70 models available in http://www.designrepository.org/datasets	[11, 36, 89, 107]
3.6.1	Volumetric Error	Volume	No	High	Yes	High	Yes	Yes	No	- Data set of human bones (hands and spine)	[23, 62, 89]
3.6.2	Weighted Point Sets	Mesh = Triangulation	No	High	Yes	High	Yes	No	Yes	- Not identified, but they used a database	[15, 23, 88, 89]

Table 3 (continued)

Section	Descriptor Name	Shape Model	3D object with 2D projections	Efficiency	Matching / Similarity	Robustness	Pose Normalization required	Voxelization required	Visual Salience	Tested in the Database	Some references
3.6.3	Extended Gaussian Image	All models	Unknown	High	Yes	High	Yes	No	No	consisting of 133 models such as mugs, cars and boats, and a database consisting of 512 models, mostly air planes	[15, 23, 26, 35, 38, 41, 82, 107]
3.6.4	Canonical 3D Morphing	Mesh = Triangulation	No	High	Yes	Medium	Yes	No	No	- Not identified, but they used a test database of 34 categories with 52 sample objects, generating more sample objects by creating 13 new versions of each sample object by arbitrary rotation, reflection, and non-uniform scaling to total 728 database samples	[21, 23, 100]
3.6.5	Canonical 3D Hough Transform	Mesh = Triangulation	No	High	Yes	Medium	No	No	No	- MPEG-7 3D model database, with 1300 mesh in VRML 2.0 format	[15, 103]
3.6.6	Heat Kernel Signatures	All models	No	High	Yes	High	No	No	No	- Not identified, but they used a dataset with 175 models, of girls dancing	[23, 84]
3.6.7	View Based	Mesh = Triangulation	Yes	Medium	Yes	High	No	No	No	- Not identified	[23, 89]
3.6.8	Deformation Based	Mesh = Triangulation	Unknown	Low	Yes	High	No	No	No	- Not identified, but they experimented with regular polygons and contours of different shapes	[8, 89]
3.6.9	Surface Normal Directions	Mesh = Triangulation	No	Unknown	Yes	Medium	Unknown	Unknown	Unknown	- Not identified	[15, 70, 73]
3.7.1	Ray Based (Complex Shading)	Mesh = Triangulation	No	High	Yes	High	Yes	Unknown	No	- MPEG-7 3D model database, with a own collection of 1,839 objects	[1, 15, 97, 98, 107]

Table 3 (continued)

Section	Descriptor Name	Shape Model	3D object with 2D projections	Efficiency	Matching / Similarity	Robustness	Pose Normalization required	Voxelization required	Visual Salience	Tested in the Database	Some references
3.8.1	Discretized model volume	Unknown	Unknown	Unknown	Yes	Unknown	Yes	Unknown	Unknown	- Not identified	[15]
3.8.2	Rotation Invariant Point Cloud	Mesh = Triangulation	No	High	Yes	Medium	Yes	No	No	- Not identified, but they experimented with a database of 1,500 furniture models	[15, 86]
3.8.3	Voxelized Volume	Mesh = Triangulation	No	Unknown	Yes	Unknown	Yes	Yes	No	- MPEG-7 3D model database	[15, 73]
3.8.4	Reflective Symmetry	Mesh = Triangulation	Yes	High	Yes	High	Yes	Yes	No	- Test database with 1890 models of "household" and "miscellaneous" objects provided by Viewpoint, containing between 120 and 120,392 triangles, with average of triangles	[1, 15, 21, 41, 76]
3.9.1	3D Zernike Moments	Mesh = Triangulation	No	High	Yes	Medium	Yes	Yes	No	- Experiments with a proprietary database, with a class of 21 chairs and another one consisting of 28 planes - Plus a small set of 655 objects of general categories downloaded from www.3dcatfe.com	[1, 21, 64, 107]
3.9.2	Spherical Moments	Mesh = Triangulation	No	Medium	Yes	High	Yes	Yes	No	- PSD databases, with 1814 models	[21, 53]
3.9.3	Relevance Feedback	Mesh = Triangulation	No	High	Yes	High	Yes	No	No	- The first dataset contained 52 objects from 34 categories	[5, 15, 23]
										- The second set, the Utrecht Database, contained 512 aircraft in six categories: delta jets, conventional airplanes, multifuseages, biplanes, helicopters, and other aircraft	
										- The third set is a subset of the Princeton Database,	

Table 3 (continued)

Section	Descriptor Name	Shape Model	3D object with 2D projections	Efficiency	Matching / Similarity	Robustness	Pose Normalization required	Voxelization required	Visual Saliency	Tested in the Database	Some references
3.9.4	Bag of Features	Mesh = Triangulation	No	High	Yes	High	Yes	No	No	containing 1236 objects in 52 categories - The total test database contained 1910 objects - PSB databases, with 907 models of the “testing part” of the benchmark	[23, 55]
3.9.5	Topological matching	Mesh = Triangulation	No	High	Yes	Medium	Yes	No	No	- 230 different polyhedral meshes selected from the Viewpoint models at http://www.viewpoint.com . 3DCAFE free stuff at http://www.3dcafe.com ; Stanford University dataset and their original data at http://www-graphics.stanford.edu/data/3Dscanrep	[15, 33]

*References indicate which papers provide the specified property. If no reference is given the property is valid in general

4.4 Matching or similarity

Shape matching is a task applied in a large number of geometric applications from different research areas such as computer vision, robotics, and molecular biology. It is the measure of similarity to another shape [94]. In order to measure the shape similarity of two objects, the distances between pairs of descriptors must be calculated using a dissimilarity measure. The term similarity is commonly used, however dissimilarity is best associated with the notion of distance: a small distance means small dissimilarity, and large similarity [89].

We found that the principal similarity measures used are: (1) Euclidean distance: the ordinary measure between points of the Cartesian coordinate system of 2D or 3D spaces; (2) Mahalanobis distance: the distance measure that calculates the similarity between two random multidimensional points; (3) Hausdorff distance: based on two sets of points of different size and no one-to-one correspondence between all points; another measure of dissimilarity may be necessary in this case [94]. All the shape descriptors described here can be used to get a matching measure with respect to a similarity measure.

Chen et al. [17] described their own matching measure method. Their procedure supports many classification algorithms and consists of four steps: (1) handcraft a feature-feature similarity measure; (2) for each given shape, find the optimal matching correspondence with another form of the set of shapes for comparison and maximize the feature-feature similarity sum of the shapes to get the correspondence between two shapes; (3) use the score of the optimal match as a similarity measure between shapes, and (4) classify the shape based on this similarity measure.

In the matching measure it is also possible to use partial matching methods on surfaces represented by triangular meshes (see Section 4.1). This procedure looks for a match with regions of the surface that are numerically and topologically different, but approximately similar. Another way of searching for a match between two objects is to use salient geometric features based on indexing rotation-invariant features and a voting scheme accelerated by geometric hashing [27]. An object matching process does not have to be calculated on the total shape of an object, it can be applied on regions or parts of the object's surface. Furthermore, the matching measure process can be used for both local and global object features.

4.5 Robustness

Vranić et al. [97] considered the robustness of a shape descriptor from the viewpoint of its level of detail and the representation of its outliers. Other authors evaluated the robustness of their descriptor by sampling the points from a model surface [42], for which purpose it is important that the model does not require the point set to be triangulated. The robustness of the descriptor is mainly evaluated for properties such as invariant to translation and rotation, robust against connectivity changes caused by simplification and subdivision in the mesh [33], changes in the parameters (scale), noise resistance, and changes due to deformation by small extra features, and independence of 3D object representation. The robustness of several descriptors was considered with respect to support radius variations, Gaussian noise, shot noise, varying mesh resolution, distance to the mesh boundary, key-point localization error, occlusion, clutter, and dataset size [29].

4.6 Pose normalization

Pose normalization is a procedure for assuring that the object state conforms to certain user-defined characteristics; it transforms a model into a canonical coordinate frame. Using this procedure, a scale, position, rotation, or orientation of a model are chosen whose representation in canonical coordinates would be unchanged. All the objects may have different levels of detail but their normalized representations should be as similar as possible for comparison [78]. Pose normalization is not necessary when local features are used (e.g., curvature) [97]. This procedure has to be applied when heterogeneous databases are used [23].

4.7 Voxelization

A 3D object has volume and can be analyzed as such [70]. To do this, a voxelized object in a 3D discrete representation of small blocks or called voxels is required. Several descriptors require the voxelization process, which produces a set of values on a regular 3D grid [74].

4.8 Visual salience

Humans have the ability to identify or locate objects using their basic features such as size, color or shape. This provides relevant information about the object and its areas. For the shape identification of an object in a 2D image or a 3D model within a database, different methods use relevant information such as bag of words (areas that define the object); characteristics of a region (variable characteristics), the local aspect (salient areas), etc. [91]. Also, the surface of an object plays an important role in visual salience, measuring the similarity between subparts of the shape. Salient feature matching is able to search the similar parts of a single surface or series of surfaces. Each salient feature is associated with rotation- and scale-invariant indices to accelerate matching processes and similarity measures. These descriptors are based on the surface, and consequently efficiency depends directly on the quality of the mesh and the curve analysis as they are significantly important for comparing the degree of similarity [27]. Other option proposed by Godil et al. [28] uses a new formulation for the 3D salient local features based on the voxel grid inspired by the Scale Invariant Feature Transform (SIFT); this method can be applied in the non-rigid 3D model retrieval. The advantage of this is that the new approach can be used to both rigid models as well as to articulated and deformable 3D models. In this review, we identified two descriptors that clearly specify the importance of the visual salience of the objects using object shape features: planar-reflective symmetry and weighted point set descriptors. The planar-reflective symmetry descriptor uses local symmetries of salient parts from a transform of the space of points to the space of planes. It gets a continuous measure of the symmetry of an object with respect to all its planes through its bounding volume [76]. The second descriptor generates a weighted point set for each object, which represents each non-empty grid cell as a salient point [88].

4.9 Databases used for testing

The most used databases for testing were Princeton Shape Benchmark (PSB), with 1,814 models, MPEG-7 3D model database, with 1300 mesh in VRML 2.0 format, and private

databases; other more recently database used for test is the Toyohashi Shape Benchmarks, which contains 10,000 models. Schmitt et al. [79], for example, validated their descriptor based on the distribution of two global features measured in a 3D shape, with this database.

In Table 3 are related the aforementioned characteristics to the 58 described descriptors. Each column was evaluated as follows: efficiency and robustness were rated as: high, medium and low. The label in the 3D object with 2D projections, matching/similarity, pose normalization required, voxelization required and visual salience columns is Yes or No, as the case may be. When no information was available, the characteristic was labeled as Unknown.

Table 4 summarizes the calculated frequencies and percentages of Table 3 with information retrieved from the consulted sources. The frequencies and percentages include all the known or unknown data about the analyzed characteristics of each descriptor. Based on this information, we can state that:

- The most commonly used format for representing 3D models is mesh=triangulation, which accounts for two-thirds of the total, not counting the descriptors that do not specify the used representation model. This is might be because the mesh is better adjusted to the shape of large or small objects or because it is better adapted to the technical requirements for its calculation. Mesh = triangulation has the following advantages: the vertices are easy to number, it is clear which vertices are connected to others, and they are suitable for handling an object's geometric information.
- The term “solid” is synonymous with object [34]. A set of flat surfaces (also called views or images) can be extracted from the object in order to represent and describe a solid object. This was an alternative format used to represent a 3D object. The set of descriptors included in the 2D image-based category combines this representation format with another format listed in Section 4.1, and less than a quarter of the total number of descriptors used this representation.
- The efficiency of more than two-thirds of the descriptors was rated as medium or high. We believe that this information was based primarily on the evaluation and interpretation of the results for task execution (i.e., classification, retrieval, clustering, matching/similarity of objects) in the best time and with best utilization of the computer resources.
- All descriptors take a local or global approach to the aspects of the shape similarity or matching among 3D models. The closeness between the vectors output by calculating the model descriptors is able to determine the similarity between shapes. The information contained in these vectors is related to the model's geometric information, which commonly corresponds to the vertices, angles, curved or linear dimensions.
- The robustness of two-thirds of these descriptors was, like efficiency, rated as medium or high. This is probably due to the resistance of object characteristics during the preprocessing processes prior to the descriptor calculation.
- Nearly two-thirds of the descriptors apply a pose normalization process to 3D models before calculation. We observed that the efficiency and robustness of the descriptors that enact this type of process are rated as high. This confirms what El Wardani [24] et al., and Tangelder et al. [89] stated about shape descriptors that use pose normalization preprocessing about having a better performances for retrieving rigid models compared with other approaches; and, that the most popular and practical methods in the field of 3D shape retrieval with this characteristics are the view-based descriptors.

Table 4 Frequencies and percentages from Table 3

Characteristics	H-B		T-B		G-B		2D IB		FB		GB		EB		VB		OM		
	(F)	(P)	(F)	(P)	(F)	(P)	(F)	(P)	(F)	(P)	(F)	(P)	(F)	(P)	(F)	(P)	(F)	(P)	
Shape model																			
Mesh = Triangulation	13	22.4	5	8.6	1	1.7	4	6.9	1	1.7	6	10.3	1	1.7	3	5.2	5	8.6	
All model	2	3.4	1	1.7	0	0.0	0	0.0	1	1.7	2	3.4	0	0.0	0	0.0	0	0.0	
Point Cloud	2	3.4	0	0.0	0	0.0	1	1.7	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	
Volume	0	0.0	1	1.7	3	5.2	0	0.0	0	0.0	1	1.7	0	0.0	0	0.0	0	0.0	
Solid	0	0.0	0	0.0	3	5.2	1	1.7	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	
Unknown	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	1	1.7	0	0.0	
3D Object with 2D projections																			
Yes	2	3.4	1	1.7	0	0.0	6	10.3	0	0.0	1	1.7	0	0.0	2	3.4	0	0.0	
No	12	20.7	5	8.6	7	12.1	0	0.0	2	3.4	6	10.3	1	1.7	1	1.7	5	8.6	
Unknown	3	5.2	1	1.7	0	0.0	0	0.0	0	0.0	2	3.4	0	0.0	1	1.7	0	0.0	
Efficiency																			
High	11	19	5	8.6	0	0.0	3	5.2	1	1.7	6	10.3	1	1.7	2	3.4	4	6.9	
Medium	4	6.9	1	1.7	6	10.3	2	3.4	1	1.7	1	1.7	0	0.0	0	0.0	1	1.7	
Low	1	1.7	1	1.7	0	0.0	1	1.7	0	0.0	1	1.7	0	0.0	0	0.0	0	0.0	
Unknown	1	1.7	0	0.0	1	1.7	0	0.0	0	0.0	1	1.7	0	0.0	2	3.4	0	0.0	
Matching/Similarity																			
Yes	17	29	7	12.1	7	12.1	6	10.3	2	3.4	9	15.5	1	1.7	4	6.9	5	8.6	
No	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	
Unknown	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	
Robustness																			
High	10	17.2	4	6.9	0	0.0	4	6.9	1	1.7	6	10.3	1	1.7	1	1.7	3	5.2	
Medium	3	5.2	3	5.2	6	10.3	2	3.4	1	1.7	3	5.2	0	0.0	1	1.7	2	3.4	
Unknown	4	6.9	0	0.0	1	1.7	0	0.0	0	0.0	0	0.0	0	0.0	2	3.4	0	0.0	

Table 4 (continued)

Characteristics	H-B		T-B		G-B		2D IB		FB		GB		EB		VB		OM	
	(F)	(P)	(F)	(P)	(F)	(P)	(F)	(P)	(F)	(P)	(F)	(P)	(F)	(P)	(F)	(P)	(F)	(P)
Pose normalization																		
Yes	11	19	5	8.6	1	1.7	4	6.9	1	1.7	4	6.9	1	1.7	4	6.9	5	8.6
No	4	6.9	2	3.4	5	8.6	2	3.4	1	1.7	4	6.9	0	0	0	0	0	0
Unknown	2	3.4	0	0	1	1.7	0	0	0	0	1	1.7	0	0	0	0	0	0
Voxelization Required																		
Yes	3	5.2	5	8.6	1	1.7	1	1.7	1	1.7	1	1.7	0	0	2	3.4	2	3.4
No	12	20.7	1	1.7	5	8.6	5	8.6	1	1.7	7	12.1	0	0	1	1.7	3	5.2
Unknown	2	3.4	1	1.7	1	1.7	0	0.0	0	0.0	1	1.7	1	1.7	1	1.7	0	0.0
Visual Saliency																		
Yes	0	0	1	1.7	0	0	0	0	0	0	1	1.7	0	0	0	0	0	0
No	17	29.3	6	10.3	7	12.1	6	10.3	2	3.4	7	12.1	1	1.7	3	5.2	5	8.6
Unknown	0	0	0	0.0	0	0	0	0	0	0	1	1.7	0	0	1	1.7	0	0

Abbreviations: *H-B* histogram-based descriptors, *T-B* transform-based descriptors, *G-B* graph-based descriptors, *2D IB* 2D image-based descriptors, *FB* feature-based descriptors, *GB* geometry-based descriptors, *EB* extension-based descriptors, *VB* volume-based descriptors, *OM* other methods, *(F)* frequency, *(P)* percentage

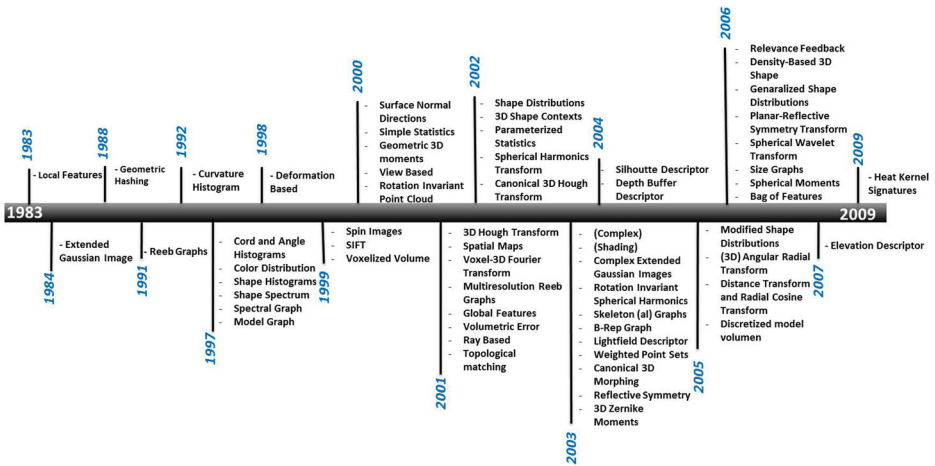


Fig. 4 Timeline of analyzed set of descriptors

- Just over a quarter of the descriptors require a voxelization process. This small amount is possibly due to the fact that voxelization is a disadvantage when the amount and resolution of the data is not very good, that is, the algorithm processing time is

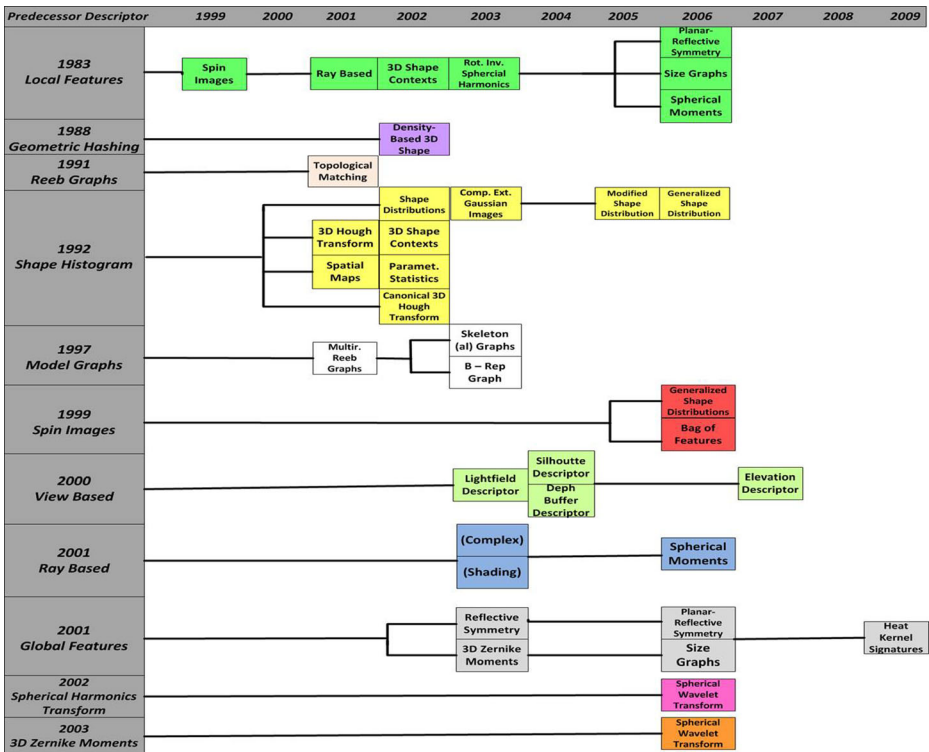


Fig. 5 Descriptors that laid the groundwork for the development of other descriptors (from left to right: predecessor descriptor, year of development, and name of descriptor developed from the predecessor)

- computationally expensive when the size of the voxels is small and the size of the 3D model is large.
- Of the many analyzed descriptors, only a very few considered aspects of object visual salience to get information about their shape. The few descriptors that considered visual salience take into account the geometric characteristics of the object. We take this to mean that visual salience is not a key factor for representing and recognizing the shape of an object in tasks such as classification, retrieval, clustering, similarity search and matching.

Based on the statistics listed in Table 3, we can conclude that the standard of performance of this set of descriptors is good. However, the shape descriptors could be evaluated from many different perspectives.

Figure 4 is a timeline of the descriptor set discussed in this paper. It shows the shape descriptor methods developed from 1983 to 2009. Most of the shape descriptors, 51 out of the 58 discussed in this paper, were developed during the decade from 1997 to 2006.

Figure 5 identifies the descriptors that preceded others. We find that the shape histogram, local features and global features descriptors were important precursors for the development of a considerable number of newer descriptors, on which they had a big impact. The shape histogram descriptor was the basis for the development of nine different descriptors, all of which fall in the histogram-based category. This group of descriptors uses solid objects in a mesh=triangulation or point cloud format. Note that none of these descriptors requires the expensive and complex voxelization process.

The local feature descriptor promoted the development of seven descriptors, which are classed within different categories. This set of descriptors uses methods of representation based on mesh=triangulation, point cloud, or 3D objects with 2D projections and are rated as having a range of medium to high efficiency. Some of them apply the voxelization process.

The global feature descriptor supported the development of five descriptors, which are also part of different categories. This group of descriptors considered the different formats of representation of a 3D model. The descriptors of this set are rated as having a medium to high efficiency, although a voxelization process is required in some cases. In this regard, note that the global feature descriptor is the most intuitive and easiest to obtain for the human visual system (providing visual salience).

On the other hand, these three groups of descriptors are rated as being highly robust, probably because a pose normalization process is applied in advance to the 3D models, as opposed to descriptors that do not require this normalization, and whose robustness is rated as being medium.

In recent years, the concept of deep feature learning based on multiple low level-descriptors, from those here described, has grown to provide high-discriminative representation of local regions of several 3D shape descriptors [14, 111].

5 Conclusions

The computational and mathematical representation of the shape of a 3D model is a complex task in which aspects of the 3D model such as its scale, position, and orientation have to be taken into account. This task has been tackled using different approaches for developing the shape descriptors.

Our motivation for this review was our interest in finding out how to characterize the shape of a 3D object within a virtual environment and how to compare shape similarities. The descriptors have been developed over almost three decades, and they are a powerful and efficient tool for characterizing the shape of an object.

The review of the 58 descriptor types is indicative of the wide variety of methods that have been used in various applications. Each application depends on the context and the goal to be achieved, which are mainly linked to tasks such as retrieval, classification, similarity search, and clustering.

The contributions of this study are to compile and analyze six shape descriptor taxonomies, to provide a general description of each of the descriptors, and to classify them into nine categories. We also analyzed and compared the aspects that we considered to be the most important for the evaluation of these methods.

Acknowledgments Graciela Lara holds a PROMEP scholarship in partnership with the UDG (UDG-685), Mexico.

References

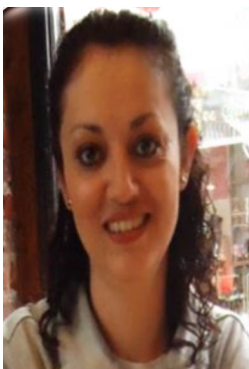
1. Akgül CB (2007) Density-based shape descriptors and similarity learning for 3D object retrieval. Boşaziçi University, Istanbul. PhD. Thesis
2. Akgül CB, Sankur B, Yemez Y, Schmitt F (2006) A framework for histogram-induced 3D descriptors. 14th European signal processing conference (EUSIPCO 2006). EURASIP Florence, Italy
3. Akgül CB, Sankur B, Yemez Y, Schmitt F (2009) 3D model retrieval using probability density-based shape descriptors. *IEEE Trans Pattern Anal Mach Intell* 31(6):1117–1133
4. Ankerst M, Kastenmüller G, Kriegel H-P, Seidl T (1999) 3D shape histograms for similarity search and classification in spatial databases In: Hartmut GR, Papadias D, Lochovsky F (eds) Proceedings of the 6th international symposium on spatial database, vol 1651. Springer Berlin Heidelberg, Hong Kong China
5. Atmosukarto I, Leow WK, Huang Z (2005) Feature combination and relevance feedback for 3D model retrieval. *IEEE proceedings of the 11th international multimedia modelling conference*. IEEE, pp 334–339
6. Bakhadyrov I, Jafari MA (1999) Inertia tensor as a way of feature vector definition for one-dimensional signatures *IEEE SMC '99*. Proceedings international conference on systems, man, and cybernetics, vol 2. IEEE Tokyo, pp 904–909
7. Barrios JM, Bustos B (2011) Automatic weight selection for multi-metric distances. *ACM*. In proceeding of the 4th international conference on similarity search and applications (SISAP'11). ACM New York, NY, USA, pp 61–68
8. Basri R, Costa L, Geiger D, Jacobs D (1998) Determining the similarity of deformable shapes. *Sci Vis Res* 38:2365–2385, Elsevier
9. Behley J, Steinhage V, Cremers AB (2012) Performance of histogram descriptors for the classification of 3D laser range data in urban environments *IEEE*. International conference on robotics and automation (ICRA). IEEE Saint Paul, MN, pp 4391–4398
10. Belongie S, Malik J, Puzicha J (2002) Shape matching and object recognition using shape contexts. *IEEE Trans Pattern Anal Mach Intell* 24(4):509–522
11. Bespalov D, Regli WC, Shokoufandeh A (2006) Local feature extraction and matching partial objects. *Comput Aided Des* 38(9):1020–1037, Elsevier
12. Biasotti S, Giorgi D, Spanuolo M, Falcidieno B (2006) Size functions for 3D shape retrieval. In: Polthier K, Sheffer A (eds) Eurographics symposium on geometry processing. The Eurographics Association Cigliari, Sardinia, Italy, pp 239–242
13. Bober M (2001) MPEG-7 visual shape descriptors. *IEEE Trans Circ Syst Video Technol* 11(6):716–719
14. Bu S, Han P, Lui Z, Han J, Lin H (2015) Local deep feature learning framework for 3D shape. *Comput Graph* 46:117–129
15. Bustos B, Keim DA, Saupé D, Schreck T, Vranić DV (2005) Feature-based similarity search in 3D object databases. *ACM Comput Surv* 37(4):345–387

16. Cerri A, Biasotti S, Giorgi D (2007) K-dimensional size functions for shape description and comparison. IEEE. 14th International Conference on Image Analysis and Processing, vol IEEE Modena, pp 795–800
17. Chen L, McAuley JJ, Feris RS, Caetano TS, Turk M (2009) Shape classification through structured learning of matching measures. Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on. IEEE Miami, FL., pp 365–372
18. Chen D-Y, Tian X-P, Shen Y-T, Ouhyoung M (2003) On visual similarity based 3D model retrieval. The Eurographics Association and Blackwell Publishers 22(3):223–232
19. Chung FRK (1997) Spectral graph theory. Regional conference series in mathematics American mathematical society. Published for the conference board of the mathematical sciences Washington DC
20. Dorai C, Jain AK (1997) Shape spectrum based view grouping and matching of 3D free-form objects. IEEE Trans Pattern Anal Mach Intell 19(10):1139–1146
21. Dos Santos FJ (2007) Retrieval of 3D models using partial matching Universidade Tecnica de Lisboa. PhD. Thesis
22. Dutağacı H, Sankur B, Yemez Y (2005) Transform-based methods for Indexing and retrieval of 3D objects. IEEE proceedings of the fifth international conference on 3D - digital imaging and modeling. computer society. (3DIM'05). IEEE Washington DC, USA., pp 188–195
23. ElNaghy H, Hamad S, Khalifa E (2013) Taxonomy for 3D content-based object retrieval methods. IJARRAS 14(2):412–446
24. El Wardani D, El Mostafa D, Taddonji C (2012) Improving 3D shape retrieval methods based on Bag-of-feature approach by using local codebooks. Int J Futur Gener Commun Netw 5(4):29–38
25. El-Mehalawi M, Miller RA (2003) A database system of mechanical components based on geometric and topological similarity. Comput Aided Des 35:83–94, Elsevier
26. Funkhouser T, Min P, Kazhdan M, Chen J, Halderman A, Dobkin D (2003) A search engine for 3D models. ACM Trans Graph 22(1):83–105
27. Gal R, Cohen-Or D (2006) Salient geometric features for partial shape matching and similarity. ACM Trans Graph 25(1):130–150
28. Godil A, Wagan AI (2011) Salient local 3D features for 3D shape retrieval. In IS&T/SPIE electronic imaging. Int. Soc. Opt. Photonics 78640S - 78640S-78648
29. Guo Y, Bennamoun M, Sohel F, Lu M, Wan J, Kwok NM (2015) A comprehensive performance evaluation of 3D local feature descriptors. Int. J Comput Vis 116:66–89
30. Healy DM, Rockmore DN, Kostelec PJ (2003) FFTs for the 2-sphere-improvements and variations. J Fourier Anal Appl 9(4):341–385, Springer
31. Heczko M, Keim DA, Saube D, Vranic DV (2002) Methods for similarity search os 3D databases (verfahren zur ähnlichkeitssuche auf 3D-objekten). Datenbank-Spektrum (DBSK) 2(2):54–63
32. Heider P, Pierre-Pierre A, Li R, Mueller R, Grimm C (2012) Comparing local shape descriptors. Vis Comput 28(9):919–929
33. Hilaga M, Shinagawa Y, Kohmura T (2001) Topology matching for fully automatic similarity estimation of 3D shapes. ACM. SIGGRAPH '01. Proceedings of the 28th annual conference on computer graphics and interactive techniques. ACM New York, NY, USA, pp 203–212
34. Hoffmann CM (1989) Geometric and solid modeling. Morgan Kaufmann Pub <https://www.cs.purdue.edu/homes/cmh/distribution/books/geo.html>, p 338
35. Horn BKP (1984) Extended gaussian images. Proc IEEE 72(12):1671–1686
36. Huang P, Starck J, Hilton A (2007) A study of shape similarity for temporal surface sequences of people. International conference on 3-D imaging and modeling - 3DIM pp 408–418
37. Icke I (2004) Content based 3D shape retrieval a survey of state of the art. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.303.353&rep=rep1&type=pdf>. Accessed May 2014
38. Ikeuchi K (1981) Recognition of 3-D objects using the extended gaussian image. Proc. 7th International joint conference on artificial intelligence, pp 595–600
39. Johnson AE, Hebert M (1999) Using spin images for efficient object recognition in cluttered 3D scenes. IEEE Trans Pattern Anal Mach Intell 21(5):433–449
40. Kang SB, Ikeuchi K (1993) The complex EGI: a new representation for 3-D pose determination. IEEE Trans Pattern Anal Mach Intell 5(7):707–721
41. Kazhdan M, Chazelle B, Dobkin D, Furkhouser T, Rusinkiewicz S (2003) A reflective symmetry descriptor for 3D models. ACM J Algorithmica 38(1):201–225
42. Kazhdan M, Funkhouser T, Rusinkiewicz S (2003b) Rotation invariant spherical harmonic representation of 3D shape descriptors eurographics symposium on geometry processing. The eurographics association, pp 156–164
43. Keim DA (1999) Efficient geometry-based similarity search of 3D spatial databases ACM. In proceedings of the international conference on management of data (SIGMOD'99), vol 28(2). ACM New York, NY, pp 419–430

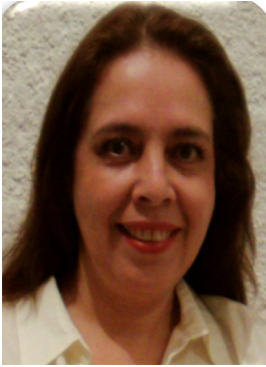
44. Koenderink JJ, Van Doorn AJ (1992) Surface shape and curvature scales. *ACM Image Vis Comput* 10(8): 557–565
45. Körtgen M, Park G-J, Novotni M, Klein R (2003) 3D shape matching with 3D shape contexts. In proceedings of the 7th central european seminar on computer graphics Budmerice, Slovakia, pp 34–43
46. Laga H, Takahashi H, Nakajima M (2006) Spherical wavelet descriptors for content based 3D model retrieval. *IEEE proceedings of the international conference on shape modeling and applications*. IEEE. Computer Society Washington DC, USA, pp 15–25
47. Lamdan Y, Wolfson HJ (1988) Geometric hashing: a general and efficient model-based recognition scheme. *IEEE proceeding of the second international conference on computer vision*. IEEE Tampa, FL, pp 238–249
48. Latecki LJ, Lakämper R, Eckhardt U (2000) Shape descriptors for non-rigid shapes with a single closed contour. *IEEE Conference on computer vision and pattern recognition (CVPR) vol 1 Hilton Head Island, SC*, pp 424–429
49. Lazebnik S, Cordelia, Ponce J (2005) A sparse texture representation using affine invariant regions. *IEEE proceeding computer society. Comp Vision Recog Pattern* 27(8):1265–1278
50. Leifman G, Meir R, Tal A (2005) Semantic-oriented 3d shape retrieval using relevance feedback. *Vis Comput* 21(8–10):865–875, Springer-Verlang
51. Li C, Hamza B (2013) A multiresolution descriptor for deformable 3D shape retrieval. *Vis Comput* 29(6–8):513–514
52. Ling H, Jacobs DW (2007) Shape classification using the inner-distance. *IEEE Trans Pattern Anal Mach Intell* 29(2):286–299
53. Liu W, He Y-j (2006) Multi-level spherical moments based 3D model retrieval. *J Zhejiang Univ (Sci)* 7(9): 1500–1507
54. Liu Y, Zha H, Qin H (2006a) The generalized shape distributions for shape matching and analysis. *IEEE proceedings of the international conference on shape modeling and applications (SMI'06)*. Computer Society Matsushima Japan
55. Liu Y, Zha H, Qin H (2006b) Shape topics: a compact representation and new algorithms for 3D partial shape retrieval. *IEEE proceedings of the computer society conference on computer vision and pattern recognition* 2025–2032
56. Lowe DG (1999) Object recognition from local scale-invariant features. *IEEE proceedings of the seventh international conference on computer vision* 2:1150–1157
57. Lowe DG (2004) Distinctive image feature from scale-invariant keypoints. *J Comput Vis* 60(2):91–110, Kluwer Academic Publishers
58. Mateus LDC (2010) Spectral tools for unsupervised modeling of articulated objects from multiple-view videos. Institut National Polytechnique de Grenoble. PhD. Thesis
59. Mikolajczyk K, Schmid C (2006) A performance evaluation of local descriptors. *IEEE Trans Pattern Anal Mach Intell* 27(10):1615–1630
60. Mitsumoto H, Tamura S, Okazaki K, Kajimi N, Fukui Y (1992) Reconstruction using mirror images based on a plane symmetry recovering method. *IEEE Trans Pattern Anal Mach Intell* 14(9):941–946
61. Mortara M, Patané G (2002) Shape-covering for skeleton extraction. World scientific publishing company. *Int J Shape Model* 8(2):139–158
62. Mortensen EN, Deng H, Shapiro L (2005) A SIFT Descriptor with Global Context. *Proc IEEE Conf Comput. Vis Pattern Recogit (CVPR)* 1:184–190
63. Novotni M, Klein R (2001a) A geometric approach to 3D object comparison. *IEEE international conference on shape modeling and applications*. IEEE, pp 166–175
64. Novotni M, Klein R (2003) 3D Zernike descriptors for content based shape retrieval. *ACM proceedings of the eighth ACM symposium on solid modeling and applications*. ACM Seattle, Washington, USA, pp 216–225
65. O'Hara S, Draper B (2010) Introduction to the bag of features paradigm for image classification and retrieval
66. Ohbuchi R, Minamitani T, Takei T (2005) Shape-similarity search of 3D models by using enhanced shape functions. *Int J Comput. Appl Technol (IJCAT)* 23(2/3/4):1–27
67. Ohbuchi R, Otagiri T, Ibato M, Takei T (2002) Shape-similarity search of three-dimensional models using parameterized statistics. *IEEE proceedings of the 10th pacific conference on computer graphics and applications*. Computer Society. IEEE Washington DC, pp 265–274
68. Osada R, Furkhouser T, Chazelle B, Dobkin D (2002) Shape distributions. *ACM Trans Graph* 21(4):807–832
69. Papadakis P, Pratikakis I, Theoharis T, Perantonis S (2010) PANORAMA: a 3D shape descriptor based on panoramic views for unsupervised 3D object retrieval. *Int J Comput Vis* 89(2–3):177–192

70. Paquet E, Rioux M (1997) Nefertiti: a query by content software for three-dimensional models databases management. IEEE proceedings international conference on recent advances in 3D digital imaging and modeling Ottawa, Ont, pp 345–352
71. Paquet E, Rioux M (1998) Content-based access of VRML libraries. In: Horace HSI, Smeulders AWM (eds) Springer, Berlin Heidelberg. IAPR. Proceeding international workshop MINAR' 98 multimedia information analysis and retrieval, vol 1464, Lectures notes in computer science. Springer, Berlin Heidelberg Hong Kong, China, pp 20–32
72. Paquet E, Rioux M (1999) The MPEG-7 standard and the content-based management of three-dimensional data: a case study. IEEE Proc Int Conf Multimed Comput Syst (ICMCS'99) 1:375–380. Florence
73. Paquet E, Rioux M, Murching A, Naveen T, Tabatabai A (2000) Description of shape information for 2-D and 3D objects. Signal Process Image Commun 16:103–122, Elsevier
74. Passalis G, Kakadiaris IA, Theoharis T (2004) Efficient hardware voxelization. IEEE Proc Comput Graph Int 374–377
75. Peleg S, Werman M, Rom H (1989) A unified approach to the change of resolution: space and gray-level. IEEE Trans Pattern Anal Mach Intell 11(7):739–742
76. Podolak J, Shilane P, Golovinsky A, Rusinkiewicz S, Funkhouser T (2006) A planar-reflective symmetry transform for 3D shapes. ACM proceedings of international conference on computer graphics and interactive techniques: ACM SIGGRAPH, vol 25. ACM. Press New York. USA, pp 549–559
77. Ricard J, Coeurjolly D, Baskurt A (2005) Generalizations of angular radial transform for 2D and 3D shape retrieval. Pattern Recog Letters 26(14):2174–2186
78. Saupé D, Vranić DV (2001) 3D model retrieval with spherical harmonics and moments. Springer proceedings of the 23rd DAGM-symposium on pattern recognition. Springer. Berlin, Germany, pp 392–397
79. Schmitt W, Sotomayor JL, Telea A, Silva CT, Comba JLD (2015) A 3D shape descriptor based on depth complexity and thickness histograms in graphics, patterns and images (SIBGRAPI), 2015 28th SIBGRAPI conference on, pp 226–233
80. Shamir A, Scharf A, Cohen-or D (2004) Enhanced hierarchical shape matching for shape transformation. World scientific publishing company. Int J Shape Model 9(2):203–222
81. Shih J-L, Lee C-H, Wang JT (2007) A New model retrieval approach based on the elevation descriptor. Pattern Recogn 40(1):283–295, Elsevier
82. Shilane P, Min P, Kazhdan M, Funkhouser T (2004) The princeton shape benchmark. IEEE proceedings of the shape modeling international. IEEE Washington, DC, USA, pp 167–168
83. Shinagawa Y, Kunii TL (1991) Constructing a reeb graph automatically from cross sections. IEEE Comput Graph Appl 11(6):44–51
84. Sun J, Ovsjanikov M, Guibas L (2009) A concise and provably informative multi-scale signature based on heat diffusion. Eurographics association and Blackwell publishing. Proc Symp Geom Proc 28(5):1383–1392
85. Sundar H, Silver D, Gagvani N, Dickinson S (2003) Skeleton based shape matching and retrieval. IEEE proceedings of the shape modeling international. IEEE Computer Society Washington, DC, pp 130–142
86. Suzuki MT, Kato T, Otsu N (2000) A similarity retrieval of 3D polygonal model using rotation invariant shape descriptors. IEEE pp 2946–2952
87. Taimouri V, Hua J (2014) Deformation similarity measurement in quasi-conformal shape space. Graph Model 76:57–69, Elsevier
88. Tangelder JWH, Veltkamp RC (2003) Polyhedral model retrieval using weighted point sets. Proc Int Conf Shape Model Appl 3(1):119–129
89. Tangelder JWH, Veltkamp RC (2008) A survey of content based 3D shape retrieval methods. Sci Multimed Tools Appl 39:441–471, Springer
90. Tung T, Schmitt F (2004) Augmented reeb graphs for content-based retrieval of 3D mesh model. IEEE Proc Shape Model Appl pp 157–166
91. Undurraga C, Mery D (2011) Improving tracking algorithms using saliency. In: César SM, Kim S-W (eds) Springer-Verlag Berlin Heidelberg. Progress in pattern recognition, image analysis, computer vision, and applications. Proceedings 16th iberoamerican congress (CIARP). Springer-Verlag Berlin Heidelberg Chile, pp 141–148
92. Vandeborre J-P, Couillet V, Daoudi M (2002) A practical approach for 3D model indexing by combining local and global invariants. IEEE. 1st international symposium on 3D data processing visualization transmission. IEEE Padova, Italy, pp 19–21
93. Vaxman A, Ben-Chen M, Gotsman C (2010) A multi-resolution approach to heat kernels on discrete surfaces. ACM Trans Graph 29(4):121
94. Veltkamp RC (2001) Shape matching: similarity measures and algorithms. IEEE. In shape modeling and applications, SMI. International conference. IEEE computer society, pp 188–197

95. Venkatraman V, Lee S, Kihara D (2009) Potential for protein surface shape analysis using spherical harmonics and 3D zernike descriptors. Vol 54(1–3). Springer <http://link.springer.com/article/10.1007/s12013-009-9051-x/fulltext.html>, pp 23–32
96. Vranić DV (2004) 3D model retrieval. University of Leipzig. PhD. Thesis
97. Vranić DV, Saupe D (2001a) 3D shape descriptor based on 3D fourier transform. Proceedings of the EURASIP conference on digital signal processing for multimedia communications and services (ECMCS 2001) Budapest, Hungary, pp 271–274
98. Vranić DV, Saupe D (2002) Description of 3D-shape using a complex function on the sphere. In proceedings of: IEEE Int Conf Multimed Expo 1:177–180
99. Vranić DV, Saupe D (2001b) Tool for 3D-object retrieval: karhunen-loeve transform and spherical harmonics. IEEE proceedings of the 4th workshop on multimedia signal processing. IEEE Cannes, pp 293–298
100. Yu M, Atmosukarto A, Leow WK, Huan Z, Xu R (2003) 3D model retrieval with morphing-based geometric and topological feature maps. Proc Comput Soc Conf Comput Vision Pattern Recog 2:656–661
101. Zaharia T, Prêteux F (2001b) Hough transform-based 3D mesh retrieval. Proceeding of the SPIE conference on vision geometry X, pp 175–185
102. Zaharia T, Prêteux F (2001a) 3D shape-based retrieval within the MPEG-7 framework. Proceedings of the SPIE conference nonlinear image processing and pattern analysis XII., vol 4304 San Jose, CA, pp 133–145
103. Zaharia T, Prêteux F (2002) Shape-based retrieval of 3D mesh models. IEEE proceedings international conference on multimedia and Expo. (ICME). IEEE, pp 437–440
104. Zhang C, Chen T (2001) Efficient feature extraction for 2D/3D object in mesh representation. IEEE proceedings international conference on image processing, vol 3. IEEE Thessaloniki, pp 935–938
105. Zhang L, Gao Y, Xia Y, Dai Q, Li X (2015) A fine-grained image categorization system by cellet-encoded spatial pyramid modeling. IEEE Trans Ind Electron 62(1):564–571
106. Zhang L, Han Y, Yang Y, Song M, Yan S (2013) Discovering discriminative graphlets for aerial image categories recognition. IEEE Trans Image Process 22(12):5071–5084
107. Zhang L, Joao DS, Manuel, Ferreira A (2004) Survey on 3D shape descriptors. In: POSC/EIA/59938 DR (ed) Republica Portuguesa, pp 1–28
108. Zhang D, Lu G (2004) Review of shape representation and description techniques. Pattern Recogn 37:1–19, Elsevier
109. Zhang L, Yang Y, Gao Y, Yu Y, Wang C, Li X (2014) A probabilistic associative model for segmenting weakly supervised images. IEEE Trans Image Process 23(9):4150–4159
110. Zhao X, Lu M (2013) 3D object retrieval based on PSO-K-modes method academy publisher. J Softw 8(4): 963–970
111. Zhu Z, Wang X, Bai S, Yao C, Bai X (2014) Deep learning representation using autoencoder for 3D shape retrieval in security, pattern analysis, and cybernetics (SPAC), 2014 international conference on, pp 279–284



Graciela Lara López is currently a Ph.D. student in Computer Science at the Universidad Politécnica de Madrid, Spain. Her areas of interest are in virtual reality, mainly on its application for training, 3D object modeling, and spatial mental models. She received a Ph. D. in Teaching Methodology from the Instituto Mexicano de Estudios Pedagógicos in 2012. She is an associate professor at the Computer Science Department at the CUCEI of the Universidad de Guadalajara, Mexico.



Adriana Peña Pérez Negrón received her Ph.D. in Computer Science in 2009 from the Universidad Politécnica de Madrid, Spain. Her main research interest is on the user's avatar display of nonverbal communication in Collaborative Virtual Environments. She is a research professor at the Computer Science Department at the CUCEI of the Universidad de Guadalajara, Mexico.



Angélica De Antonio Jiménez received her Ph.D. in Computer Science in 1994 from the Universidad Politécnica de Madrid, Spain. She is Director of the laboratory “Decoroso Crespo” since 1995. Her areas of interest are educational software, particularly intelligent tutoring systems; virtual and augmented reality, mainly on its application for education and training; intelligent virtual agents, specially working as pedagogical agents; human-computer interaction techniques for 3D environments; and the software engineering methods and techniques that make it possible to develop such complex systems. She is a senior lecturer of University in the Universidad Politécnica de Madrid.



Jaime Ramírez Rodríguez received his Ph.D. in Computer Science in 2002 from the Universidad Politécnica de Madrid, Spain. He is member of the laboratory “Decoroso Crespo” since 1996. His areas of interest are virtual environments for education and training, intelligent tutoring systems, ontological engineering and software architecture. He is an associate professor of University in the Universidad Politécnica de Madrid.



Ricardo Imbert Paredes received his Ph.D. in Computer Science in 2005 from the Universidad Politécnica de Madrid, Spain. He is member of the laboratory “Decoroso Crespo” since 1996. His areas of interest are software agents, cognitive architectures for agents, intelligent tutoring systems and human-computer interaction, including advanced interaction in 3D environments. He is an associate professor of University in the Universidad Politécnica de Madrid.