A VISUAL SIMILARITY- BASED 3D SEARCH ENGINE

Elmustapha Ait Lmaati[1], Ahmed El Oirrak[1], and M. N. Kaddioui[1]***

*[*1] Department of Computer Science, Faculty of Science, Cadi Ayyad University, P.O. Box 2390, Marrakech, Morocco*
Email: lmaatimustapha@yahoo.fr;,e.aitlmaati@ucam.ac.ma;. oirrak@yahoo.fr

ABSTRACT

Retrieval systems for 3D objects are required because 3D databases used around the web are growing. In this paper, we propose a visual similarity based search engine for 3D objects. The system is based on a new representation of 3D objects given by a 3D closed curve that captures all information about the surface of the 3D object. We propose a new 3D descriptor, which is a combination of three signatures of this new representation, and we implement it in our interactive web based search engine. Our method is compared to some state of the art methods, tested using the Princeton-Shape Benchmark as a large database of 3D objects. The experimental results show that the enhanced curve analysis descriptor performs well.

Keywords: Retrieval system, Search engine, 3D objects, Combination of 3D descriptors, Information retrieval

1. INTRODUCTION

Three dimensional databases are growing both in number and size around the web due to the rapid development of tools for acquisition and storage of 3D objects. Therefore, the navigation through those databases to find desirable objects is a major problem. Content based indexing is an important way to manage those large databases. Many search engines for 3D objects are available on the web. These search engines give users the possibility of navigating through the databases to visualize models in 3D space, using web navigators, and to search by visual similarity similar models for a given 3D model query. The most popular web based search engines where the links are given in references are the Princeton University search engine, the 'CCCC' Konstanz University search engine, the Fox MIIRE 3D search engine, the Ogden 3D search engine, and *Informatics and Telematics* Institute 3D search engine.

Since 1997, authors have proposed algorithms to describe 3D objects. These algorithms are categorized, in general, taking into account their representation of 3D objects. Hilaga et al. (2001) used the Reeb graph based descriptor that represents an object by a structure that captures important information about the structure of model. The authors used the geodesic distance to compute distance between graphs. Sundar et al. (2003) proposed a skeleton graph based descriptor. The authors used a thinning algorithm proposed by Gagvani & Silver (1999) on the voxelization of a solid object. In order to extract similarity between 3D objects, they compared the corresponding skeleton graphs of the objects. Hekzko et al. (2002) proposed using images based 3D descriptors. The authors generated images obtained from orthogonal projections of the object, and applied 2D shape descriptors to these images in order to compute the feature vectors for 3D objects. Chen et al (2003) and Ansari et al. (2007) extracted several views as silhouettes from the object and apply 2D shape descriptors to generate feature vectors. We note that the image based descriptors and view based descriptors are similar. However, the difference between them is that the image based approach requires the pose normalization using the CPCA; continuous principal component analysis (Vranic et al. 2001) and the views based approach do not require the normalization of 3D objects.

The voxel based approach was used by Vranic et al. (2001). This method was based on a volumetric representation of 3D objects presented in frequency domain by applying a 2D discrete Fourier transform. The authors chose the lower frequencies as components of feature vectors of 3D objects because the high frequencies are affected by noise. Kazhdan et al. (2002) proposed to apply spherical harmonics to the voxelized model so as to generate the feature vectors. This descriptor obtained in frequency domain used the rotational invariance propriety of Spherical

Harmonics. Osada et al. (2002) proposed a statistical method named D2 (shape distribution) to describe 3D objects. Vranic et al. (2001b) proposed a descriptor named Ray with Spherical Harmonics to present the extents from the centre of mass of an object to its surface in spectral domain as components of feature vectors for 3D objects.

In this paper, we present a new approach to extract similarity between 3D objects. This approach is based on a 3D closed curve that represents the 3D object. In order to extract feature vectors for our 3D model, we firstly apply CPCA (Continuous Principal Component Analysis) as pose normalization in order to align the model into canonical position. Secondly, we extract a 3D closed curve that represents this object. Finally, we extract three signatures of this curve, which are combined in a descriptor named Enhanced Curve Analysis Descriptor (ECA). We present the design of our web based search engine where we implement the proposed descriptor. We evaluate our method using the Princeton Shape Benchmark Database (Shilan et al., 2004) using measures widely used in the information retrieval community. We end with our conclusion.

2 RETRIEVAL SYSTEM

The proposed retrieval system is made up of two steps: O*ff-line* processes and *On-line* processes.

- In the *off-line* processes, the retrieval system computes feature vectors for 3D objects. It firstly aligns the model into a canonical position using CPCA, and then reconstructs the 3D closed curve that represents the 3D object. Finally, the system extracts three signatures, which are the area descriptor, the dot product descriptor, and the torsion descriptor. The system combines those descriptors to generate the enhanced curve analysis descriptor.
- In the *on-line* processes, the web server receives the query model and then compares its feature vector to the feature vectors of all the models in the database. The closest models are extracted and returned to the user in less then a second. Our system uses the Manhattan distance in order to compute the similarity between two objects.

Figure 1 shows the architecture of our retrieval system comprising the two discussed steps.

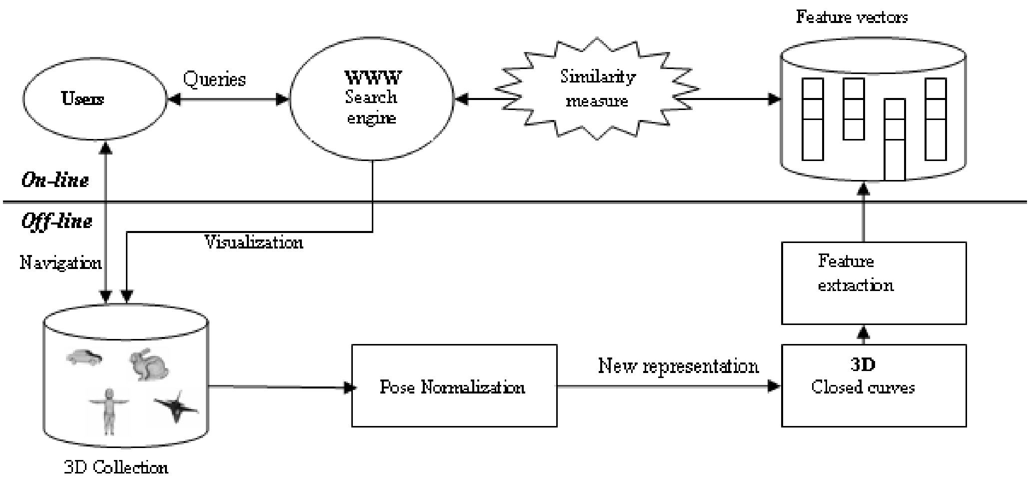

Figure 1: Retrieval system architecture

3 3D CURVE BASED REPRESENTATION FOR 3D OBJECT

In order to extract feature vectors for any 3D model, we represent the object by its 3D closed curve that we defined in our previous work (Lmaati et al., 2009). Our representation is important because many problems are posed when the surfaces of 3D objects are directly studied. The greatest problem solved by a closed curve representation is that of parameterization because 3D objects are generally given in an arbitrary topology by polygonal meshes. The proposed closed curve is homeomorphic with the spherical helix curve defined on the unit sphere. This curve can represent 3D objects without loss information on its surface (Figure 2).

We first use CPCA (Vranic et al., 2001) to achieve translation, rotation, reflection, and scale invariance. These transformations are summarised by the equation,

$$\phi(P) = s^{-1} F.V.(P - G),$$

where P is the vertex of mesh, s is the average distance between the centre of mass of the object and its surface, F is the matrix of reflection, and V is the matrix of rotation.

The ray casting triangular meshes algorithm (Bunyk et al., 1997) is used to compute the points that constitute the curve, defined as the farthest vertex intersection between directions defined by the spherical helix and the surface of the object. If there is no intersection for a direction, then we take the center of coordinates as the point intersection. The result is that we reconstruct a 3D parametric curve, which is periodic with a period T=2π, given by the following equation,

$$X(t) = \begin{cases} x(t) \\ y(t) \\ z(t) \end{cases} \qquad t \in [0, T]$$

Figure 2 shows two 3D objects and their reconstructed 3D closed curves using 4096 points.

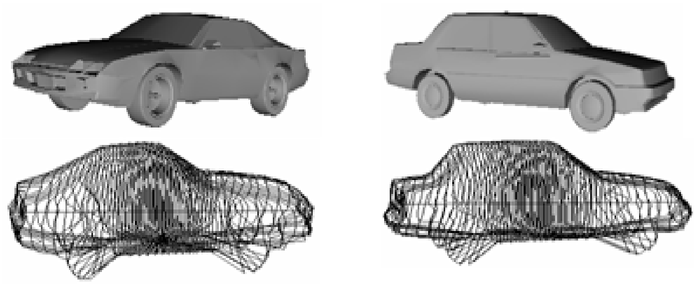

Figure 2: Models m1523 and m1524 from the PSB database (upper) and their similar 3D closed curves (lower)

4 ENHANCED CURVE ANALYSIS DESCRIPTOR

In this section, we present a new shape descriptor, which is a combination of three descriptors considered to be signatures of the reconstructed 3D closed curve. We combine the area descriptor, the dot product descriptor, and the torsion descriptor into a descriptor named enhanced curve analysis.

- **The Area Descriptor**

Let O be the centre of system coordinates, P_i the i-*th* point on the 3D closed curve that represents 3D object, P_{i+1} the (i+1)-th point on the 3D closed curve. The area of triangle OP_iP_{i+1} is defined as,

$$\text{Area}_i = \frac{1}{2} \| (OP_i \times OP_{i+1}) \|,$$

where "×" is the cross product defined for vectors. The feature vector components for this descriptor as mentioned in our previous work (Lmaati et al., 2009) are the magnitudes of the first coefficients of the fast Fourier transform of the area defined by the formula,

$$FT(\text{Area})(k) = \frac{1}{N} \sum_{n=0}^{N-1} \text{Area}_n . e^{\frac{-jk2n\pi}{N}},$$

where N is the number of points that constitute the reconstructed close curve.

-

- **The Dot Product Descriptor**

In order to capture more information about the reconstructed 3D closed curve, we propose to generate the dot product, using the absolute value that captures the angle between OP_i and P_iP_{i+1} as a descriptor. It is gives by the formula,

$$DP_i = | (OP_i \bullet P_iP_{i+1}) | .$$

The magnitudes of the first fast Fourier transform coefficients of the dot product are taken as components of the feature vector for this descriptor. They are computed using the formula,

$$FT(DP)(k) = \frac{1}{N} \sum_{n=0}^{N-1} DP_n . e^{-\frac{jk2n\pi}{N}},$$

where N is the number of points that constitute the reconstructed closed curve.

- **The Torsion Descriptor**

Torsion is an important quantity that can describe a 3D curve, being a local measure that captures more information about the geometric form of the 3D closed curve. In order to compute it at each point on the reconstructed curve, we need to express it in terms of derivatives of x, y, z.

The absolute value of the torsion is computed on each point P_i by the formula,

$$PT = | det(dX, d^2X, d^3X) | .$$

X is a coordinate of points on the surface of the object; dX is its first derivative; d^2X is its second derivative, and d^3X is its third derivative. The “det” denotes the popular determinant defined for matrices. Note that we choose to not normalise the torsion formula because we are applying the CPCA, which solves the problems of translation, rotation, reflection and scale invariance.

The feature vectors for this descriptor are composed of the first coefficients of the Fourier transform given by the formula,

$$FT(PT)(k) = \frac{1}{N} \sum_{n=0}^{N-1} PT_n . e^{-\frac{jk2n\pi}{N}},$$

where N is the number of points that constitute the reconstructed closed curve. Note that the 1D fast Fourier transform is applied to curves that are closed, and the use of the Fourier transform helps us to define descriptors that are robust with respect to noise.

- **The Enhanced Curve Descriptor**

The enhanced curve descriptor is a combination of the three descriptors described above. These descriptors cooperate so as to describe the reconstructed 3D closed curve and to give enhanced retrieval of 3D models. The feature vector for the enhanced curve descriptor (*ECAFv*) is given by the formula,

$$ECAFv = \left(\frac{Fv1(1)}{Fv1_{avr}}, \ldots, \frac{Fv1(Dim1)}{Fv1_{avr}}, \frac{Fv2(1)}{Fv2_{avr}}, \ldots, \frac{Fv2(Dim2)}{Fv2_{avr}}, \frac{Fv3(1)}{Fv3_{avr}}, \ldots, \frac{Fv3(Dim3)}{Fv3_{avr}}\right), \quad (1)$$

where *Fv1, Fv2, Fv3* are the feature vectors with dimensions *Dim1, Dim2, and Dim3* respectively for the area descriptor, the dot product descriptor, and the torsion descriptor. $Fv1_{avr}$, $Fv2_{avr}$, and $Fv2_{avr}$ are the averages of the three feature vectors computed using the following equation:

$$Fvp_{avr} = \frac{1}{Dimp} \sum_{i=1}^{Dimp} Fvp(i),$$

where p has values 1, 2, or 3. Using this method of combination, we define a new descriptor that we call Enhanced Curve Analysis Descriptor (ECA). This descriptor is effective for describing 3D objects. Experimentally, our method gives better performance, as discussed in Section 6.

5 THE VISUAL SIMILARITY BASED 3D SEARCH ENGINE

In this section, we describe the design of our web based search engine for 3D objects, which implement the ECA descriptor given in the last section. Our similarity based search for three dimensional objects is based on the idea that if two objects are similar, their 3D closed curves are similar (Figure 2). The process of 3D indexing and retrieval used by our system is summarised in Figure 3.

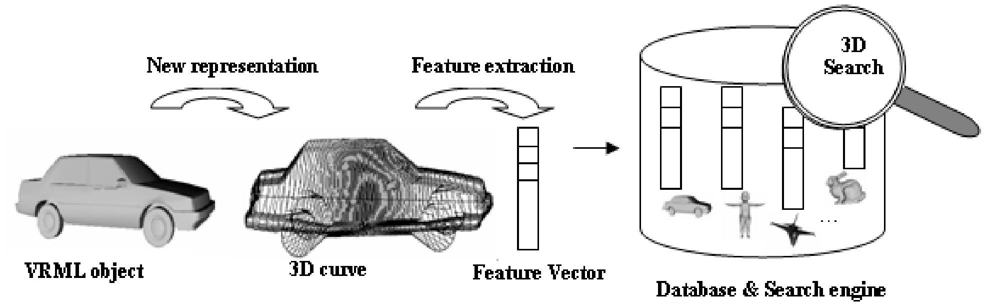

Figure 3: The process of the three dimensional indexing and retrieval

Our retrieval system is a web based application. The search engine allows users to navigate through a collection of 3D objects, select a query model, and then submit it to the web server. The system computes the Manhattan distances between the query and models in the database and reorders them in decreasing order. The most similar objects are extracted and returned to the user. The user can select another query from the retrieved models for a new search. Figure 5 shows the screen shot of the visual similarity 3D search engine. All models from the database can be visualised in the web navigator in 3D space using VRML2.0, which is the standard of the web.

The Princeton Shape Benchmark Database is used by the search engine in order to test our method (Princeton). It is a large digital database of 3D objects provided by the shape retrieval and analysis group of the Princeton University. It contains 1814 3D objects given by triangular meshes. This database was split into two sets, a train set containing 907 3D models, classified into 90 classes, and a test set containing 907 3D models, classified into 92 classes.

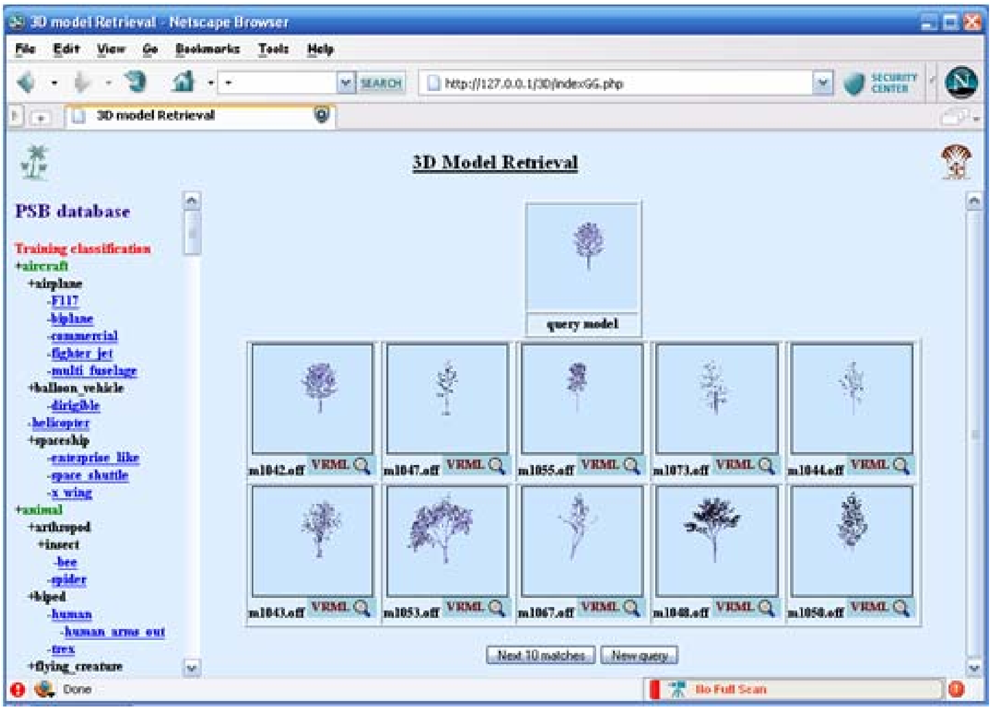

Figure 4: Screen shot of our web based search engine

Java and C/C++ languages are used to implement all the *off-line* process (generate feature vectors). The system consumes 1.3s as the average time to compute the feature vector for a model from the Princeton Shape Benchmark Database, under a Windows platform, using a 1.4 GHz Celeron M machine with 512 MB memory. The Pre-Processor language (PHP) is used to implement the *on-line* process. The Apache web server system is used to accept queries as HTTP queries and returns the response as HTTP responses to users. Our search engine answers the user in less than a second using the Princeton shape benchmark.

6 EXPERIMENTAL RESULTS

In this section, we give the measures used to evaluate the proposed method, providing a comparison of our descriptor to other methods. We test our descriptor by using the Princeton Shape Benchmark Database.

- **Method Evaluation**

In order to evaluate our method, we use measures that are widely used in the multimedia information retrieval community, Nearest Neighbour (NN), First Tier (FT), and Second Tier (ST). The recall versus precision curve (Raghavan et al., 1989) is an important tool that we use in the evaluation. For a given query Q in a class C with n models, let Rk be the number of correctly retrieved models among the K best matches. The recall is defined as a ratio of relevant models Rk to (n-1), and the precision is the ratio of the relevant results and returned results K. The First Tier measure is the same as the precision value when K is equal to $(n - 1)$, and the Second Tier is the same as the precision value when K is equal to $2(n - 1)$.

The computation of dissimilarity between two 3D objects is achieved by computing the metric distance between the feature vectors corresponding to the objects. Our system uses the Manhattan distance L_1, the Euclidian distance L_2, and the maximum distance L_{max} distance so as to choose which gives the better performance for the method.

Our method gives better performance using a curve with 4096 points, using the L_1 distance and 400 as the dimension of the feature vector of the enhanced curve analysis descriptor. Note that, we take *Dim1=150, Dim2=150* and

Dim3=100 the dimensions of the area descriptor, the dot product descriptor and the torsion descriptor respectively. Figure 5 shows a query for a sword from the Princeton shape benchmark database.

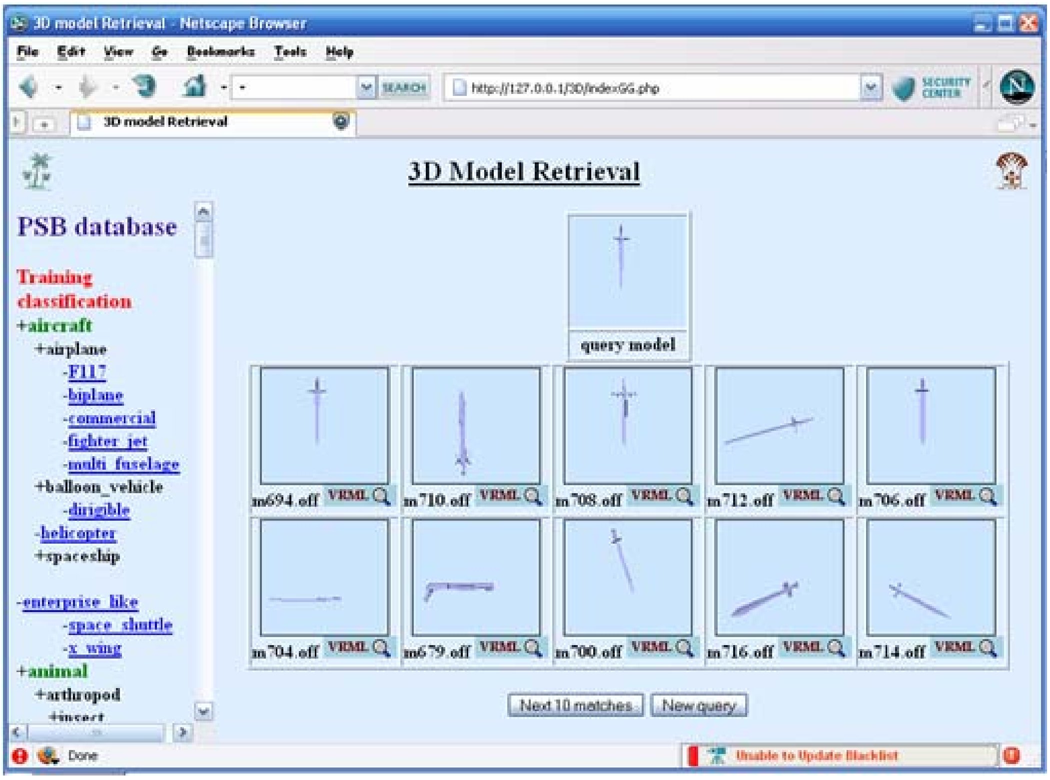

Figure 5: Query for a sword from the PSB database

- **Comparison to other descriptors**

We compare the Enhanced Curve Analysis descriptor to four other methods found in the literature. The first method used in the comparison is the descriptor 3DCA that we proposed in our previous work (Lmaati et al., 2009). The 3D curve analysis descriptor is an efficient descriptor based on the reconstructed 3D curve. The 3D curve analysis descriptor (3DCA) uses the CPCA to align models into canonical position. A 3D closed curve is used to represent the object without loss of information on the surface of the object. The authors extracted two descriptors, which are the area descriptor and the dot product descriptor and combine them in a 3D curve analysis descriptor.

The second method is the Ray with Spherical Harmonics descriptor, which was proposed by Vranic et al. (2001). The authors aligned models into canonical positions using the CPCA and compute the extents defined by the centre of object and its surface. Finally, the authors presented those extents in the spectral domain by applying spherical harmonics.

The third comparison descriptor is the Silhouette Based Feature vector (SIL), proposed by Hekzko et al. (2001). In order to extract feature vectors, the authors aligned models using the CPCA, generating three binary images as silhouettes, which are obtained by the orthogonal projections of the 3D object. The authors applied the 2D discrete Fourier transform to each silhouette and took the first coefficients of this transformation as components of a feature vector.

The Depth Buffer is an image based feature vector that is the fourth used in this comparison. It is an efficient method proposed by Hekzko et al. (2001) that aligns each model using the CPCA, computes six images in gray level, and for each image, applies the 2D Fourier transform. The magnitudes of the first coefficients are taken as a feature vector.

Figures 6 and 7 show the recall versus precision plots for all the four descriptors using the test set and the training set using the Princeton Shape Benchmark Database. Table 1 shows the measures NN, FT, ST and storage size given in bytes for the five descriptors. The Enhanced Curve Analysis descriptor outperforms three of the comparison methods: 3D Curve Analysis (3DCA) descriptor, Silhouette (SIL) descriptor, and Ray with Spherical Harmonics (RSH) descriptor. The Depth Buffer (DBD) descriptor outperforms the proposed descriptor. However, our method computes feature vectors faster than the Depth Buffer because our method is based on the 1D FFT, which has complexity of the comparison order $o(N \log(N))$, while the Depth Buffer is based on the 2D FFT, which has complexity of order $o(N^2 \log_2(N))$. It also needs more storage size. Note that N is the number of points or pixels used by the methods.

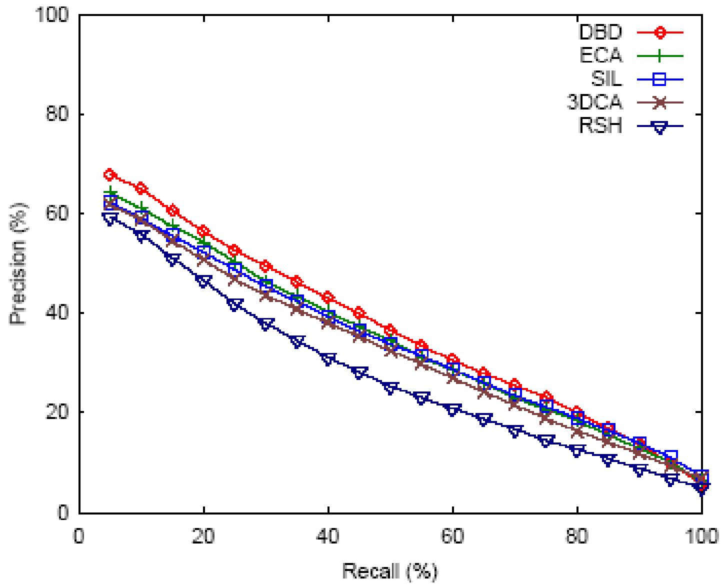

Figure 6: Recall vs. precision plots for ECA, DBD, 3DCA, SIL, and RSH descriptors using the test set from the PSB database

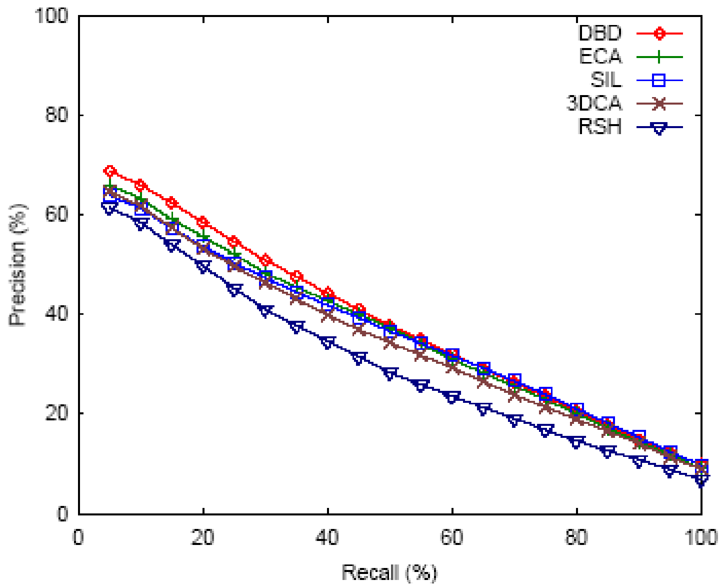

Figure 7: Recall vs. precision plots for ECA, DBD, 3DCA, SIL, and RSH descriptors using the training set from the PSB database

	Storage size	ST (%)	FT (%)	NN (%)
Depth buffer (DBD)	1752	42.1	33.2	60.9
Enhanced curve analysis (ECA)	1600	41.2	31.4	56.1
Silhouette (SIL)	1200	40.9	30.5	54.7
3D Curve analysis (3DCA)	1200	39.7	29.7	53.9
Ray with spherical harmonics (RSH)	544	34.6	25.6	51.5

Table 1: Comparison of measurements: FT, ST and NN for different descriptors using test set from PSB database.

7 CONCLUSION

In this paper, we propose a new method for describing 3D objects as well as describe a web based 3D search engine that implements our method. The enhanced curve analysis descriptor represents a 3D object using a 3D closed curve and uses a combination of three 3D descriptors extracted from the reconstructed curve. This method is implemented in our interactive search engine that allows users to search 3D models by visual similarity in a large database of 3D objects. The method is tested using the Princeton Shape Benchmark Database and evaluated using measures that are widely used in the information retrieval community. Our proposed method shows strong performance in comparison with other common similarity methods.

8 REFERENCES

Bunyk, P., Kaufman, A.E., & Silva, C.T. (1997) Simple, fast, and robust ray casting of irregular grids. *In: Scientific Visualization Conference Dagstuhl'97, Dagstuhl, Germany,* pp. 30–36

Chen, D.Y., Ouhyoung, M., Tian, X.P., & Shen, Y.T. (2003) On visual similarity based 3D model retrieval. *Eurographics, Granada, Spain,* pp. 223-232

Dejan Vranic's 3D Search Engine (Konstanz University). Retrieved from the WWW, February 25, 2009: http://merkur01.inf.uni-konstanz.de/CCCC/

Filali, T.A., Daoudi, M., & Vandeborre, J.P. (2007) A bayesian 3d search engine using adaptive views clustering. *IEEE Trans. Multimedia* 9, 78–88 (2007)

Fox MIIRE search engine. Retrieved from the WWW, February 27, 2009: http://www-rech.enic.fr/3dretrieval/

Gagvani, N. & Silver, D. (1999) Parameter-controlled VOLUME thinning. *CVGIP: Graphical Model and Image Processing,* 61(3):149-164.

Heczko, M., Keim, D.A., Saupe, D., & Vranic, D.V. (2002) Verfahren zur ahnlichkeitssuche auf 3d objekten: (methods for similarity search on 3d databases), vol. 2. *Datenbank-Spektrum*, pp. 54–63

Hilaga, M., Shinagawa, Y., Kohmura, T., & Kunii, T.L. (2001) Topology matching for fully automatic similarity estimation of 3D shapes. *ACM SIGGRAPH 28th Annual Conference on Computer Graphics and Interactive Techniques, Los Angeles,* pp. 203-212.

Informatics and Telematics Institute 3D search engine. Retrieved from the WWW, February 27, 2009: http://3d-search.iti.gr/3DSearch

Kazhdan, M. & Funkhouser, T. (2002) Harmonic 3D Shape Matching. *in SIGGRAPH 2002 Technical Sketches, San antonio, TX, July 2002*, p. 191, ACM SIGGRAPH.

Lmaati, E. A., El Oirrak, A., & Kaddioui, M.N. (2009) A 3D search engine based on 3D curve analysis. *SIViP: Signal, Image and Video processing*. DOI: 10.1007/s11760-008-0091-2

Ogden search engine. Retrieved from the WWW, February 27, 2009: http://ship.nime.ac.jp/~motofumi/Ogden/

Osada, R., Funkhouser, T., Chazelle, B., & Dobkin, D. (2002) Shape distributions. *ACM Trans. Graph., 21(4),* pp. 807–832.

Princeton University search engine. Retrieved from the WWW, February 25, 2009: http://shape.cs.princeton.edu/search.html

Raghavan, V., Bollmann, P., & Jung, G.S. (1989) A critical investigation of recall and precision as measures of retrieval system performance. *ACM Trans. Inf. Syst.,* pp. 205-229.

Shilane, P., Kazhdan, M., Min, P., & Funkhouser, T. (2004) The Princeton shape benchmark. *Shape Modeling International,Washington, DC,*pp. 167–178

Sundar, H., Silver, D., Gagvani, N., & Dickinson, S.J. (2003) Skeleton based shape matching and retrieval. *Shape Modeling International (SMI'03), IEEE Computer Society, Washington, DC,* pp. 130—142

Vranic, D. & Saupe, D.(2001) 3D shape descriptor based on 3D Fourier transform. *In: Proceedings of the EURASIP Conference on Digital Signal Processing for Multimedia Communications and Services, Budapest, Hungary*, pp. 271-274

Vranic, D.V. & Saupe, D. (2001b) 3d model retrieval with spherical harmonics and moments. *DAGM 2001. In: Radig B., Florczyk S. (eds.) Munich, Germany*, pp 392–397

(Article history: Received 1 December 2008, Accepted 25 February 2009, Available online 24 April 2009)