Scientific
Research

3D Recognition Using Affine Invariants

Ahmed El Oirrak, Mustapha Elhachloufi, Elmustapha Ait Lmaati, Mohamed Sadgal, Mohammed Najib Kaddioui

Faculty of Science Semlalia, LISI Laboratory, Marrakech, Morocco.
Email: oirrak@yahoo.fr

ABSTRACT

The size of 3D data stored around the web has become bigger. Therefore the development of recognition applications and retrieval systems of 3D models is important. This paper deals with invariants for 3D models recognition. Thus under general affine transform we propose in a first time determinants of three points to realize invariance under affinity. To solve starting point problem we needs Fourier Series (FS) to extract affine invariant descriptors, called Fourier Series Descriptor (FSD). The difference between first and second approaches: in first approach determinants are computed on cartesian coordinates directly while in the second one determinants are computed from FS coefficients to eliminate dependency on starting point. The FS are also applied on 2D slices to generate affine invariants for 3D volume. FS can be computed based on hole points of volume, but this technique. The principal advantages of proposed approaches are the possibility to handle affine transform and 3D volume. Two types of 3D objects are used in the experimentations: mesh and volume, the Princeton Shape Benchmarek (PSB) is also used to test our descriptor based on FSD.

Keywords: Affine Invariant; FSD; Mesh; Volume; 2D Slice

1. Introduction

The digital databases of 3D objects which are used in various domains (e-commerce, games, medicine, etc.) become large. Therefore, an efficient method that allows users to find similar 3D objects for a given 3D model query is becoming necessary. Content based indexing and retrieval is an important way to manage these databases. Many content based retrieval systems and search engines for 3D models are available on the web [1-3]. Many methods in this way are developed. The shape spectrum descriptor proposed by Zaharia *et al*. based on surface geometry, is recommended by MPEG-7 [4]. Filali *et al*. proposed the descriptor based on 2D views named Adaptive Views Clustering (AVC). It is a probabilistic Bayesian method that selects the most interesting views from several views of a 3D model [1]. Based on the statistics, Osada *et al*. proposed the descriptor named shape distribution (D2) [5]. Paquet *et al*. proposed the method of cord histograms [6]. The ratio area/volume is used as a feature vector to describe the 3D models by Zhang and Chen [7]. Even if this descriptor computes the feature vectors easily and quickly, it needs a high quality of mesh. Vranic and Saupe [8] proposed the method named Ray based descriptor. It uses the extents obtained from the center of mass of the model to intersection with its surface in directions which are constructed by an ico-

sahedron. This approach is not robust to noise and needs a high dimension of feature vectors, this is why the authors construct the feature vectors from the complex function on a sphere, composed with Ray based feature vectors and shading based feature vectors, presented in frequency domain by applying the spherical harmonics [9]. In order to capture the information in the interior of a 3D model, the authors proposed the descriptor named Layered Depth Spheres [10], where they use the property of spherical harmonics to achieve the rotation invariance. All previous work cited above can handle only 3D similitude (rotation, scale, translation).

This paper deals with invariants for 3D models recognition. Thus under general affine transform we propose in a first time determinants of three points to realize invariance under affinity. To solve starting point problem we needs Fourier Series (FS) to extract affine invariant descriptors, called Fourier Series Descriptor (FSD). The difference between first and second approaches: in first approach determinants are computed on cartesian coordinates directly while in the second one determinants are computed from FS coefficients to eliminate dependency on starting point. The FS are also applied on 2D slices to generate affine invariants for 3D volume. FS can be computed based on hole points of volume, but this technique This technique is obviously time-consuming.

The principal advantages of proposed approaches is the possibility to handle affine transform and 3D volume. Two types of 3D objects are used in the experimentations: mesh and volume, the Princeton Shape Benchmarek (PSB) is also used to test our descriptor based on FSD.

The paper is organized as follow: in Section 2, determinants of three points are introduced as affine invariants for 3D mesh. However to solve the starting point problem we apply Fourier series on separates coordinates to obtain what we call FSD in Section 3. In Section 4 FSs coefficients are derived from 2D slices to generate affine invariants for 3D volume. PSB database is used in Section 5 as a direct application of FSD method on 3D mesh. Section 6 presents some mathematical and experimental comparisons.

2. Invariants under 3D Affinity Using Determinant of Three Points

A general affine transform A Equation (1) is defined by the following matrix:

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \tag{1}$$

Let X be a set of 3D points, then the analytic expression Equation (2) of the transformed set $\tilde{X}$ is given by:

$$\begin{cases} \tilde{x}_i = a_{11}x_i + a_{12}y_i + a_{13}z_i \\ \tilde{y}_i = a_{21}x_i + a_{22}y_i + a_{23}z_i \\ \tilde{z}_i = a_{31}x_i + a_{32}y_i + a_{33}z_i \end{cases} \tag{2}$$

The determinant of three points Equation (3) is given by:

$$D_{ijk} = \det \begin{pmatrix} x_i & x_j & x_k \\ y_i & y_j & y_k \\ z_i & z_j & z_k \end{pmatrix} \tag{3}$$

For each three points we can verify that:

$$\tilde{D}_{ijk} = \det(A) \cdot D_{ijk} \tag{4}$$

The resulting (3×3) matrix D_{ijk} depends on choice of index i, j and k for X and $\tilde{X}$.

3. Starting Point Problem and FSD Method

The starting point problem can be seen as dependency on index of points, the following example of a 2D curve illustrates this problem: a circle can be generated by:

$$(x(t), y(t)) = (\cos t(t), \sin(t))$$

where $t \in [0; 2\pi]$.

Therefore the following equation:

$$(x(t + t_0), y(t + t_0)) = (\cos(t + t_0), \sin(t + t_0))$$

where $t \in [0; 2\pi]$ and t_0 a constant can give a circle but the order of points are not the same.

In many shape descriptors, the curve is re-parameterized by normalized arc length Equation (5), s which is defined as follows:

$$s(t) = \int (\dot{x} + \dot{y} + \dot{z}) \tag{5}$$

Arc length is preserved under similarity transforms *i.e.* the corresponding points between the original and the transformed curve have the same normalized arc length under change in orientation, translation, and differ by a scale factor under scaling. Arc length is not preserved under general affine transforms [11]. In some applications, where curves are subject to general affine transforms, arc length is replaced by affine length Equation (6) which is defined as follows and is proved to be affine invariant.

$$\tau(t) = \int \left(x(\dot{y}\ddot{z} - \dot{z}\ddot{y}) - y(\dot{x}\ddot{z} - \dot{z}\ddot{x}) + z(\dot{x}\ddot{y} - \dot{y}\ddot{x}) \right) \tag{6}$$

The analytic expression of the transformed and reparameterized set $\tilde{X}$ Equation (7) is:

$$\begin{cases} \tilde{x}(\tau) = a_{11}x(\tau) + a_{12}y(\tau) + a_{13}z(\tau) \\ \tilde{y}(\tau) = a_{21}x(\tau) + a_{22}y(\tau) + a_{23}z(\tau) \\ \tilde{z}(\tau) = a_{31}x(\tau) + a_{32}y(\tau) + a_{33}z(\tau) \end{cases} \tag{7}$$

To extract invariant we develop separately x, y and z into FSs as following:

$$\begin{cases} \text{FS}(x(\tau)) = \sum_{n=-\infty}^{n=\infty} c_n^{(x)} e^{jn2\pi\tau} \\ \text{FS}(y(\tau)) = \sum_{n=-\infty}^{n=\infty} c_n^{(y)} e^{jn2\pi\tau} \\ \text{FS}(z(\tau)) = \sum_{n=-\infty}^{n=\infty} c_n^{(z)} e^{jn2\pi\tau} \end{cases} \tag{8}$$

where

$$\begin{cases} c_n^{(x)} = \dfrac{1}{T} \int_{-T/2}^{T/2} x(\tau) e^{-j2n\pi\tau} \, d\tau \\ c_n^{(y)} = \dfrac{1}{T} \int_{-T/2}^{T/2} y(\tau) e^{-j2n\pi\tau} \, d\tau \\ c_n^{(z)} = \dfrac{1}{T} \int_{-T/2}^{T/2} z(\tau) e^{-j2n\pi\tau} \, d\tau \end{cases} \tag{9}$$

T denote the period of parameter τ.

To verify invariance take the coordinate x, the same prove can be done for coordinates y and z, then:

$$c_n^{(\tilde{x})} = \frac{1}{T}\int_{-T/2}^{T/2}\tilde{x}(\tau)e^{-j2n\pi\tau}\mathrm{d}\tau$$

$$= \frac{1}{T}\int_{-T/2}^{T/2}\left(a_{11}x(\tau)+a_{12}y(\tau)+a_{13}z(\tau)\right)e^{-j2n\pi\tau}\mathrm{d}\tau$$

$$= a_{11}\left(\frac{1}{T}\int_{-T/2}^{T/2}x(\tau)e^{-j2n\pi\tau}\mathrm{d}\tau\right)$$

$$+ a_{12}\left(\frac{1}{T}\int_{-T/2}^{T/2}y(\tau)e^{-j2n\pi\tau}\mathrm{d}\tau\right)$$

$$+ a_{13}\left(\frac{1}{T}\int_{-T/2}^{T/2}z(\tau)e^{-j2n\pi\tau}\mathrm{d}\tau\right)$$

$$= a_{11}c_n^{(x)}+a_{12}c_n^{(y)}+a_{13}c_n^{(z)}$$

Thus

$$\begin{cases} c_n^{(\tilde{x})} = a_{11}c_n^{(x)}+a_{12}c_n^{(y)}+a_{13}c_n^{(z)} \\ c_n^{(\tilde{y})} = a_{21}c_n^{(x)}+a_{22}c_n^{(y)}+a_{23}c_n^{(z)} \\ c_n^{(\tilde{z})} = a_{31}c_n^{(x)}+a_{32}c_n^{(y)}+a_{33}c_n^{(z)} \end{cases} \quad (10)$$

We can construct relative invariant, based on determinant of three FS coefficients (indexed m, n and p) as follow:

$$C_{mnp} = \begin{vmatrix} c_m^{(x)} & c_m^{(y)} & c_m^{(z)} \\ c_n^{(x)} & c_n^{(y)} & c_n^{(z)} \\ c_p^{(x)} & c_p^{(y)} & c_p^{(z)} \end{vmatrix} \quad (11)$$

Thus absolute invariant (FSD) can be derived by dividing by another relative invariant $C_{m_0 n_0 p_0}$, (m_0, n_0 and p_0 are three fixed coefficients).

$$I_{mnp} = \frac{C_{mnp}}{C_{m_0 n_0 p_0}} \quad (12)$$

4. Affine Invariants for 3D Volume Using FS Applied on 2D Slice

Typical scalar volume data is composed of a 3-D array of data and three coordinate arrays of the same dimensions. The coordinate arrays specify the x-, y-, and z-coordinates for each data point. The units of the coordinates depend on the type of data. For example, flow data might have coordinate units of inches and data units of psi. Slice planes provide a way to explore the distribution of data values within the volume by mapping values to colors. You can orient slice planes at arbitrary angles, as well as use nonplanar slices. In this case each volume V is characterized by a set of slices, each slice is a 2D image having coordinates (x, y) and color information.

To extract affine invariants we apply FS on coordinates x and y using color as parameter [12,13]. From our image we first define the set of parameterized points:

$$(x(\tau), y(\tau)), \tau = \frac{R(x,y)G(x,y)B(x,y)}{R^3(x,y)+G^3(x,y)+B^3(x,y)} \quad (13)$$

Then, we develop separately x and y into FS Equation (14) as following:

$$\begin{cases} \mathrm{FS}(x(\tau)) = \sum_{n=-\infty}^{n=\infty} c_n^{(x)}e^{jn2\pi\tau} \\ \mathrm{FS}(y(\tau)) = \sum_{n=-\infty}^{n=\infty} c_n^{(y)}e^{jn2\pi\tau} \end{cases} \quad (14)$$

We can construct relative invariant Equation (15), based on two FS coefficients (indexed m and n) as follow:

$$C_{mn} = \begin{vmatrix} c_m^{(x)} & c_m^{(y)} \\ c_n^{(x)} & c_n^{(y)} \end{vmatrix} \quad (15)$$

Absolute invariant (FSD) can be derived by :

$$I_{mn} = \frac{C_{mn}}{C_{m_0 n_0}} \quad (16)$$

In **Figures 2** and **3** we present five slices of a 3D volume and it's transformed, those volumes **Figure 1** are initially stocked into 24 slices.

In this experiment only slices 1, 8 and 27 are considered. For each one six normalized color coefficients are extracted. In **Table 1** invariant color coefficients are presented for original slices 1, 8 and 27. In **Table 2** invariant color coefficients are presented for transformed slices 1, 8 and 27. The result obtained allow to detect

(a)   (b)
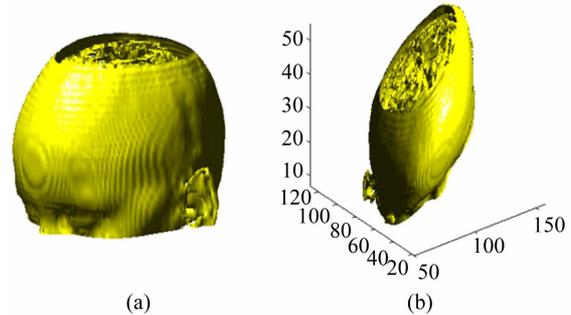
Figure 1. A 3D volume and its transformed. (a) MRI volume; (b) Transformed volume.
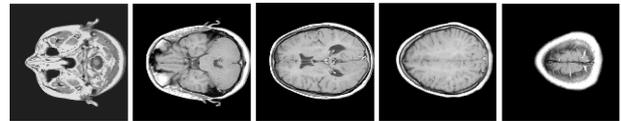
Figure 2. Five 2D slices of 3D volume.
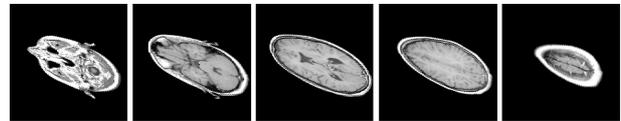
Figure 3. Five 2D slices of transformed 3D volume.

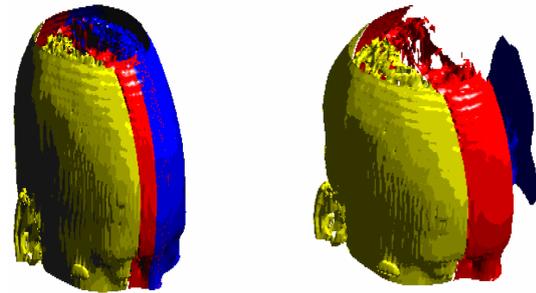**Table 1. Six invariant coefficients for each slice of V.**

slice 1	slice 8	slice 27
2.1298	3.5516	1.3817
0.3842	0.1887	0.2907
0.2912	0.3892	0.2188
0.4825	0.5262	0.6793
0.3140	0.3626	0.5440
0.4386	0.4805	0.5930

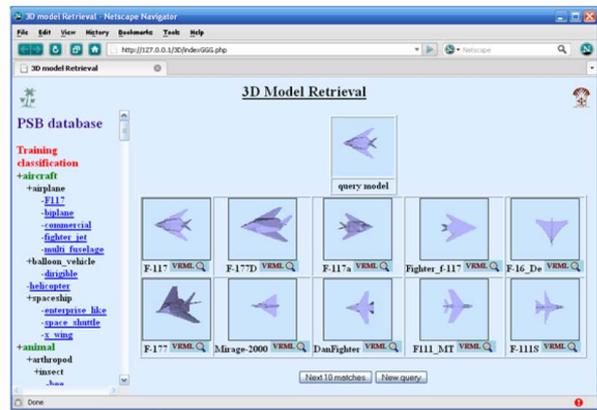**Table 2. Six invariant coefficients for each slice of $\tilde{V}$.**

slice 1	slice 8	slice 27
2.1252	3.5407	1.3851
0.3845	0.1880	0.2927
0.2892	0.3924	0.2173
0.4832	0.5315	0.6810
0.3139	0.3642	0.5477
0.4392	0.4768	0.5940

similarity between slices and consequently between 3D volume.

This technique suppose that V and $\tilde{V}$ are stocked with the same number of slices. We can use another technique based on isosurface [14]. An isosurafce define the 3D volume bounded by a particular isovalue. The volume inside contains values greater (or less) than the isovalue. The volume outside contains values less (or greater) than the isovalue. The isovalue can be an interval [min, max] in this case the isosurface is a list of all cells for which the isovalue is contained in the interval [min, max]. Also the choice of isovalue on V and $\tilde{V}$ is another problem. **Figure 4** shows some subvolumes obtained by this technique.
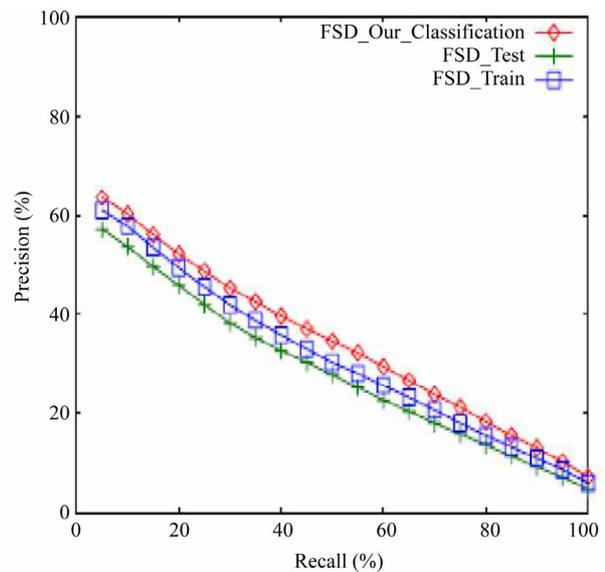
5. 3D Search Engine for 3D Mesh

In order to test our 3D descriptor **Figure 5**, we use the Princeton Shape Benchmark database (PSB) [15]. It consists of 1814 3D models given in format Object File Format (OFF), with two sets (test set and train set) of classified 3D models. The test classification consists of 907 models classified into 92 classes. The training classification consists of 907 models classified into 90 classes. We add some 3D models to the PSB database, which have been created using QSlim software [16] so as to provide models with different levels of detail. The modified PSB consists of 1889 models reclassified with geometrical aspects, our classification consists of 982 models classified into 93 classes. We use the average precision versus recall plots **Figure 6** and the First Tier (FT), Second Tier (ST), and Nearest Neighbor (NN) quantities widely used in shape retrieval community to

Figure 4. Subvolumes obtained by isosurface technique.

Figure 5. Screen-shot of the 3D search engine.

Figure 6. Precision/recall plots for our classification, test classification and train classification using PSB database.

evaluate the performance of the descriptors. 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, given by the

following formulas:

$$\text{Recall} = Rk/(N-1)$$

$$\text{Precision} = Rk/K$$

The First Tier is the same as precision value when K is equal to n − 1, the Second Tier is also the same as precision value when K is equal to 2(n − 1) and the Nearest Neighbor measure is the percentage of the closest matches that belong to the same class as the query. Obviously, an ideal score is 100, and higher scores represent better results.

Figure 6 shows the recall versus precision of test set, the training set and PSB Database. After training and testing the precision becomes higher using the hole database which mean a considerable amelioration and best performance of FSD method.

6. Comparison

6.1. Mathematical Comparison

Feature vectors for 3D model retrieval can be based also on 3D moments. In [17] a complete set of orthogonal 3D Zernike polynomials is proposed based on 3D moments. Let $\rho(x, y, z)$ be a local continuous density function, for example, this can be 1 inside voxels belonging to an object and 0 in free space. Traditional 3D moments μ_{pqr} are defined Equation (17) as:

$$\mu_{pqr} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^p y^q z^r \rho(x, y, z)\,\mathrm{d}x\mathrm{d}y\mathrm{d}z \qquad (17)$$

But those moments can not handle affine transform, so we propose separable 1D moment Equation (18).

$$\begin{cases} m_n^{(x)} = \int_{-\infty}^{\infty} \tau^{(n)} x(\tau)\,\mathrm{d}\tau \\ m_n^{(y)} = \int_{-\infty}^{\infty} \tau^{(n)} y(\tau)\,\mathrm{d}\tau \\ m_n^{(z)} = \int_{-\infty}^{\infty} \tau^{(n)} z(\tau)\,\mathrm{d}\tau \end{cases} \qquad (18)$$

We can verify as we have done in Section 3 that those moments are affine invariants.

6.2. Experimental Comparison

We compare our approach to the descriptor named Ray with spherical harmonic (RSH), proposed by Vranic and Saupe [21]. In order to extract the feature vector for a 3D model the authors use the continuous principal component analysis to align the model into the canonical position. Then they extract the maximal extents from the center of mass of the model to its surface, finally the spherical harmonic is applied to represent these rays in the frequency domain. The second descriptor used in comparison is the Silhouette based feature vectors (SIL), proposed by Heczko *et al.* [18]. This approach aligns

models using the CPCA (Continuous Principle Component Analysis), capture shape characteristics of models in three monochromes images. Then the authors extract the three contours and for each contours they apply the discrete Fourier Transform in order to present feature vectors in spectral domain. The third descriptor used in the comparison is the descriptor named Depth Buffer (DBD), it is an efficient image-based descriptor proposed by Heczko *et al.* [18]. It needs the CPCA to align the model in a canonical position and scale into the canonical unit cube. Six grey-scale images are rendered using parallel projection. Then the authors apply 2D Fourier Transform, and their feature vectors are composed in the low frequency coefficients. **Figure 7** shows the average precision versus recall plots comparing the four descriptors, where the dimension of the DBD, SIL, FSD and RSH are 438, 300, 300 and 136 respectively. **Table 3** provides comparative performance measures of the different descriptors (DBD, SIL, FSD and RSH). It is clear from **Table 3** that the proposed approach gives lower performance than DBD and SIL. However, it is worth mentioning that DBD and SIL methods are 2D view based approaches.
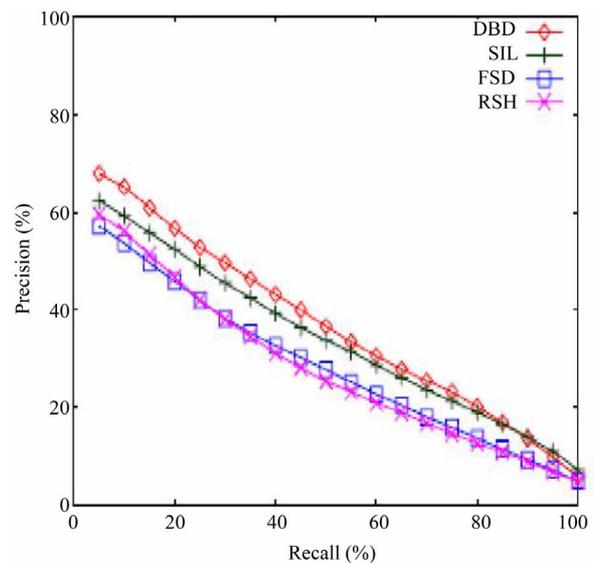

Figure 7. Precision/recall plots comparing FSD method to RSH, SIL and DBD using PSB database.

Table 3. Comparison of measurements: NN, FT and ST for different methods using PSB database.

	ST (%)	FT (%)	NN (%)
DBD	44.7	34.2	61.3
SIL	43.6	32.6	55.7
FSD	38.9	28.4	52.3
RSH	38.4	27.8	53.3

7. Conclusion

This paper advocate the use of affine invariants to describe 3D objects. Those 3D descriptors are based on determinants and Fourier series. The principal advantages of our description is that similarity is achieved without aligning models orientation by CPCA (Continuous Principal Component Analysis) and a general affine transform is considered. CPCA can gives erroneous alignments for example aircrafts with long and small wings can be aligned in tow different canonical positions. In future the 3D search engine can be extended for medical databases (MRI magnetic resonance imaging), this data typically contains a number of slice planes taken through a volume, such as the human body. One of the key difficulty when using proposed invariants is the lack of clear interpretation of the invariant's meaning relatively to the shape. It may interesting to study the behavior of these invariants for families of synthetic shapes controlled by only a few parameters.

REFERENCES

[1] T. F. Ansary, M. Daoudi and J. P. Vandeborre, "A Bayesian 3D Search Engine Using Adaptive Views Clustering," *IEEE Transaction on Multimedia*, Vol. 9, 2007, pp. 78-88. doi:10.1109/TMM.2006.886359

[2] E. A. Lmaati, A. El Oirrak and M. N. Kaddioui, "3D Model Retrieval Based on 3D Discrete Cosine Transform," *International Arab Journal of Information Technology*, Vol. 7, No. 3, 2010, pp. 264-271.

[3] E. A. Lmaati, A. El Oirrak and M. N. Kaddioui, "A 3D Search Engine Based on 3d Curve Analysis," *Signal, Image and Video Processing*, Vol. 4, No. 1, 2009, pp. 89-98.

[4] MPEG-7 Video Group, "Information Technology Multimedia Content Description Interfac*e*," Part 3: Visual, ISO/ IEC FCD, 15938:3/N4062, MPEG-7, 2001.

[5] R. Osada, T. Funkhouser, B. Chazelle and D. Dobkin, "Matching 3D Models with Shape Distributions," *Proceedings of the International Conference on Shape Modeling & Applications*, Genova, 2001, pp. 154-166.

[6] E. Paquet and M. Rioux, "A Query by Content Software for Three-Dimensional Models Databases Management," *Conference on Recent Advances* 3*-D Digital Imaging and Modeling*, Ottawa, 12-15 May 1999, pp. 345-352.

[7] D. Y. Chen, M. Ouhyoung, X. P. Tian and Y. T. Shen, "On Visual Similarity Based 3D Model Retrieval," *Eurographics*, Vol. 22, No. 3, 2003, pp. 223-232.

[8] D. V. Vranic and D. Saupe, "3D Model Retrieval with Spherical Harmonics and Moments," B. Radig and S. Florczyk, Eds., *Proceedings of the* 23*rd DAGM-Symposium on Pattern Recognition*, Munich, 2001, pp. 392-397.

[9] D. V. Vranic and D. Saupe, "Description of 3D Shape Using a Complex Function on the Sphere," *IEEE International Conference on Multimedia ICME*, Vol. 1, 2002, pp. 177-180.

[10] D. V. Vranic, "An Improvement of Rotation Invariant 3D-Shape Descriptor Based on Functions on Concentric Spheres," *IEEE International Conference on Image Processing*, 14-17 September 2003, pp. 757-760.

[11] J.-L. Dugelay, A. Baskurt and M. Daoudi, "3D Object Processing: Indexing Compression and Watermarking," John Wiley Sons Inc., New York, 2008.

[12] A. Eloirrak, M. Daoudi and D. Aboutajdine, "Affine Invariant Descriptors Using Fourier Series," *Pattern Recognition Letters*, Vol. 23, No. 10, 2002, pp. 1109-1118. doi:10.1016/S0167-8655(02)00027-2

[13] A. Eloirrak, M. Daoudi and D. Aboutajdine, "Affine Invariant Descriptors for Color Images Using Fourier Series," *Pattern Recognition Letters*, Vol. 24, No. 9, 2003, pp. 1339-1348. doi:10.1016/S0167-8655(02)00375-6

[14] Udeepta, D. Bordoloi and H.-W. Shen, "Space Efficient Fast Isosurface Extension for Large Datasets," *IEEE Visualization*, Washington DC, 19-24 October 2003, pp. 201-208.

[15] P. Shilane, M. Kazhdan, P. Min and T. Funkhouser, "The Princeton Shape Benchmark," *Shape Modeling International*, 2004, pp. 345-352.

[16] M. Garland and P. Heckbert, "Surface Simplification Using Quadric Error Metrics," *Proceedings of the* 24*th Annual Conference on Computer Graphics and Interactive Techniques*, 3-8 August 1997.

[17] M. Novotni and R. Klein, "3D Zernike Descriptors for Content Based Shape Retrieval," *Proceedings of the Eighth ACM Symposium on Solid Modeling and Applications*, Seattle, 16-20 June 2003, pp. 216-225. doi:10.1145/781606.781639

[18] M. Heczko, D. A. Keim, Saupe and D. Vranic, "Verfahren zur Ahnlichkeitssuche auf 3D Objekten (Methods for Similarity Search on 3D Databases)," *Datenbank-Spektrum*, Vol. 2, No. 2, 2002, pp. 54-63.