3D Search and Retrieval using Krawtchouk Moments

Petros Daras [1], Dimitrios Tzovaras [1], Athanasios Mademlis [2],
Apostolos Axenopoulos [1], Dimitrios Zarpalas [1] and Michael. G. Strintzis [1,2]

[1]Informatics and Telematics Institute
1st Km Thermi-Panorama Road,
Thessaloniki 57001, Greece
Email: daras@iti.gr

[2]Information Processing Laboratory
Electrical and Computer Engineering Department
Aristotle University of Thessaloniki, Greece
540 06 Thessaloniki, Greece

Abstract

In this paper a combination of a rotation invariant method with a method which utilizes rotation normalization is proposed. Both methods used are based on 2D/3D Krawtchouk moments. The first method is an extension of that which was originally introduced in [1] and utilizes 2D Krawtchouk moments while the second method was originally introduced in [2] and utilizes 3D Krawtchouk moments.[1]

1 Introduction

The huge amount of 3D models available and the increasingly important role of multimedia in many areas such as medicine, engineering, architecture, graphics design etc, showed up the need for efficient data access in 3D model databases. An important question arises, is how to search efficiently for 3D objects into many freely available 3D model databases. A query by content approach seems to be the simpler and more efficient way.

2 Spherical 3D Trace Transform Approach

Given a 3D object, its volumetric binary function is calculated $f_b(\boldsymbol{x})$, where $\boldsymbol{x} = (x, y, z)$ and $x, y, z \in [0, 2N)$, by voxelizing the whole object, which is defined as:

$$f_b(\boldsymbol{x}) = \begin{cases} 1, & \text{when } \boldsymbol{x} \text{ lies within the 3D model's volume,} \\ 0, & \text{otherwise.} \end{cases}$$

In order to achieve translation invariance the 3D object's center of mass is calculated and the model is translated so as its center of mass coincides with the coordinates system origin. Afterwards, the maximum distance $maxD$ of the most distant voxel from the origin is found and the model is scaled by the factor $1/maxD$, hence scaling invariance is also accomplished.

Then, every eight neighboring voxels are grouped, forming a bigger one and $f_b(\boldsymbol{x})$ is transformed to the integer volumetric function of the model $f(\boldsymbol{x})$, which takes values from 0, none of the eight voxels lying inside the object's volume, to 8, all of them lie inside, and $\boldsymbol{x} = (x, y, z)$, $x, y, z \in [0, N)$. This transformation denotes that more attention is given to the voxels lying inside the object's volume, which characterize more reliable the 3D object.

2.1 Decomposition of $f(\boldsymbol{x})$

The next step involves the decomposition of $f(\boldsymbol{x})$ into planes. Each plane in the 3D space can be fully described by the spherical coordinates (ρ, θ, ϕ) of the point on which the plane is tangential to a sphere originated from the center of the coordinate system. Imagine concentric spheres, simulated by icosahedra whose triangles are subdivided in many smaller equilateral triangles. The barycenters of these triangle are considered to be the characteristic (tangential) points for the planes.

The intersection of each plane with the object's volume provides a spline of the object, which can be treated as a 2D image with dimensions $N \times N$. Consider a 2D functional F, which is applied to this 2D image, producing a single value. Let us assume that the result of that functional when applied to all splines, will be a function whose domain is the set of the aforementioned points, and its range is the results of the functional. The mathematical expression of that transformation can be written as:

$$F\{f(\boldsymbol{x})\} = g(\rho, \theta, \phi)$$

Restricting to different values of ρ, $g(\rho, \theta, \phi)$ can be considered as a set of functions $g_\rho(\theta, \phi)$ whose domain is con-

[1]This work was supported by the PENED2003 project of the Greek Secretariat of Research and Technology

centric spheres. Now, let T be a functional which can be applied to a function defined on a sphere, producing a single value. Then, the result of the T functional to every $g_\rho(\theta,\phi)$ is a vector with length equal to the number of the spherical functions. The Krawtchouk moments were used as F functionals and the Spherical Fourier Transform as T functional.

2.2 Rotation Invariance Requirements

In order to produce rotation invariant descriptor vectors two requirements should be met. Imagine that the model is rotated, hence $f(\boldsymbol{x})$ is rotated, producing $f'(\boldsymbol{x}) = f(\boldsymbol{A}\boldsymbol{x})$, where $\boldsymbol{A}$ a 3D rotation matrix. The splines that will be derived from $f'(\boldsymbol{x})$ will be the same with the ones derived from $f(\boldsymbol{x})$, but the characteristic points of the planes will also be rotated by the same rotation matrix. This transformation can be translated to a rotation of $g_\rho(\theta,\phi)$ by $\boldsymbol{A}$. This problem can be settled by using a rotation invariant T functional. Moreover, the planes that are perpendicular to the axes of rotation, will be rotated around their characteristic point, resulting to a rotated version of their 2D image. Thus, the F functionals should also be rotation invariant.

2.3 2D Krawtchouk Moments

Krawtchouk moments [3] is a set of moments formed by using Krawtchouk polynomials as the basis function set. The n^{th} order classical Krawtchouk polynomials are defined as:

$$K_n(x;p,N) = \sum_{\kappa=0}^{N} a_{\kappa,n,p} x^\kappa \Rightarrow$$

$$\Rightarrow K_n(x;p,N) ={}_2 F_1(-n,-x;-N;\frac{1}{p}) \quad (1)$$

where $x, n - 0,1,2,\ldots,N$, $N>0$, $p \in (0,1)$, ${}_2F_1$ is the hypergeometric function defined as:

$${}_2F_1(a,b;c;z) = \sum_{\kappa=0}^{\infty} \frac{(a)_\kappa (b)_\kappa}{(c)_k} \frac{z^\kappa}{\kappa!}$$

and $(a)_\kappa$ is the Pochhammer symbol given by:

$$(a)_\kappa = a(a+1)\ldots(a+\kappa-1) = \frac{\Gamma(a+\kappa)}{\Gamma(a)}$$

where $\Gamma(.)$ is the gamma function.

For each $\hat{f}(i,j)$ with spatial dimension $N \times N$, the Krawtchouk moment invariants can be defined using the classical geometric moments:

$$M_{km} = \sum_{i=0}^{N-1}\sum_{j=0}^{N-1} i^k j^m f(i,j)$$

The standard set of geometric moment invariants, which are independent of rotation [4] can be written as:

$$\nu_{km} = \sum_{i=0}^{N-1}\sum_{j=0}^{N-1} [icos\xi + jsin\xi]^k [jcos\xi - isin\xi)]^m f(i,j)$$

where $\xi = (1/2)tan^{-1}\frac{2\mu_{11}}{\mu_{20}-\mu_{02}}$ and μ are the central moments:

$$\mu_{pq} = \sum_{i=0}^{N-1}\sum_{j=0}^{N-1} (i-\bar{x})^p (j-\bar{y})^q \hat{f}_t(i,j), \quad p,q = 0,1,2,\ldots$$

The value of ξ is limited to $-45^o \le \xi \le 45^o$. In order to obtain the exact angle ξ in the range of 0^o to 360^o modifications described in detail in [5] are required.

Following the analysis described in [3], the rotation invariant Krawtchouk moments are computed by:

$$\tilde{Q}_{km} = [\rho(k)\rho(m)]^{-(1/2)} \sum_{i=0}^{N-1}\sum_{j=0}^{N-1} a_{i,k,p_1} a_{j,m,p_2} \nu_{ij} \quad (2)$$

where the coefficients $a_{\kappa,n,p}$ can be determined by (1), and

$$\rho(n) = \rho(n;p,N) = (-1)^n \left(\frac{1-p}{p}\right)^n \frac{n!}{(-N)_n} \quad (3)$$

In this paper, parameter p of Krawtchouk polynomials has been selected to be $p = 0.5$

2.4 Spherical Fourier Transform

Spherical harmonics [6] are special functions on the unit sphere, generally denoted by $Y_{lm}(\boldsymbol{\eta})$, where $l \ge 0$, $|m| \le l$ and $\boldsymbol{\eta}$ is the unit vector in $\mathcal{R}^3$, $\boldsymbol{\eta} = [cos\phi sin\theta, sin\phi sin\theta, cos\theta]$. Using this notation, $g_\rho(\theta,\phi)$ can be rewritten as $g_\rho(\boldsymbol{\eta})$.

These functions form a complete orthonormal set on the unit sphere:

$$\sum_{i=1}^{N_s} Y_{lm}(\boldsymbol{\eta}_i) Y_{kj}(\boldsymbol{\eta}_i) = \delta_{lk}\delta_{mj} \quad (4)$$

where N_s is the total number of sampled points on the unit sphere (in our case the number of the equilateral triangle barycenters of the icosahedron). Hence, each function $g_\rho(\boldsymbol{\eta})$ can be expanded as an infinite Fourier series of spherical harmonics:

$$g_\rho(\boldsymbol{\eta}_i) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} \alpha_{lm} Y_{lm}(\boldsymbol{\eta}_i), \qquad i = 1,\ldots,N_s \quad (5)$$

where the expansion coefficients α_{lm} are determined by:

$$\alpha_{lm} = \sum_{i=1}^{N_s} g_\rho(\boldsymbol{\eta}_i) Y_{lm}(\boldsymbol{\eta}_i) \Delta\boldsymbol{\eta} \quad (6)$$

where $\Delta\boldsymbol{\eta}$ is the area of each triangle, hence $\Delta\boldsymbol{\eta} = \frac{4\pi}{N_s}$, since all the equilateral triangles have the same area and each icosahedron is assumed to be of unit radius. The overall vector length of α_{lm} coefficients with the same l:

$$A_l^2 = \sum_m \alpha_{lm} \tag{7}$$

is preserved under rotation and this is the reason why the quantities A_l are known as the rotationally invariant shape descriptors. In the proposed method, for each l the corresponding A_l is a spherical functional T.

2.5 Descriptor Extraction

For each F functional, a descriptor vector with length $L{\cdot}N_\rho$, where $N_\rho = 20$ and $L = 26$, is produced. In our experiments the first four Krawtchouk moments ($\tilde{Q}_{00}$, $\tilde{Q}_{10}$, $\tilde{Q}_{11}$, $\tilde{Q}_{20}$) were used as F functionals and four descriptor vectors were formed ($\mathbf{D}_{3DTrace00}$, $\mathbf{D}_{3DTrace10}$, $\mathbf{D}_{3DTrace20}$ and $\mathbf{D}_{3DTrace11}$, respectively).

3 3D Krawtchouk moments Approach

In this section the necessary steps followed so as to obtain descriptor vectors based on the 3D Krawtchouk moments are given.

3.1 Rotation Estimation

An essential part of the approach contains a novel combination of the two dominant rotation estimation methods, PCA and VCA [7], which have been proposed so far in the literature.

The VCA method achieves more accurate rotation estimation results than PCA when the 3D objects are composed of large flat areas. Otherwise, PCA produces better results than VCA. The proposed fully automatic approach tracks wrong rotation estimated objects produced either from PCA, or from VCA, and selects the most appropriate one.

In this paper for every model, rotation normalization is estimated using both PCA and VCA. Then, the volume of the bounding boxes parallel to principal axes are computed and the method which leads to minimum volume is chosen.

3.2 Extraction Of Krawtchouk Descriptors

In [3], Yap *et al.* introduced Krawtchouk moments and Krawtchouk moment invariants for image analysis, 2D object recognition and region based feature extraction (2D case), based on Krawtchouk polynomials. Their work was extended in 3D case [2] and the discrete Weighted 3D Krawtchouk moments were defined. A short description of this extension is presented in the sequel.

As it was mentioned earlier (Section 2), $f(\boldsymbol{x})$ is the volumetric representation of the 3D object. Then, the 3D Krawtchouk moments of order (n+m+l) of f, are defined as:

$$\begin{aligned}\bar{Q}_{nml} \quad = \quad & \sum_{x=0}^{N-1}\sum_{y=0}^{N-1}\sum_{z=0}^{N-1} \bar{K}_n(x;p_x,N-1)\times \\ & \times \bar{K}_m(y;p_y,N-1)\bar{K}_l(z;p_z,N-1)\times \\ & \times f(x,y,z)\end{aligned} \tag{8}$$

where

$$\bar{K}(x;p,N) = K_n(x;p,N)\sqrt{\frac{w(x;p,N)}{\rho(n;p,N)}} \tag{9}$$

and

$$w(x;p,N) = \left(\begin{array}{c} N \\ x \end{array}\right) p^x(1-p)^{N-x} \tag{10}$$

The 3D Krawtchouk moments can then be used to form the descriptor vector of every object. Specifically, the descriptor vector is composed of 3D Krawtchouk Moments up to order s, where s is experimentally selected to be $s = 18$.

$$\mathbf{D}_{3DKraw} = \left[\bar{Q}_{nml}\right], n+m+l = 0\ldots s \tag{11}$$

4 Matching

The first step is the normalization of each descriptor according to:

$$\tilde{D}(i) = \frac{1}{\sum_{i=1}^{T}|D(i)|}D(i)$$

where T is the number of descriptors in the descriptor vector $\mathbf{D}$, $D(i)$ is the $i-th$ element of $\mathbf{D}$, and $\tilde{D}(i)$ is $i-th$ element of the normalized vector $\tilde{\mathbf{D}}$.

Then, the well-known L1-norm defined as:

$$L_1(A,B) = \sqrt{\sum_{i=1}^{T}|\tilde{D}^A(i) - \tilde{D}^B(i)|}$$

is used for every normalized descriptor vector $\tilde{\mathbf{D}}_{3DKraw}$, $\tilde{\mathbf{D}}_{3DTrace00}$, $\tilde{\mathbf{D}}_{3DTrace10}$, $\tilde{\mathbf{D}}_{3DTrace20}$, $\tilde{\mathbf{D}}_{3DTrace11}$ and five distances are computed: $L_{3DTrace00}$, $L_{3DTrace10}$, $L_{3DTrace20}$, $L_{3DTrace11}$ and L_{3DKraw}, each one for every normalized descriptor vector. It has to be mentioned that due to the ambiguity of axis orientation after the rotation estimation that takes place for the 3D Krawtchouk approach, the distance is selected to be the minimum for every possible orientation.

The total distance is computed as follows:

$$
\begin{aligned}
L_{tot} \quad = \quad & a_1 L_{3DTrace00} + a_2 L_{3DTrace10} + a_3 L_{3DTrace20} \\
& + a_4 L_{3DTrace11} + a_5 L_{3DKraw} \qquad (12)
\end{aligned}
$$

where $a_1 = a_4 = 0.15$, $a_2 = a_3 = 0.25$ and $a_5 = 0.2$. These values were experimentally selected.

5 Results

The proposed method was evaluated in terms of retrieval accuracy, using the Princeton Shape benchmark (PSB) database which consists of 1814 3D objects. The performance of the proposed method against the other 16 competitive ones which took part in the SHREC contest, was proved to be among the best. The results published by the contest organizers have shown that the proposed method clearly outperforms the other methods if we take into account the first 10% of the retrieved results and it is among the first 3-4 methods concerning the overall performance. Also, it should be clearly stated that the proposed method is based on a native 3D descriptor extraction algorithm.

Very useful conclusions can be derived by examining the Mean Normalized Cumulated Gain (MNCG) and the Mean Normalized Discounted Cumulated Gain (MNDCG) graphs. Both graphs visualize the performance of the retrieval methods as a function of the retrieved results. However, MNDCG applies a discount factor to devaluate late-retrieved results and, thus, it is an appropriate user-oriented evaluation metric for retrieval applications. Our method is always in the first three positions based on MNDCG and in the first four based on MNCG. It should be noticed that our method is ranked first using both MNCG and MNDCG, considering the first 10% of the retrieved results. That means that the proposed method first retrieves the more relevant to the query 3D objects.

Specifically, based on MNDCG, the proposed method is ranked first for 5% and 10% of the retrieved results, second for 25% and third for 50% and 100%. Based on MNCG, the proposed method is first at 5%, third at 25% and forth at 50% and 100%. However, by examining the MNCG graph, methods ranked in the second, third and forth position after 50% of the retrieved results change consecutively.

The proposed method is ranked third with respect to Mean First and Second Tier measures, if only the highly relevant objects considered as similar. However, if marginally relevant objects are considered as similar too, it is ranked 6-th and 7-th respectively, although the differences between the methods from 3rd to 7th position are close to 0.5%. The main reason for these results derives from the fact that marginally relevant objects are usually semantically and functionally similar rather than geometrically similar, while the proposed method does take into account only geometrical information. In addition, the proposed method is ranked third based on Mean Dynamic Average Recall metric, which measures the overall performance of the retrieval method. The Mean ADR takes into account both highly and marginally relevant objects with different weights. The results based on ADR are presented in Table 1. These results clearly rank the proposed method among the best in native 3D content-based search.

Rank	Participant	RunNr	Mean ADR
1	Makadia et al.	2	0.54986260
2	Makadia et al.	1	0.54084843
3	*Daras et al.*	*1*	*0.52424060*
4	Chaouch et al.	1	0.50018275
5	Pratikakis et al.	1	0.49523294
6	Shilane et al.	3	0.49371490
7	Zaharia et al.	1	0.49247277
8	Shilane et al.	2	0.48770607
9	Chaouch et al.	2	0.42156762
10	Shilane et al.	1	0.39706558
11	Makadia et al.	3	0.39249521
12	Makadia et al.	4	0.37667266
13	Laga et al.	1	0.32631385
14	Laga et al.	2	0.30619973
15	Jayanti et al.	2	0.26785165
16	Jayanti et al.	3	0.23702210
17	Jayanti et al.	1	0.23020707

Table 1: Ranking Results based on Mean Average Dynamic Recall (ADR) value. The Second and Third column show the name of the first author and the serial number of the method respectively

References

[1] P.Daras, D.Zarpalas, D.Tzovaras and M.G.Strintzis: “3D Model Search and Retrieval Based on the Spherical Trace Transform”,*IEEE International Workshop on Multimedia Signal Processing (MMSP 2004)*, Siena, Italy, October 2004.

[2] A. Mademlis, A. Axenopoulos, P. Daras, D. Tzovaras and M. G. Strintzis, “3D Content-based Search Based on 3D Krawtchouk Moments”, *In Proc. of Third International Symposium on 3D Data Processing, Visualization and Transmission* (3DPVT 2006), accepted for publication.

[3] P-T. Yap, R. Paramesran, “Image Analysis by Krawtcouk Moments”, *IEEE Trans. On Image Processing*, Vol. 12, No.11 pp. 1367-1377, November 2003.

[4] M.K.Hu, “Visual Pattern Recognition by Moment Invariants”, *IRE Trans. on Information Theory*, Vol 8, pp. 179-197, 1962.

[5] M. R. Teague, “Image analysis via the general theory of moments”, *Journal of Optical Society of America*, vol. 70, pp. 920930, 1979.

[6] D.W.Ritchie, “Parametric Protein Shale Recognition”, *PhD Thesis*, University of Aberdeen, 1998.

[7] J. Pu, and K. Ramani “ An Approach to Drawing-Like View Generation From 3D Models”, *In Proc. of IDETC/CIE 2005, ASME 2005*