

# Generalized Radon Transform Based 3D Feature Extraction for 3D Object Retrieval

P. Daras<sup>1</sup>, D. Zarpalas<sup>2</sup>, D. Tzovaras<sup>2</sup> and M. G. Strintzis<sup>1,2</sup> *Fellow, IEEE*

<sup>1</sup>Information Processing Laboratory  
Electrical and Computer Engineering Department  
Aristotle University of Thessaloniki  
540 06 Thessaloniki, Greece

<sup>2</sup>Informatics and Telematics Institute  
1st km Thermi-Panorama Road  
570 01 Thessaloniki, Greece  
Email: daras@iti.gr

## Abstract

The present paper proposes a novel method for 3D model content-based search based on the 3D Generalized Radon Transform. A set of descriptor vectors is extracted, which represent significant shape characteristics and completely invariant in terms of translation, scaling and rotation. Experimental results show that the proposed method can be used for 3D model search and retrieval in a highly efficient manner.

## 1 Introduction

Determining the similarity between 3D objects is a challenging task in content-based recognition, retrieval, clustering and classification. Its main applications have traditionally been in computer vision, mechanical engineering, education, e-commerce, entertainment and molecular biology.

Many methods have already been proposed dealing with 3D shape matching. A descriptor based on the 3D Discrete Fourier Transform is introduced by Vranic and Saupe [1]. Kazhdan [2] describes a reflective symmetry descriptor associating a measure of reflective symmetry to every plane through the model's centroid. Ohbuchi *et al.* [3] employ shape histograms that are discretely parameterized along the principal axes of inertia of the model. In [4] a fast querying-by-3D-model approach, is presented which utilizes both volume-based (the binary 3D shape mask) and edge-based (the set of paths outlining the shape of the 3D object) descriptors. The descriptors chosen seem to mimic the basic criteria that humans use for the same purpose. In [5] another approach to 3D shape comparison and retrieval for arbitrary objects described by 3D polyhedral models, is described. The signature of an object is represented as a weighted point set that represents the salient points of the object. These weighted point sets are compared using a variation of the "Earth Mover's Distance". Additionally, the

MPEG-7 standard provides two kind of 3D descriptors [9]:

1. the MPEG-7 3D shape descriptor, which is based on the shape spectrum concept. Shape spectrum is defined as the histogram of the shape index, computed over the entire 3D surface.
2. the 2D/3D descriptor, which can be used to combine 2D descriptors representing a visual feature of a 3D object seen from different view angles.

In this paper, a novel general 3D search and retrieval method based on the Generalized Radon Transform (GRT) [6] is proposed. Two forms of the GRT are implemented: (a) the Radial Integration Transform (RIT), which integrates the 3D model's information on lines passing through its center of mass and contains all the radial information of the model, and (b) the Spherical Integration Transform (SIT), which integrates the 3D model's information on the surfaces of concentric spheres and contains all the spherical information of the model [7]. Combining RIT's and SIT's properties, a completely invariant, in terms of scaling and rotation, transform is created. Further, two descriptor vectors are produced, which "carry" information characteristic of the model. Additionally, an approach for reducing the dimension of the descriptor vectors is proposed, providing a more compact representation, which makes the procedure for the comparison of two models very efficient. A descriptor-based approach for 3D-model matching is finally defined.

The motivation for the implementation of these two types of the 3D Generalized Radon Transform stems from the fact that a point in 3D space is fully described by the parameters  $(\rho, \phi, \theta)$ , using spherical coordinates. The feature vector extracted using the SIT consists of information regarding  $\rho$  while the feature vector extracted using the RIT contains information regarding  $\phi, \theta$ . These simple but efficient descriptors do not require large storage capacity and are not computationally expensive to retrieve, because they are extracted from the actual content of the 3D models.

The paper is organized as follows. Section 2 describes mathematically the proposed transforms and their most significant properties. Section 3 presents in detail the proposed descriptor extraction method. In Section 4 the matching algorithm used is described. Experimental results evaluating the proposed method and comparing it with other methods are presented in Section 5. Finally, conclusions are drawn in Section 6.

## 2 The 3D Generalized Radon Transform

In this section the mathematic framework of the 3D Generalized Radon transform is described. All definitions and the basic properties of the proposed transform are proven using its continuous form, in order to facilitate reader's perception. These are valid for the discrete form as well.

Let  $M$  be a 3D model and  $f(\mathbf{x})$  the volumetric binary function of  $M$ , which is defined as:

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

Let also,  $\boldsymbol{\eta}$  be the unit vector in  $\mathbb{R}^3$ .

### 2.1 The Radial Integration Transform

The Radial Integration Transform  $RIT_f(\boldsymbol{\eta})$  of a 3D model's binary function  $f(\mathbf{x})$  is a function which associates to each  $\boldsymbol{\eta}$  the integral of  $f(\mathbf{x})$  on the line  $L(\boldsymbol{\eta}) = \{\mathbf{x} | \frac{\mathbf{x}}{|\mathbf{x}|} = \boldsymbol{\eta}\}$  passing through the origin:

$$RIT_f(\boldsymbol{\eta}) = \int_{\mathbf{x} \in L(\boldsymbol{\eta})} f(\mathbf{x}) d\mathbf{x} \Rightarrow \quad (1)$$

$$\Rightarrow RIT_f(\boldsymbol{\eta}) = \int_{-\infty}^{+\infty} f(\mathbf{x}) \delta\left(\frac{\mathbf{x}}{|\mathbf{x}|} - \boldsymbol{\eta}\right) d\mathbf{x} \quad (2)$$

Now, let  $h(\mathbf{x})$  be  $f(\mathbf{x})$  rotated by a rotation matrix  $\mathbf{A}$ , i.e.  $h(\mathbf{x}) = f(\mathbf{A}\mathbf{x})$ . Then:

$$RIT_h(\boldsymbol{\eta}) = RIT_f(\mathbf{A}\boldsymbol{\eta}) \quad (3)$$

In other words, the model's rotation rotates the output of the RIT transform. Further, let  $h(\mathbf{x})$  be  $f(\mathbf{x})$  scaled by a factor  $\alpha$ ,  $\alpha > 0$ , i.e.  $h(\mathbf{x}) = f(\frac{\mathbf{x}}{\alpha})$ . Then:

$$RIT_h(\boldsymbol{\eta}) = \alpha RIT_f(\boldsymbol{\eta}) \quad (4)$$

In other words, the RIT amplitude of the scaled model is only multiplied by the factor  $\alpha$ .

### 2.2 The Spherical Integration Transform

The Spherical Integration Transform  $SIT_f(\rho)$  of a 3D model's binary function  $f(\mathbf{x})$  associates to each  $\rho$  the integral of  $f(\mathbf{x})$  on the sphere  $S(\rho) = \{\mathbf{x} | |\mathbf{x}| = \rho\}$ , with center  $(0, 0, 0)$  and radius  $\rho$ :

$$SIT_f(\rho) = \int_{\mathbf{x} \in S(\rho)} f(\mathbf{x}) d\mathbf{x} \Rightarrow \quad (5)$$

$$\Rightarrow SIT_f(\rho) = \int_{-\infty}^{+\infty} f(\mathbf{x}) \delta(|\mathbf{x}| - \rho) d\mathbf{x} \quad (6)$$

Now, let  $h(\mathbf{x})$  be  $f(\mathbf{x})$  rotated by a rotation matrix, i.e.  $\mathbf{A}$ ,  $h(\mathbf{x}) = f(\mathbf{A}\mathbf{x})$ . Then:

$$SIT_h(\rho) = SIT_f(\rho) \quad (7)$$

In other words, the SIT of a 3D-model is rotation invariant. Further, let  $h(\mathbf{x})$  be  $f(\mathbf{x})$  scaled by a factor  $\alpha$ ,  $\alpha > 0$ , i.e.  $h(\mathbf{x}) = f(\frac{\mathbf{x}}{\alpha})$ . Then:

$$SIT_h(\rho) = SIT_f\left(\frac{\rho}{\alpha}\right) \quad (8)$$

In other words, the SIT of the scaled model is scaled by the same scaling factor.

### 2.3 Discrete RIT and SIT

The discrete forms of RIT and SIT will be used for the actual extraction of the shape descriptors. The discrete form of the RIT (2) is given by:

$$RIT(\boldsymbol{\eta}_i) = \sum_{j=1}^J f(\mathbf{x}_j) \delta\left(\frac{\mathbf{x}_j}{|\mathbf{x}_j|} - \boldsymbol{\eta}_i\right), i = 1, \dots, N_{RIT} \quad (9)$$

where  $N_{RIT}$  is the total number of lines with orientation  $\boldsymbol{\eta}_i$ . Similarly, the discrete form of the SIT, (6), is given by:

$$SIT(\rho_i) = \sum_{j=1}^J f(\mathbf{x}_j) \delta(|\mathbf{x}_j| - \rho_i), i = 1, \dots, N_{SIT} \quad (10)$$

where  $N_{SIT}$  is the total number of spheres with radius  $\rho_i$ . As is easily verifiable, the discrete RIT and SIT have exactly the same properties with the continuous transforms.

## 3 Descriptor Extraction Procedure

### 3.1 Preprocessing

A 3D model  $M$  is generally described by a 3D mesh. Let  $N \times N \times N$  be the size of the smallest cube bounding the mesh. The bounding cube is partitioned in  $R^3$  voxels  $\mathbf{u}_i$

with centers  $v_i$ . The discrete binary volume function  $\hat{f}(v)$  of  $M$ , is defined as:

$$\hat{f}(v) = \begin{cases} 1, & \text{when } v \text{ lies inside model's volume,} \\ 0, & \text{otherwise.} \end{cases}$$

In general, 3D models have various levels-of-detail, depending on their mesh representation, ranging from a very few to thousands of vertices and triangles. The number  $R^3$  of voxels is kept constant for all models in order to achieve robustness with respect to level-of-detail.

In order to achieve translation invariance the center of mass of the model is first calculated and then the model is translated so that the center of mass coincides with the center of the bounding cube. Scaling invariance is achieved by calculating the maximum distance  $d_{max}$  between the center of mass and the most distant voxel where  $\hat{f}(v) = 1$ . Then,  $\hat{f}(v)$  is scaled so that  $d_{max} = 1$ .

## 3.2 Descriptor Extraction

### 3.2.1 SIT-based Descriptors

In order to extract the SIT-based feature vector, equation (10) is applied to  $\hat{f}(v)$ , producing the SIT feature vector  $\mathbf{u}_S(i) = SIT_f(\rho_i)$ , where  $i = 1, \dots, N_{SIT}$  and  $\rho_i = i \cdot \tau$ , where  $\tau$  is a preselected step.

Since SIT is invariant to rotation (7) the descriptor vector  $\mathbf{u}_S(i)$  is rotation invariant and also due to preprocessing scaling and translation invariant.

### 3.2.2 RIT-based Descriptors

Since RIT is not rotation invariant (3) the Principal Component Analysis (PCA) method is used to compensate rotation.

After translation, scaling and rotation compensation, equation (9) is applied to  $\hat{f}(v)$ , producing the RIT vector with elements  $RIT_f(\eta_i)$ , where  $i \in C_R = \{1, \dots, N_{RIT}\}$  and  $\eta_i$  are the sample points.

In order to obtain a more compact representation of the information contained in the RIT vector, a clustering of  $\eta_i$  where  $i \in C_R$ , is performed. A cluster is defined as:  $Cluster(k) = \{\eta_i | |\eta_k - \eta_i| \leq d_c\}$ , where  $d_c$  is a preselected threshold,  $k \in C_{R1} \subset C_R$ ,  $C_{R1} = \{1, \dots, N_{cluster}\}$  and  $N_{cluster}$  is the total number of clusters. For each cluster the mean of its RIT values is calculated and finally the RIT feature vector  $\mathbf{u}_R(k)$ ,  $k \in C_{R1}$  is created as the set of all these mean values.

### 3.2.3 Enhanced RIT-based Descriptors

The large and important amount of information contained in the  $RIT_f(\eta)$ , or  $RIT_f(\theta, \phi)$ , using spherical coordinates since  $\eta_i = [\cos\phi_i \sin\theta_i, \sin\phi_i \sin\theta_i, \cos\theta_i]$ , can be further

exploited in order to enhance the RIT-based descriptor vectors. For this reason, an approach similar to the one introduced in [8], was followed. According to [8], a set of invariant functionals are applied to  $RIT_f(\theta, \phi)$  in order to produce a new, enhanced descriptor vector  $\mathbf{u}_{ER}$ , which represents a set of features of  $RIT_f(\theta, \phi)$ .

The most suitable set of functionals for the proposed application is the following:

$$\begin{aligned} 1. F_1^t(g) &= \max\{g(t_i)\}, & 2. F_2^t(g) &= \sum_{i=1}^N |g'(t_i)|, \\ 3. F_3^t(g) &= \sum_{i=1}^N g(t_i), & 4. F_4^t(g) &= F_1^t(g) - \min\{g(t_i)\} \end{aligned}$$

where  $i = 1, \dots, N$ ,  $g$  is a differentiable function,  $g'$  its derivative,  $t_i, i = 1, \dots, N$  are sample points for  $g$  and  $N$  is their total number. The superscript  $t$  indicates the variable, which is eliminated by the functional.

The goal is the gradual reduction of the dimensions of  $RIT_f(\theta, \phi)$ , so as to produce a compact representation, which could be, for example, a single number, the descriptor. For this reason we define:

$$G_k(\phi) = F_k^\theta[RIT_f(\theta, \phi)], G_k(\theta) = F_k^\phi[RIT_f(\theta, \phi)],$$

where  $k=1,2,3,4$ , and choose as descriptors:

$$A_{kj} = F_j^\theta[G_k(\theta)], B_{kj} = F_j^\phi[G_k(\phi)], \quad k,j=1,2,3,4.$$

A set of  $N_{ER} = 32$  descriptor values  $A_{kj}, B_{kj}$  is produced. The enhanced RIT based feature vector is defined by  $\mathbf{u}_{ER}(i) = \{A_{kj}, B_{kj}\}$ , where  $k, j = 1, 2, 3, 4$  and  $i = 1, \dots, N_{ER}$ .

## 4 Matching Algorithm

Let  $A, B$ , be two 3D models. The similarity between two models is calculated using the following formulas:

$$S_S = (1 - \sum_{i=1}^{N_S} \frac{|\mathbf{u}_{SA}(i) - \mathbf{u}_{SB}(i)|}{|\mathbf{u}_{SA}(i) + \mathbf{u}_{SB}(i)|/2}) \cdot 100\% \quad (11)$$

$$S_R = (1 - \sum_{i=1}^{N_{cluster}} \frac{|\mathbf{u}_{RA}(i) - \mathbf{u}_{RB}(i)|}{|\mathbf{u}_{RA}(i) + \mathbf{u}_{RB}(i)|/2}) \cdot 100\% \quad (12)$$

$$S_{ER} = (1 - \sum_{i=1}^{N_{ER}} \frac{|\mathbf{u}_{ERA}(i) - \mathbf{u}_{ERB}(i)|}{|\mathbf{u}_{ERA}(i) + \mathbf{u}_{ERB}(i)|/2}) \cdot 100\% \quad (13)$$

where  $S_S$  denotes the similarity of the objects using only SIT descriptors,  $S_R$  using only RIT descriptors and  $S_{ER}$  using only Enhanced RIT descriptors. The overall similarity measure is determined by:

$$S_{Total} = a_1 \cdot S_R + a_2 \cdot S_{ER} + (1 - a_1 - a_2) \cdot S_S$$

where  $a_1, a_2$  are weight factors. Our experiments presented in the sequel were performed using the values:  $a_1 = 0.45$ ,  $a_2 = 0.15$ , hence  $1 - a_1 - a_2 = 0.40$ , reflecting the experimentally determined discriminative power of each feature vector. Also, the dimension of the SIT feature vector was experimentally selected to be  $N_{SIT} = 123$ , while that of the RIT feature vector was selected to be  $N_{cluster} = 252$ .

## 5 Experimental results

The proposed method was tested using the database formed in Utrecht University and also used by Tangelder *et al.* in [5], which consists of 512 3D models in the following six categories (classes): 242 conventional airplanes, 60 delta-jets, 45 multi-fuselages, 19 biplanes, 10 helicopters and 136 other models. All the models containing in the Utrecht database are in VRML 2 format. The average number of vertices and triangles of the models is 953 and 1616 respectively. The bigger model contains 35947 vertices and 69451 triangles.

To evaluate the ability of the proposed method to discriminate between classes of objects, each 3D model was used as a query object. Our results were compared with those of the method described in [5] and with the MPEG-7 Shape 3D descriptor [9]. The retrieval performance was evaluated in terms of “precision”. Figure 1 shows the precision as a function of the number of returned models. The proposed method outperforms the method in [5], since it results in improvements in precision, ranging from 3% to 10%. Additionally, it outperforms the MPEG-7 3D shape descriptor by more than 30%. Figure 2, illustrates the results produced by the proposed method when applied to the Utrecht database. The models in the first horizontal line are the query models (each one belongs to a different class) while the rest are the first five retrieved models.

These results were obtained using a PC with a 2.4 MHz Pentium IV processor running Windows 2000. The average time needed for the extraction of the feature vectors for one 3D model is 20 seconds, while the time needed for the comparison of two feature vectors is 0.1 msec. Clearly, even though the time needed for the extraction of the feature vectors could be further improved, the retrieval performance is excellent.

## 6 Conclusions

A novel method for 3D model search and retrieval based on two Generalized Radon Transforms (GRT) was presented. A set of descriptor vectors completely invariant in terms of translation, scaling and rotation, is extracted. The proposed method is characterized by properties highly desirable for

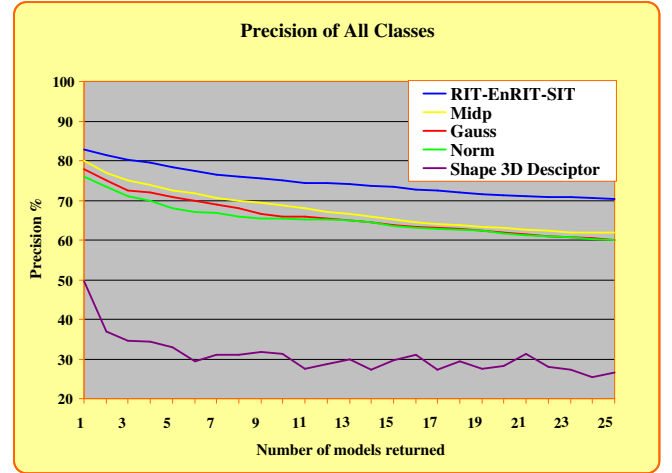


Figure 1: Comparison of the proposed method (**RIT-EnRIT-SIT**) against the Gaussian curvature (**Gauss**) [5], the normal variation (**Norm**) [5] and the midpoint method (**Midp**) [5].

efficient 3D model search and retrieval since it is very accurate, it provides robustness with respect to level-of-detail and it is very fast since matching involves simple comparison of vectors.

## References

- [1] D. V. Vranic, and D. Saupe, “3D Shape Descriptor Based on 3D Fourier Transform”, *In Proc. of the EURASIP Conference on Digital Signal Processing for Multimedia Communications and Services (ECMCS2001)*, Budapest, Hungary, September 2001.
- [2] M. Kazhdan, B. Chazelle, D. Dobkin, A. Finkelstein, and T. Funkhouser, “A Reflective Symmetry Descriptor”, *In Proc. of European Conf. on Computer Vision (ECCV2002)*, LNCS 2351, Springer, 2002, II:642-656.
- [3] R. Ohbuchi, T. Otagiri, M. Ibato, and T. Takei, “Shape-Similarity Search of Three-Dimensional Models Using Parameterized Statistics”, *Pacific Graphics 2002*, pp. 265-274, Beijing, China, (IEEE CS Press).
- [4] I. Kolonias, D. Tzovaras, S. Malassiotis and M. G. Strintzis, “Fast Content-Based Search of VRML Models Based on Shape Descriptors”, *IEEE Trans. on Multimedia*, to appear.
- [5] J. W. H. Tangelder and R. C. Veltkamp, “Polyhedral Model Retrieval Using Weighted Point Sets” *In Proc. of Inter. Conf. on Shape Modeling and Applications (SMI2003)*, 2003.
- [6] P. Toft, “The Radon Transform: Theory and Implementation,” *Ph.D. thesis*, Technical University of Denmark, 1996, <http://www.sslug.dk/pto/PhD/>

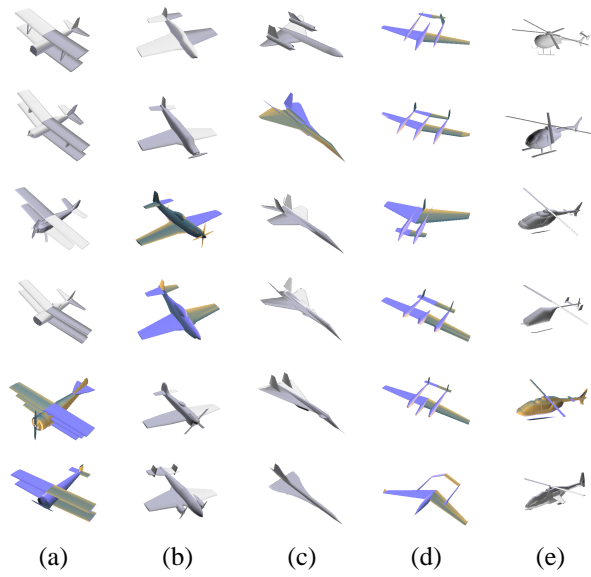


Figure 2: Query results using the proposed method in the Utrecht database. The query models are depicted in the first horizontal line.

- [7] D. Simitopoulos, D.E. Koutsonanos, M.G. Strintzis, "Robust image watermarking based on generalized radon transformations", *IEEE Trans. on Circuits and Systems for Video Technology*, Vol. 13, Issue 8, August 2003, pp. 732-745.
- [8] A. Kadyrov and M. Petrou, "The Trace Transform and Its Applications" *IEEE Transactions on Pattern Analysis and machine Intelligence (PAMI)*, VOL. 23, No. 8, pp. 811-828, August 2001.
- [9] M. Bober, "MPEG-7 Visual Shape Descriptors" *IEEE Transactions on Circuits and Systems for Video Technology (CSVT)*, VOL. 11, No. 6, pp. 716-719, June 2001.