UGC Approved Journal

IJIREEICE

ISSN (Online) 2321 – 2004
ISSN (Print) 2321 – 5526

International Journal of Innovative Research in
Electrical, Electronics, Instrumentation and Control Engineering

ISO 3297:2007 Certified

Vol. 5, Issue 8, August 2017

3D Object Retrieval by using View- Model Joint Relevance Technique

K. Sritha Rani

Department of ECE, JNTU College of Engineering (Autonomous), Anantapuramu, Andhra Pradesh, India

Abstract: Extensive research efforts have been dedicated to 3D model retrieval in recent decades. Recently, view-based methods have attracted much research attention due to the high discriminative property of multi-views for 3D object representation.3-D object retrieval has been extensively used wide range in various digital image processing applications which includes medicine, security, biometrics, genetics etc. In this work, collaboration of both view-model relevance among 3D objects for retrieval and 3D objects perception is performed based on various graph structures. Object hypergraph structure (view information) is implemented at initial stage to perceive the 3D objects in multiple views and an object graph is constructed for model data for obtaining the information about the relationship between the different features of the obtained 3D objects. Better performance and high efficiency show the supremacy of proposed work over traditional state-of-art methods.

Keywords: 3D object retrieval, Object hypergraph structure, View-model relevance, object graph.

1. INTRODUCTION

With the improvement of the 3D modeling tools and scanning devices, as well as the development of computer software and hardware technology, 3D objects become a type of important multimedia data with many applications, whose amount increases at geometric series. Generally, starting by 2000, 3D objects progressively came into center of media recovery look into. Reusing the models by retrieving 3D models in a huge database turns into an essential issue in the entertainment, computer aided design/CAM, game designing and medicinal imaging. This has prompted the innovative work of 3D shape recovery techniques.

Clearly, the 3D models can't be effectively and decisively portrayed just by text, content-based 3D model retrieval (below CB3DR) system was created. These frameworks utilize the color, texture and shape data. The shape information is represented by rotation, scaling-, translation-invariable feature descriptors based on topology, views and shape.

i. Part matching uses local features.

There exist two major research problems concerning the design of content-based multimedia retrieval systems. In the first problem, one is concerned with finding robust representation schemes describing the content of multimedia objects in terms of compact surrogates. In the context of 3D objects, content description is synonymous to 3D shape description. Several effective and efficient description algorithms have been proposed in the last decade and promising performance results have been obtained on standard benchmarks. In the second issue, one looks for computational comparability measures between descriptors that well approximate the semantic similarity between objects, in view of the grounds of client prerequisites and perceptual judgments. This second issue constitutes the principle center of the present paper.

In particular, we propose novel similarity learning algorithms for 3D object retrieval (3DOR) and test them against existing ones. With the quick advancement of web innovation, computer hardware, and software, 3D models have been generally utilized as a part of numerous applications, for example, PC illustrations, computer vision, computer aided design and medical imaging. Effectively and effectively retrieve 3D display recovery has pulled in much research consideration nowadays.

3D model retrieval methods can be divided into two categories: model-based methods and view-based methods. Early works are mainly model-based methods, in which low-level feature based methods (e.g. the geometric moment, surface distribution, volumetric descriptor and surface geometry or high-level structure-based methods are employed. Due to the requirement of 3D models, these methods are limited in the practical applications. Because of the necessity of 3D models, these methods are restricted in the practical applications.

It is noticed that the most of existing techniques isolate the model based method and the view-based method, and utilize either model information or view features for 3D object retrieval. In this work, we propose to together employ both the model and the view data for 3D object relevance estimation. In the view part, representative views are firstly selected for each object, and then the view-level distances are calculated. A object hypergraph is developed utilizing the view star extension. In the model part, the spatial structure circular descriptor is extracted and a basic graph is produced utilizing the pair wise object distances. In this way, the view information and the model data can be formulated in two

IJIREEICE
ISSN (Online) 2321 – 2004
ISSN (Print) 2321 – 5526
International Journal of Innovative Research in
Electrical, Electronics, Instrumentation and Control Engineering
ISO 3297:2007 Certified
Vol. 5, Issue 8, August 2017

graph structures. Learning on the two graphs is conducted to estimate the relevance among 3D objects, in which the graph weights can be also optimized.

2. FROM 2D PHOTOGRAPHY TO 3D OBJECT RETRIEVAL

Due to the compactness of global 3D object descriptors, their performance in capturing inter/intra class variability's are known to be poor in practice. In contrast, local geometric descriptors, even though computationally expensive, achieve relatively good performance and captures inter/intra class variability's (including deformations) better than global ones. The framework presented in this paper is based on local features and also cares about computational issues while keeping advantages in terms of precision and robustness.
Our objective is searching 3D databases of objects utilizing one or various 2D views; this plan will be alluded to as "2D-to-3D". We characterize our test set as an accumulation of single or multiple views of a similar scene or question while our exhibition set relates to an extensive arrangement of 3D models. A question, in the test set, will either be (i) multiple views of a same object, for example stereo-pair, or (ii) a 3D object model processed in order to extract several views; so ending with the "2D-to-3D" querying paradigm in both cases (iii).Gallery data are also processed in order to extract several views for each 3D object

At least two reasons motivate the use of the 2D-to-3D querying paradigm:
• The difficulty of getting "3D query models" when only multiple views of an object of interest are available. This might happen when 3D reconstruction techniques fail or when 3D acquisition systems are not available. 2D-to-3D approaches should then be applied instead.
• 3D gallery models can be manipulated via different similarity and affine transformations, in order to generate multiple views which fit the 2D probe data, so "2D-to-3D" recognition and retrieval paradigm can be achieved.

3. LITERATURE SURVEY

Elad et al. [2001] have utilized minutes (up to the seventh request) of surface focuses, abusing the way that, unique in relation to the instance of 2D images, 3D models calculation of minutes is not influenced without anyone else's input impediments. In Zhang [2001], a portrayal in view of minute invariants and Fourier change coefficients has been joined with dynamic figuring out how to consider client importance input and enhance the viability of recovery. In Novotni [2003], a technique has been displayed to figure 3D Zernike descriptors from voxelized models. 3D Zernike descriptors catch question intelligibility the outspread way and toward the path along a circle.
Not with standing, the adequacy of the approach is unequivocally dependent on the nature of the voxelization procedure. View-based portrayals utilize an arrangement of 2D views of the model and proper descriptors of their substance to speak to the 3D object shape. One issue with this approach is concerns the requirement for portrayals that are computationally tractable .In Mahmoudi [2002] and Ohbuchi [2003], various perspectives of the 3D object is taken and, for each view, the 2D views is considered. Consequently, PCA has been utilized to decrease all question perspectives to a restricted arrangement of delegate sees that are utilized to speak to the entire 3D object shape. Shape portrayals in view of factual models consider the circulation of nearby components measured at the vertices of the 3D object mesh.
The least difficult approach approximates a component dissemination with its histogram. Any metric can be utilized to process the likeness between the circulations of two models. In Vandeborre [2002], portrayal of 3D objects is caught utilizing histograms of the ebb and flow of work vertices. In Osada et al. [2002], the creators have presented shape works as disseminations of shape properties. Each distribution is approximated through the histogram of the values of the shape function.

4. PROPOSED METHOD

(A) 3D-object retrieval and recognition with hypergraph analysis

View based 3D object retrieval and recognition has become popular in practice , e.g., in CAD. It is hard to assess the difference between two objects which are represented by multiple views . In this way, view based 3-D object retrieval and recognition techniques may not perform well. In this paper, we propose a hypergraph analysis **to** deal with trouble by keeping away from the estimation of the distance between objects: specifically, we build different hypergraphs for a set 3-D objects based on 2D views. In these hypergraphs , every vertex is a object, and each edge is a cluster of views. Consequently, an edge associates various vertices. The weights of each edge based on the similarities in any two views inside the cluster. Retrieval and Recognition are performed based on the Hypergraphs. Accordingly, this technique can investigate the higher order relationship among objects and does not use the distance between objects

 DOI 10.17148/IJIREEICE.2017.5821

UGC Approved Journal

IJIREEICE

ISSN (Online) 2321 – 2004
ISSN (Print) 2321 – 5526

International Journal of Innovative Research in
Electrical, Electronics, Instrumentation and Control Engineering

ISO 3297:2007 Certified

Vol. 5, Issue 8, August 2017

(B) View-based hypergraph generation

Here the view-based hypergraph is generated by using the method in [20] and briefly introduced as follows. Let $O=\{O_i,...,O_n\}$denote then3D objects in the data set, and $V_i=\{v_{i1},,...,v_{in}\}$denote the n_i views of the i_{th} 3D object O_i. Here, we aim to explore the relevance among 3D object with multiple view information.

Generally, although multiple views can represent by more information about 3D objects, they also bring in redundant data, which may result to much computational cost and even lead to false results. we first select representative views for each 3D object, and only these representative views are employed in the 3D object retrieval process.

Given then n_i views $V_i=\{v_{i1},...,v_{ini}\}$of O_i, we conduct hierarchical agglomerative clustering (HAC) to group these views into view clusters. The HAC method is considered here due to that it can guarantee the intra cluster distance between each pair of views cannot exceed the threshold. Here the widely employed Zernike moments are used as the view features, which are robust to image rotation, scaling and translation and have been used in many 3D object retrieval tasks. The 49-D Zernike moments are extracted from each view of 3D objects. With the view clustering results, one representative view is selected from each view cluster. Here we let $V_i= \{v_{i1},...,v_{im}\}$ denote the $m_i$representative views for O_i. In our experiments, m_i mostly ranges from 5 to 20.

Hypergraph has been used in many multimedia information retrieval tasks, such as image retrieval. Hypergraph has shown its superior on high-order information representation. In our work, we propose to employ star expansion to construct an object hypergraph with views to formulate the relationship among 3D objects. Here we denote the object hypergraph as $G_H= (V_H,E_H,W_H)$. For the n objects in the dataset, there are n vertices inG_H, where each vertex represents one 3D object.

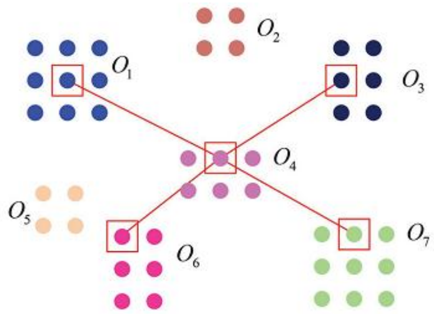

Fig. 1: An illustration of hyper edge construction. In this figure, there are seven objects with representative views. Here one view from O4 is selected as the centre view, and its four closest views are located in the figure, which are from O1, O3, O6 and O7. Then the corresponding hyper edge connects O1, O3, O4, O6 and O7

The hyper edges are generated as follows. We assume there are totally nr representative views for all n objects. We first calculate the Zernike moments-based distance between each two views, and the top K closest views can be generated for each representative view. For each representative view, one hyper edge is constructed, which connects the objects with views in the top K closest views. In our experiment, K is set as 10. Figure 3 shows an example of hyper edge generation. Generally, n_r hyper edges can be generated for G_H. The weight of one hyper edge e_H can be calculated by

$$w_H(e) = \frac{1}{K}\sum \exp(-\frac{d(V_x,V_c)^2}{\sigma^2 H}) \quad (1)$$

Where V_c is the centra view of the hyper edge, V_x is one of the top K closest view to V_c, $d(V_x,V_c)$ is the distance between V_c and V_x , and σ_H is empirically set as the median of all view pair distances.

Given the object hyper graph $G_H= (V_H,E_H,W_H)$, the incidence matrix H can be generated by

$$h(v_{H,}e_H)=\begin{cases}1 & \text{if } V_H \in e_H \\ 0 & \text{if } v_H \notin e_H\end{cases} \quad (2)$$

The vertex degree of V_H can be defined as

$$\rho(V_H) = \sum_{e_H \epsilon E_H} \omega(e_H)\, h(v_H, e_H) \quad (3)$$

The edge degree of e_H can be defined as

$$\rho(e_H) = \sum_{V_H \epsilon V_H} h(V_H, e_H) \quad (4)$$

UGC Approved Journal
IJIREEICE
ISSN (Online) 2321 – 2004
ISSN (Print) 2321 – 5526
International Journal of Innovative Research in
Electrical, Electronics, Instrumentation and Control Engineering
ISO 3297:2007 Certified
Vol. 5, Issue 8, August 2017

The vertex degree matrix and the edge degree matrix can be denoted by two diagonal matrices D_v and D_c.
In the constructed hypergraph, when two 3D objects share more similar views, they can be connected by more hyper edges with high weights, which indicates the high correlation among these 3D objects.

(C) Model-based graph generation
Given the model data of 3D objects, here we further explore the model-based object relationship. Here the spatial structure circular descriptor (SSCD) is employed as the model feature. SSCD aims to represent the depth information of the model surface on the projection minimal bounding box of the3D model. The depth histogram is generated as the feature for the 3D model. Following [21], the bipartite graph matching is conducted to measure the distance between each two 3D models, i.e.,d_{SSCD} (O_i,O_j)
Here, the relationship among 3D objects is formulated in a simple object graph structure G=(V,E,W). Here each vertex in G represents one 3D object, i.e., there are n vertices in G. The weight of an edge e (i,j) in G is calculated by using the similarity between two corresponding 3D objects O_i and O_j as

$$W(V_{i,}V_j) = exp\left(-\frac{d_{SSCD}\ \ (v_{i,}\ v_j)^2}{\sigma^2{}_s}\right) \qquad (5)$$

Where $d_{SSCD}(V_i,V_j)$is distance between O_i and O_j , and σ_s is set as the median of all modal pair distances.

(D) Learning on the joint graph
Now we have two types of formulation of relationship among 3D objects, i.e., view-based and model-based. Here these two formulations are jointly explored to estimate the relevance among 3D objects.
In this part, first we introduce the learning framework when n the view-based and model-based information are regarded with equal weight, and then we propose a jointly learning framework to learn the optimal combination weights for each modality.

(1) The initial learning framework
Here we start from the learning framework which regards different modalities, i.e., model and view, as equal. The 3D object retrieval task can be formulated as the one-class classification work as shown in[51]. The main objective is to learn the optimal pairwise object relevance under both the graph and hypergraph structure. Given the initial labeled data (the query object in our case), an empirical loss term can be added as a constraint for the learning process. The transductive inference can be formulated as a regularization as

$$arg\ \min_f\{\Omega_v(f) + \mu R(f)\} \quad (6)$$

In this formulation, f is the to-be-learnt relevance vector, Ω_v(f)is the regularizer term on the view-based hypergraph structure, Ω_M(f)is the regularizer term on the model-based graph structure, R(f) is the empirical loss. This objective function aims to minimize the empirical loss and the regularizers on the model-based graph and the view-based hypergraph simultaneously which can lead to the optimal relevance vector f for retrieval. The two regularizers and the empirical loss term are defined as follows.

The view-based hypergraph regularizer Ω_v (f) is defined as

$$\Omega_v(f) = \sum_{e_H}\sum_{u,v\epsilon v_H}\frac{w_H(eH_i)h(u,eH_i)h(v,eH_i)}{p(eH_i)}\left(\frac{f(u)}{\sqrt{p(u)}} - \frac{f(u)}{\sqrt{p(v)}}\right)^2 = \frac{1}{2}\sum_{e_H}\sum_{u,v\epsilon v_H}\frac{w_H(eH_i)h(u,eH_i)h(v,eH_i)}{p(eH_i)}\left(\frac{f^2(U)}{p(u)} - \frac{f(u)f(v)}{\sqrt{p(u)p(v)}}\right)$$
$$= f^T(\text{I-}\theta_v)\text{f}, \qquad (7)$$

Where $\theta_H$is defined as θ_H=$D_V^{\frac{-1}{2}}HWD_e^{-1}H^TD_v^{\frac{-1}{2}}$ .Here we denote Δ_H=I−θ_H, Ω_v(f) can be written as

$$\Omega_v(f) = f^T\Delta_H f \qquad (8)$$

The model-based graph regularizer Ω_M(f) is defined as

$$\Omega_v(f) = \frac{1}{2}\sum_{u,v\epsilon V} w(e_i)\left(\frac{f(u)}{\sqrt{d(u)}} - \frac{f(v)}{\sqrt{d(v)}}\right)^2$$
$$=\sum_{u,v\epsilon V} w(e_i)\left(\frac{f^2}{d(u)} - \frac{f(u)f(v)}{\sqrt{d(u)d(v)}}\right) = f^T(I - \theta_s)f, \quad (9)$$

Where θ_s=$D^{-1/2}$W$D^{-1/2}$. Here we denote Δ_s=I−θ_s, Ω_M(f) can be written as

$$\Omega_M(f) = f^T\Delta_s\text{f} \quad (10)$$

The empirical loss term R(f)is defined as

$$R(f) = \|f - y\|^2 \quad (11)$$

Where y is the initial label vector. In the retrieval process, it is defined as an n×1vector, in which only the query is set as 1 and all other components are set as 0.

Now the objective function can be rewritten as

$$\arg\min_{f} \{f^T \Delta_H f + f^T \Delta_s f + \mu\|f - y\|^2\} \quad (12)$$

f can be solved by

$$f = (I + \frac{1}{\lambda}(\Delta_H + \Delta_s))^{-1} y \quad (13)$$

f is the relevance of all the objects in the dataset with respect to the query object. A large relevance value indicates high similarity between the object and the query. The higher the corresponding relevance value is, the more similar the two objects are. With the generated object relevance f, all the objects in the dataset can be sorted in a descending order according to f.

(2) Learning the combination weights

We noted that the view information and the model information may not share the same impact on 3D object representation. In some cases, the view information may be more important, and in some other cases, the model data may play an important role. Under such conditions, we further learn the optimal weights for the view information and the model data. In this part, we introduce the learning framework embedding the combination weight learning. The objective for the learning process is composed of three parts, i.e., the graph/hypergraph structure regularizers, the empirical loss and the combination weight regularizer.

Here we let α and β denote the combination weights for view-based and model-based information respectively, where α+β=1. After adding the normal the combination weights, the objective function can be further revised as

$$arg\, min_{f,\alpha,\beta} \{\alpha f^T \Delta_H f + \beta f^T \Delta_s f + \mu\|f - y\|^2 + \eta(\alpha^2 + \beta^2)\} \quad (14)$$

Where α + β=1.

The solution for the above optimization task is provided as follows. To solve the above objective function, we alternatively optimize f and α/β. We first fix α and β, and optimize f.

Now the objective function changes to

$$arg\, min_f \{\alpha f^T \Delta_H f + \beta f^T \Delta_s f + \mu\|f - y\|^2\} \quad (15)$$

According to Eq. (13), it can be solved by

$$f = \left(I + \frac{1}{\lambda}(\alpha\Delta_H + \beta\Delta_s)\right)^{-1} y \qquad (16)$$

Then we optimize α/β with fixed f. Here we employ the Lagrangian method, and the objective function changes to

$$\underset{\alpha,\beta}{\arg\min} \left\{\alpha f^T \Delta_H f + \beta f^T \Delta_s f + \eta\left(\alpha^2 + \beta^2\right) + \xi(\alpha + \beta - 1)\right\} \qquad (17)$$

Solving the above optimization problem, we can obtain

$$\xi = -\frac{f^T \Delta_H f + f^T \Delta_s f}{2} - \eta\,, \qquad (18)$$

$$\alpha = \frac{1}{2} - \frac{f^T \Delta_H f - f^T \Delta_s f}{4\eta} \qquad (19)$$

$$\beta = \frac{1}{2} - \frac{f^T \Delta_s f - f^T \Delta_H f}{4\eta} \qquad (20)$$

The above alternative optimization can be processed under the optimal f value is achieved, which can be used for the 3D object retrieval. With the learned combination weights, the model-based and view-based data can be optimally explored simultaneously and the relevance vector f can be obtained. The main merit of the proposed method is that it jointly explore the view information and the model data of 3D objects in hypergraph/graph frameworks for 3D object retrieval.

UGC Approved Journal

IJIREEICE

ISSN (Online) 2321 – 2004
ISSN (Print) 2321 – 5526

International Journal of Innovative Research in
Electrical, Electronics, Instrumentation and Control Engineering

ISO 3297:2007 Certified

Vol. 5, Issue 8, August 2017

5. RESULTS

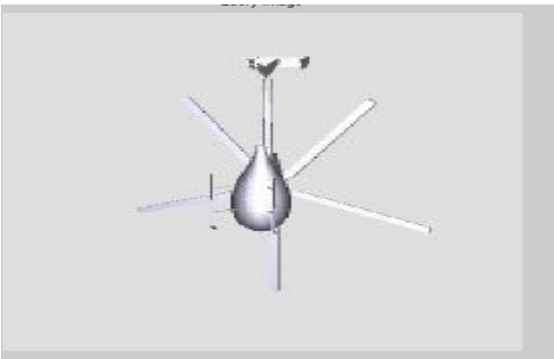

Fig. 2: Query image

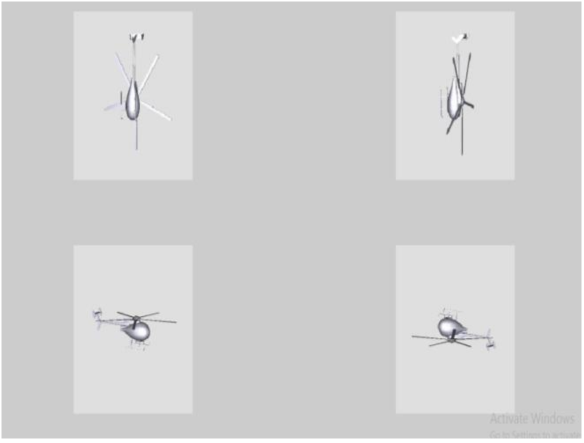

Fig 3: Retrieved image

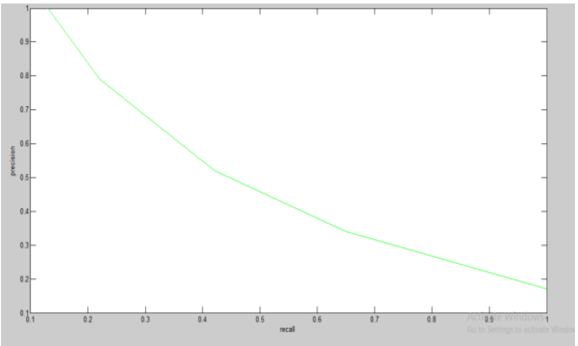

Fig 4:precision and recall curve for NTU database

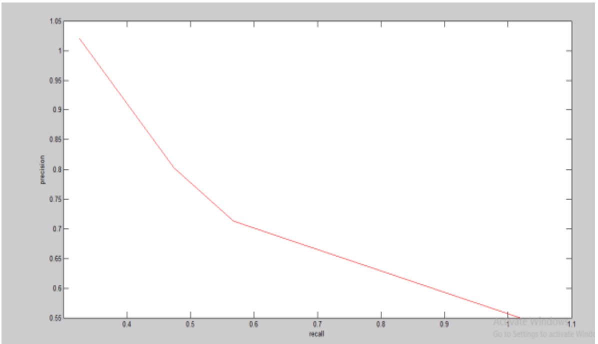

Fig 5:precision and recall curve for PSB database

 DOI 10.17148/IJIREEICE.2017.5821

UGC Approved Journal

IJIREEICE

ISSN (Online) 2321 – 2004
ISSN (Print) 2321 – 5526

International Journal of Innovative Research in
Electrical, Electronics, Instrumentation and Control Engineering

ISO 3297:2007 Certified

Vol. 5, Issue 8, August 2017

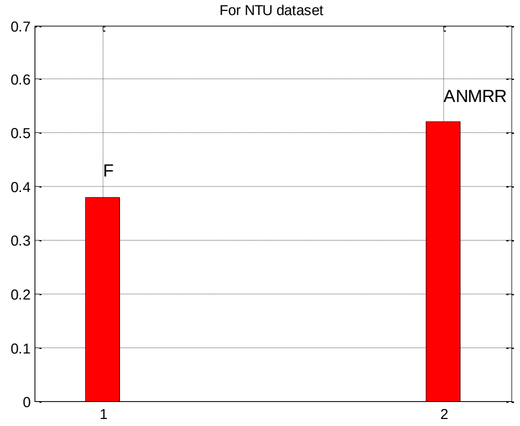

Fig. 6: F measure and ANMRR For NTU dataset

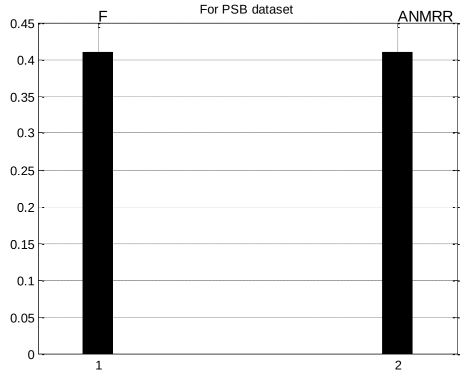

Fig. 7: F measure and ANMRR For PSB dataset

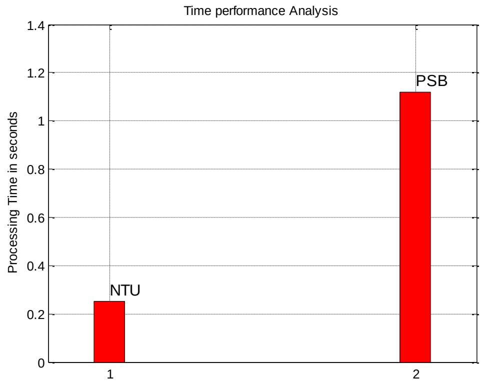

Fig. 8: Time performance analysis

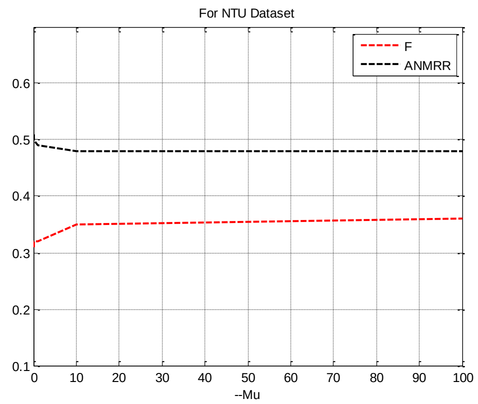

Fig. 9: For NTU dataset

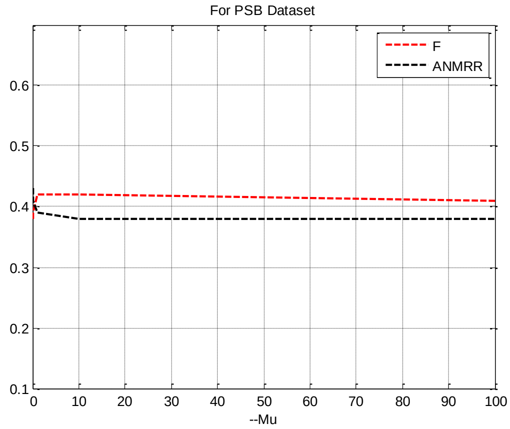

Fig. 10: For PSB dataset

Performance comparison with respect to the variation of μ of compared methods in terms of F and ANMRR in the NTU, and PSB dataset

6. CONCLUSION

Latest advancements in methods for modeling, digitizing and visualizing 3D shapes has prompted a blast in the number of accessible 3D models on the Internet and in space particular databases. In this article, we have displayed a relative assessment of 3D object representation with the end goal of recovery by content from computerized chronicles. Joint View-based and model based 3D object recognition is a fundamental subject with many rising applications. The following phase of research in this field won't just concentrate on the key advancements for see based question recovery yet in addition extend it to general spaces, which can absolutely profit by the accomplishments of view-based protest examination. Mix of both view-demonstrate importance among 3D objects for recovery and 3D objects observation is performed in light of different diagram structures. object hypergraph structure is executed at starting stage to Perceive the 3D objects in multiple views and an object graph is built for model data for acquiring the information about the connection between n the diverse features of the obtained 3D objects.

REFERENCES

[1] A. del Bimbo and P. Pala., "Content-based retrieval of 3D models," ACM Transactions on Multimedia Computing, Communications, and Applications, vol. 2, no. 1, pp. 20–43, 2006.
[2] A. Godil, H. Dutagaci, C. Akgul, A. Axenopoulos, B. Bustos, M. Chaouch, P. Daras, T. Furuya, S. Kreft, Z. Lian, T. Napoleon, A. Mademlis, R. Ohbuchi, P. L. Rosin, B. Sankur, T. Schreck, X. Sun, M. Tezuka, A. Verroust-Blondet, M. Walter, and Y. Yemez, "Shrec'09 track: Generic shape retrieval," in Proceedings of Eurographics Workshop on 3D Object Retrieval, Munich, Germany, 2009.
[3] B. Bustos, D. Keim, D. Saupe, T. Schreck, and D. Vranic, "Feature-based similarity search in 3D object databases," ACM Computing Surveys, vol. 37, no. 4, pp. 345–387, 2005
[4] J. W. H. Tangelder and R. C. Veltkamp, "A survey of content based 3D shape retrieval methods," Multimedia Tools and Applications, vol. 39, pp. 441–471, 2008
[5] Y. Yang, H. Lin, and Y. Zhang, "Content-based 3D model retrieval: A survey," IEEE Transactions on Systems, Man, and Cybernetics-Part C: Applications and Reviews, vol. 37, pp. 1081–1035, 2007.
[6] K. Lu, Q. Wang, J. Xue, and W. Pan, "3d model retrieval and classification by semi-supervised learning with content-based similarity," Information Sciences, vol. 281, pp. 703–713, 2014
[7] K. L¨ u, N. He, and J. Xue, "Content-based similarity for 3d model retrieval and classification," Progress in Natural Science, vol. 19, no. 4, pp. 495–499, 2009.
[8] A. E. Johnson and M. Hebert, "Using spin images for efficient object recognition in cluttered 3D scenes," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 21, no. 5, pp. 433–449, 1999.
[9] R. Osada, T. Funkhouser, B. Chazelle, and D. Dobkin, "Shape distributions," ACM Transactions on Graphic, vol. 21, no. 4, pp. 807–832, 2002.
[10] E. Paquet and M. Rioux, "Nefertiti: A query by content system for three dimensional model and image databases management," Image Vision Computing, vol. 17, pp. 157–166, 1999.