Mohamed El far et. al. / (IJCSE) International Journal on Computer Science and Engineering
Vol. 02, No. 04, 2010, 1327-1331

The efficiency of algorithms DATA MINING "Close + and Extraction of association rules" for indexation 2D/3D and Selection of characteristic views.

Mohamed El far[1] Lahcen Moumoun[1] Taoufiq Gadi[1] Rachid Benslimane[2]

[1]*Université Hassan 1er, Laboratoire ASTI, FST de Settat, Maroc*

simohamed.elfar@gmail.com
gtaoufiq@yahoo.fr
lahcenm@gmail.com

[2] *Université Sidi Mohamed Ben Abdellah, Laboratoire LTTI, EST de Fès, Maroc*

***Abstract*— This paper presents a new classification and search method of 3D object features views. This method is an application of algorithms:**

- **Close+ for an object views classification purpose**
- **Algorithm for extracting association rules in order to extract the characteristic view.**

We use the geometric descriptor of Zernike Moments to index 2D views of 3D object. The proposed method relies on a Bayesian probabilistic approach for search queries. The resulting outcome is presented by a collection of 120 3D models of the Princeton-based benchmark and then compared to those obtained from conventional methods.

***Keywords* — Model 3D, indexation 3D, characteristic views, Zernike Moment, CLOSE+, association rules.**

I. Introduction

The rise of new technologies and communication networks that emerged from the internet made it easier to create, use and distribute three-dimensional models. The development of modeling tools, 3D scanners, 3D graphic accelerated hardware, Web3D and so on is enabling access to three-dimensional materials of high quality.

In recent years, many systems have been proposed for efficient information retrieval from digital collections of images and videos. However, the solutions proposed so far to support retrieval of such data are not always effective in application contexts where the information is intrinsically three-dimensional. The 2D/3D shape retrieval methods are based on the computation of 2D images rendered from multiple viewpoints, considering that two models are similar when they look similar from all viewing angles. The 3D model is then indirectly represented by various 2D-shape descriptors associated with 2D views so that the 3D-shape matching is transformed into similarity measuring between 2D images. Silhouettes and depth-buffer images are the most widespread projections. A silhouette is a binary image whereas a depth buffer image contains the information of the distance between the object and the viewing plane in its pixels.

In [1], the silhouettes are encoded by their Zernike moments and Fourier descriptors. The dissimilarity between two 3D models is defined as the minimal dissimilarity of the silhouettes over all rotations and all pairs of vertices on the corresponding dodecahedrons. Filali Ansary et al [2] introduce a novel probabilistic Bayesian using a characteristic views selection. Curvature scale space (CSS) descriptor is used in [3]. Many of the 2D/3D approaches based on depth images [4, 5, 6] use the two-dimensional discrete Fourier transform (2D-DFT) as a 2D-shape signature. They differ mainly in the similarity estimation technique used and in the number of views retained. Chaouch et al [5] have introduced a technique to enhance 2D/3D shape descriptors. To take into account the dispersion of information in the views, they associate to each view a relevance index.

In this paper, we propose a data mining model of 3D object indexing based on 2D views. The goal of this model is to provide a method for optimal selection of 2D characteristic views from a 3D object, and a probabilistic Bayesian method (introduced by T. Filali Ansary and all [9] [10][11]) for 3D object indexing from these views. The data mining model based on Algorithm CLOSE+[2] for construction .The data mining model is totally independent from the 2D view descriptor used, but the 2D view descriptors should provide some properties. The data mining model has been tested with Zernike Moment 2D descriptors [15] .

Section 2 describes the different methods of 3D object indexation by views. Section 3 details our approach for characteristic views selection and search 3D object indexing and explains the principle of association rules and its algorithm. Finally, the results obtained from a collection of 3D models are presented for Zernike Moments 2D view descriptor showing the excellent performances of our Data mining Model.

II. METHODS OF 3D OBJECT INDEXATION

A. *Indexation 2D/3D*

The methods that use the 3D model views came from psychophysical results which imply that the human visual system represents 3D objects as a group of 2D views rather than a three-dimensional model. The objective of these methods is to search three-dimensional models that are the closest to the 2D image.

This process can be divided into two phases: an indexation phase and a search phase. In the indexation phase, for each and every three-dimentional model in the base, we determine the characteristic views and their associated vectors using a method of 2D form analysis : : CSS[3], ART[4], Zernike[5][6] etc. Then comes the search phase where the requested image goes through a process similar to views in the base, in which a descriptor is computed and compared to the descriptors of the base.

The two main issues are the choice of the finest indexation method of 2D views and the choice of the optimum number of views to be taken into account and which views are to be used.

The selection of the form descriptor of 2D views will not modify the main concept, but it will change the performances and efficiency according to the chosen method.

In our study, Zernike Moments were chosen to describe characteristic views.

B. *Search of association rules*

Introduced by Agrawal and al.[7], the searching method of association rules has been proposed to allow us to analyze the supermarket sells to extract rules such as : “when a customer buys bread and butter, he also buys milk 9 times over 10”. As Hébrail suggests [8], even though this association rule concept was developed for marketing purposes, it could also be implemented in another field of research like the research of frequent co-occurrences, pairs or values if the data structure is in accordance with this research.

III. PROPOSED METHOD

A. *Properties of the algorithm of selection of characteristic views*

Now that we have each initial view represented by a vector of Zernike Moments, we wish to reduce this set of vectors to the minimum by removing the redundancies that could have been created in a row of BDT (impossible to see 2 equal views in the same row of a BDT).

We use the 2 phases of the extraction algorithm of association rules (described above) between our BDT views in order to classify the set of views (the number of classes is unknown) depending on the choice of the threshold distance and to extract the characteristic view of each class.

According to CLOSE+[2], the outcome will be a certain number of groups. Each group will contain a set of minimal reduced views depending on the minimum support threshold. These views are the most frequent ones.

We use the confidence value between the views of the same class to extract the characteristic views of the 3D Object (one view per class only). The characteristic view of a class is the one that has the minimal confidence value in relation to the sets of views of its class.

B. *Generation of association rules*

The association rules that are used here are not limited to the rules which consequences are composed of one single item. To generate the rules, we take into account all non empty sub-sets of f for each frequent itemset of f. And, for each of these sub-sets h, we return a rule in the form of h => (f-h) if the ratio “support (f) /support (h)” equals at least a minimal confidence threshold minconf. We evaluate all the sub-sets of f to generate the rules which consequences are the itemsets of size greater than one.

IV. EXPERIMENTS AND RESULTS

In order to measure the performance of our algorithms, we used the Princeton benchmark base that is classified into 20 classes depending on their visual similarity.

The 3D Objects of figure 1 shows examples of 3D objects of different classes.

In our current implementation, Zernike moments are extracted from views from second order. To compare two Zernike moments, the distance Minkowski order 1 is used.

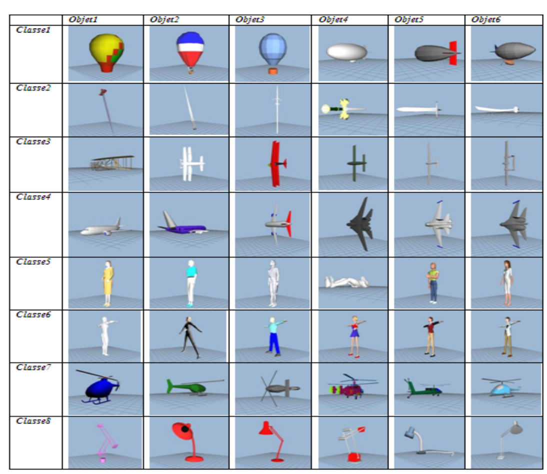

Figure 1. Examples of 3D Objects of different classes.

The choice of a minimum support and a minimum confidence threshold is justified by the number of experiences.

For a minsup <50%, we will have enough characteristic views, with unsignificant views among them.

For a minsup >60%, we will have a small number of characteristic views which represent a limited number of initial views of the 3D object.

In order to get characteristic views that represents best their class, it is better to choose a minsup that belongs to the interval [48% , 60%], as well as a minimum confidence threshold greater than 70% .

A. *Steps of the experiment*

Step1: extraction of characteristic views:

Based on a minimum support of 50% and a minimum confidence threshold of 80%, we tested our method on a 3D object of the class “arme”. The obtained outcomes are shown in figure 2:

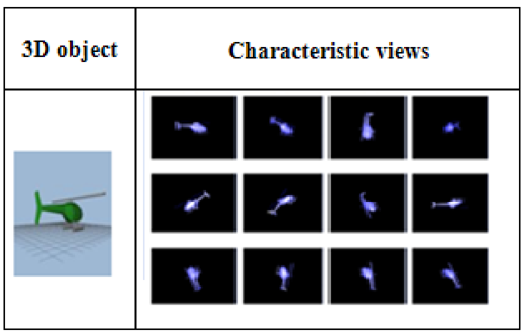

Figure 2. Characteristic views of the object « m640.off »

Step2: Research of objects that are similar to the 2D requested view

Using a minimum support of 50% and a minimum confidence threshold of 80%, the research of similarity according to our method and the use of probability on a 2D request example produces the outcomes displayed in figure 3.

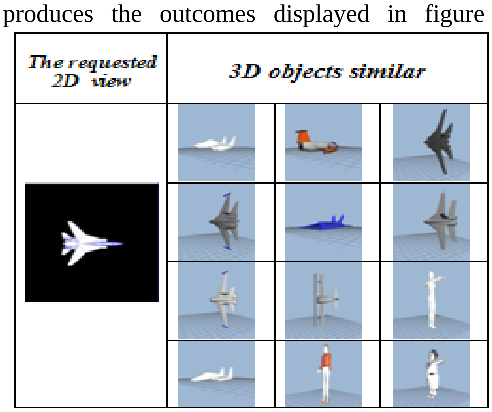

Figure 3. 3D objects that are similar to the requested 2D view of class “Fighter”

B. *Time comparison*

A general description of our 3D database is given below:

Class Number	Number of objects of the class	Object size	View size
1	9	342 views	644*645
2	6	342 views	644*645
3	14	342 views	644*645
4	8	342 views	644*645
5	37	342 views	644*645
6	14	342 views	644*645
7	23	342 views	644*645
8	11	342 views	644*645

TABLE1. A GENERAL REPRESENTATION OF THE 3D DATABASE

1. *Comparison of the extraction time for characteristic views of each object in our 3D database using both methods*

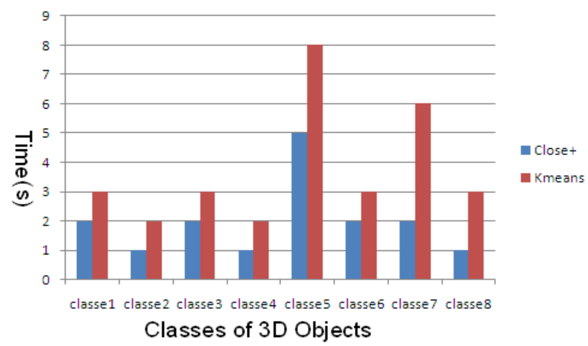

Figure 4. The extraction time for characteristic views of each 3D object class

	Average time saving (s)	Total time saving(s)	Charm average time(s)	KMEANS average time(s)
Results	1,625	13	2,125	3,75

TABLE2. The Total and Average time saving

2. *The ratio between characteristic views and the set of 3D database views of each class*

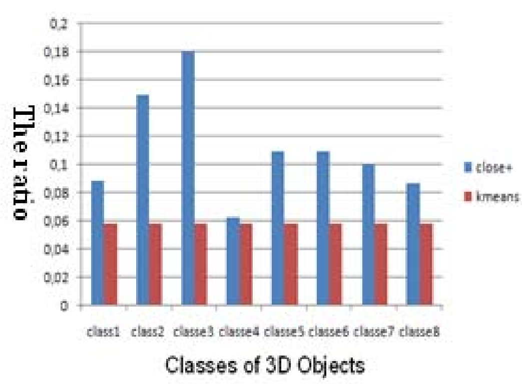

Figure 5. The ratio between characteristic views and the set of 3D database views of each class

3. Comparison of the average search time of a requested view using both methods

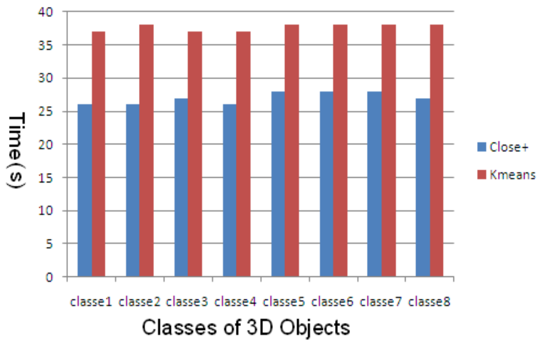

Figure 6. The average search time of similar 3D objects for requested views

	Average time saving (s)	Total time saving(s)	Charm average time(s)	KMEANS average time(s)
Results	10,625	85	27	37,625

TABLE3. The Total and Average time saving.

In figure 4, we notice that the Close+ method is faster than the Kmeans one. The average time saving of the first method compared with the Kmeans one is 1,625s. The total time saving is 13s.

This is due to the fact that Close+ starts first to clean the base to eliminate all the redundant views. So we are left, for each class, with a reduced number of views. Obviously, a research on characteristic views for a reduced number would be faster and more efficient.

Figure 5 presents the relationship between the characteristic views and the whole views of each class. We notice that as for the Kmeans method, which is a supervised method, this relationship is constant and low. This is due to the fact that we have fixed the number of characteristic views to 20 and the number of views for each object to 342. However, as far as the Close+ method is concerned, this relationship is relatively higher. This shows the importance of cleaning the base and of keeping the non-redundant views only. From now on, we can build a new 3D database that only contains real views of the 3D object and redundancy-free. This will facilitate the research phase.

Figure 6 shows the importance of using this new Base and the impact on the time of the research. In fact, the average time saving is of 10,625s. The total saving for the whole DB is of 85s.

So far, we have shown the advantages of using the proposed method in terms of time and time saving measurement. Next, we will show the advantages in terms of performance.

C. Measurement of performance indices

We apply our algorithm « CLOSE+ and Extraction of association rules between views» and the Kmeans algorithm over 10 requested views of the same object of the class avion of the Benchmark base. These views are randomly selected by a program. Figure 4 displays results comparing Kmeans (in red) and the suggested algorithm in blue.

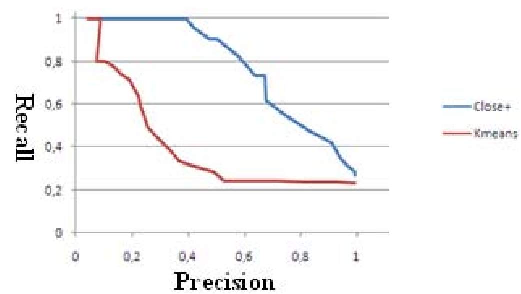

Figure 7. Curves of Recall/ Precision with values of 10 requested views

Figure 4 displays the Recall/Precision curve computed using our algorithm and the Kmeans curve. In this figure, we notice that the accuracy score of the first 35 K (first third) is close to 100% for our suggested method. This means that most of the objects that are similar to the requested object are placed at the top of the search list. And, as of the K greater than 35 (second thirds) in the search list, the Recall score goes near 100%, which signifies that the rest of the objets (belonging to the same class) that are similar to the requested object are ranked/placed* as two thirds in the search list. These indices are in accordance with the visual results displayed in Figure 3.

On the other hand, the Kmeans algorithm gives a result between 70 and 33% for the first 35K (first third). This implies that the first third of the search list contains a combination of objects: those belonging to the same class as the requested object and others belonging to other classes. Moreover, the K that is greater than 35 corresponds to a score less than 85%, which suggests that the rest of the objects belonging to the same class that are similar to the requested object are far from the top of list search.

V. CONCLUSION

In this article, we introduced the concept of 3D object indexation, in particular the indexation from the views of these 3D objects. First, we introduced an algorithm that is independent of the 2D used descriptor in order to extract the characteristic views of a 3D object. The outcomes of this method are very satisfying since this latter reduces the 3D object size (instead of using 342 initial views, the system automatically reduces this number depending on the threshold of the distance – to a smaller number-). The 3D object will then be characterized by a small number of views called “characteristic views”. Next, we used the probabilistic bayesian view (translated by T.F

Ansary, J.P. Vandeborred and M. Daoudi[5]) for the indexation of these 3D models. The displayed results highlight the good performances of this method compared to some classical methods of classification.

This method produces great results when the object size is big enough, with more than 340 views per object.

These tests are performed based on Princeton Shape benchmark.

REFERENCES

[1] Rakesh Agrawal , Ramakrishnan Srikant, Fast Algorithms for Mining Association Rules in Large Databases, Proceedings of the 20th International Conference on Very Large Data Bases, p.487-499, September 12-15, 1994

[2] Mohammed J. Zaki, Member, IEEE, and Ching-Jui Hsiao Efficient Algorithms for Mining Closed Itemsets and Their Lattice Structure IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 17, NO. 4, APRIL 2005

[3] A. Baskurt, F. Blum, M. Daoudi, J.L. Dugelay, F. Dupont, A. Dutartre, T. Filali Ansary, F. Fratani, E. Garcia, G. Lavoué, D. Lichau, F. Preteux, J. Ricard, B. Savage, J.P. Vandeborre, T. Zaharia. SEMANTIC-3D : COMPRESSION, INDEXATION ET TATOUAGE DE DONNÉES 3D Réseau National de Recherche en Télécommunications (RNRT) (2002)

[4] T.Zaharia F.Prêteux, Descripteurs de forme : Etude comparée des approches 3D et 2D/3D 3D versus 2D/3D Shape Descriptors: A Comparative study

[5] T.F.Ansary J.P.Vandeborre M.Daoudi, Recherche de modèles 3D de pièces mécaniques basée sur les moments de Zernike

[6] A. Khothanzad, Y. H. Hong, Invariant image recognition by Zernike moments, IEEE Trans. Pattern Anal. Match. Intell.,12 (5), 489-497, 1990.

[7] Agrawal R., Imielinski T., Swani A. (1993) Mining Association rules between sets of items in large databases. In : Proceedings of the ACM- SIGMOD Conference on Management of Data, Washington DC, USA.

[8] Hébrail G., Lechevallier Y. (2003) Data mining et analyse des données. In : Govaert G. Analyse des données. Ed. Lavoisier, Paris, pp 323-355

[9] T.F.Ansary J.P.Vandeborre M.Daoudi, une approche baysiénne pour l'indexation de modèles 3D basée sur les vues caractéristiques

[10] Ansary, T. F. Daoudi, M. Vandeborre, J.-P. A Bayesian 3-D Search Engine Using Adaptive Views Clustering, IEEE Transactions on Multimedia, 2007

[11] Ansary, T.F. Vandeborre, J.-P. Mahmoudi, S. Daoudi, M. A Bayesian framework for 3D models retrieval based on characteristic views, 3D Data Processing, Visualization and Transmission, 2004. 3DPVT 2004. Proceedings. 2nd International Symposium Publication Date: 6-9 Sept. 2004

[12] Agrawal R., Srikant R., Fast algorithms for mining association rules in larges databases. In Proceeding of the 20th international conference on Very Large Dada Bases (VLDB'94), pages 478-499. Morgan Kaufmann, September 1994.

[13] U. Fayyad, G.Piatetsky-Shapiro, and Padhraic Smyth, From Data Mining toKnowledge Discovery in Databases, American Association for Artificial Intelligence. All rights reserved. 0738-4602-1996

[14] S.Lallich, O.Teytaud, Évaluation et validation de l'intérêt des règles d'association

[15] W.Y. Kim et Y.S. Kim. A region-based shape descriptor using Zernike moments. Signal Processing : Image Communication, 16 :95–100, 2000.