Comparison of Multidimensional Data Access Methods for Feature-based Image Retrieval

Serdar Arslan
TUBITAK Space Technologies Research Institute
Ankara, Turkey
serdar.arslan@uzay.tubitak.gov.tr

Ahmet Saçan
School of Biomedical Engineering, Science and Health Systems, Drexel University
Philadelphia, PA, USA
ahmet.sacan@drexel.edu

Esra Açar, İ. Hakkı Toroslu, Adnan Yazıcı*
Dept. of Computer Engineering, Middle East Technical University
Ankara, Turkey
{e1455377, toroslu, yazici}@ceng.metu.edu.tr

Abstract*— Within the scope of information retrieval, efficient similarity search in large document or multimedia collections is a critical task. In this paper, we present a rigorous comparison of three different approaches to the image retrieval problem, including cluster-based indexing, distance-based indexing, and multidimensional scaling methods. The time and accuracy trade-offs for each of these methods are demonstrated on a large Corel image database. Similarity of images is obtained via a feature-based similarity measure using four MPEG-7 low-level descriptors. We show that an optimization of feature contributions to the distance measure can identify irrelevant features and is necessary to obtain the maximum accuracy. We further show that using multidimensional scaling can achieve comparable accuracy, while speeding-up the query times significantly by allowing the use of spatial access methods.

Keywords-component; Multidimensional Access Methods, CBIR, indexing, BitMatrix, SlimTree,LMDS,Fastmap, MPEG-7.

I. Introduction

As the number of applications in information technologies increases, efficient retrieval of multidimensional data for these applications becomes crucial in order to make the best use of the available data. Classical approaches for accessing multidimensional data are often inapplicable or insufficient, due to the complexity and high dimensionality of the data. Naive application of classical approaches usually incurs an unacceptable penalty in retrieval time or search accuracy. In order to meet the requirements of content based retrieval in complex data domains, several multidimensional access methods have been proposed [1].

In this study, we focus on content-based image similarity measurement and retrieval process and evaluate three different multidimensional access methods; BitMatrix cluster-based structure [2], Slim-Tree, which is an index structure on metric spaces [3], and a Landmark Multidimensional Scaling (LMDS) method which uses FastMap method for landmark selection [4]. While BitMatrix partitions and indexes in the feature space of the data objects; Slim-Tree and LMDS are based on the distance space, partitioning the data objects themselves based on the inter-image distance measurements. We evaluate the performance of these methods using query times and retrieval accuracy as the main criteria.

The similarity of images is computed by a feature-based similarity measurement technique, where four low-level MPEG-7 [5] feature descriptors are extracted and used. The Euclidean distance is used as the similarity function and the weights of the distance function are determined by either a constant weighting (CW) of each feature or by dynamically adapting the weights using the Ordered Weighting Averaging (OWA) method [6].

The performance of the multidimensional access structures are evaluated using the Corel database [7], and the retrieval accuracy is measured using the image classifications provided in the dataset. A Sequential Scan method that compares the query with every single database object is also included in the comparison to provide a base line for time and accuracy values. Section 2 below provides an overview of the multidimensional access methods proposed in this study and Section 3 presents the performance experiments.

II. Background

A. BitMatrix

BitMatrix [2] is an access structure proposed for multimedia data retrieval, and uses a data approximation approach in the spirit of VA-File [1]. BitMatrix partitions each feature dimension of the multimedia data into a set of

* This work is supported in part by a research grant from TUBITAK EEEAG with grant number 109E014.

1051-4651/10 $26.00 © 2010 IEEE
DOI 10.1109/ICPR.2010.797

ranges and assigns the multimedia object to the range that it belongs in the related feature dimension. As a result of this partitioning, bitmap signatures are generated for each multimedia object.

In order to build a BitMatrix, the images are first clustered into a predefined number of clusters along their low-level feature dimensions. During the clustering process, bitmap signatures are generated for images. The signatures contain a one (1) for the cluster to which the feature value of the image belongs and a zero (0) for the other clusters. The collection of these bitmap signatures for each image forms the BitMatrix access structure.

B. *SlimTree*

SlimTree is a balanced and dynamic paged structure that can be used to index multimedia data efficiently in metric space [3]. SlimTree has two types of nodes; leaf nodes to store the data objects, and internal nodes to store the partitioning information that guide the search and pruning during a query processing. In this structure, the multimedia data is presented by means of complex features and similarity between any two objects is defined by a possibly time consuming metric distance function. Slim-Tree indexes the distances between objects, rather than the features of the individual objects.

C. *Landmark Multidimensional Scaling (LMDS)*

The Landmark MDS (LMDS) algorithm was introduced in [8], and its aim is to decrease the cost of classical MDS algorithm by using a set of n chosen points referred as "landmarks". LMDS algorithm first applies the classical MDS to the nxn distance matrix of pair wise distances between the landmarks. A distance-based triangulation procedure is then applied over the nxN distance matrix to determine the coordinates for the remaining N points. There are two different methods for the selection of n landmarks [9];

- Random choice.
- MaxMin (greedy optimization): landmark points are chosen one at a time, and each new landmark maximizes, over all unused data points, the minimum distance to any of the existing landmarks. The first point is chosen arbitrarily.

Random choice (LMDS_{Random}) works quite well in practice and LMDS requires O (nN) space and O($CnN + knN + n^3$) time where k is the number of dimensions and C is the cost for accessing or computing each distance.

We have previously proposed a metric-preserving, landmark-guided embedding approach to represent bimolecular sequences in the vector domain, in order to allow efficient indexing and similarity [4]. Specifically, FastMap is used as a preprocessing step to identify the most effective landmark points for the LMDS algorithm ($\mathrm{LMDS}_{FastMap}$). The algorithm for choosing landmark points affects the quality of the embedding; the results in [4] show that combining the landmark selection strategy of FastMap and the stability of LMDS achieves better performance than LMDS_{MaxMin} and LMDS_{Random}.

III. EXPERIMENTS

Images from the Corel Database [7], which consists of ten groups (elephant, mountain, beach, horse, African people, flower, dinosaur, bus, food and architecture) each having 100 images, were used to evaluate the performance of the multidimensional access methods. SlimTree was implemented using the XXL API [10], whereas BitMatrix was implemented using Weka API [11] and Colt API [12]. Weka API has been used for clustering low-level features while Colt API has been used to store the constructed BitMatrix access structure and to perform queries on the BitMatrix. The $\mathrm{LMDS}_{FastMap}$ implementation in [4] has been modified to use feature-based image similarity measure for the embedding. The embedded vectors are indexed using the X-tree spatial access method [13].

In order to compare images in the feature-based similarity measurement method, the low-level features are first extracted from the images. These features are described using MPEG-7 descriptors [5]. In [14], MPEG-7 visual descriptors are analyzed from the statistical point of view. The main results show that the best descriptors for combination are Color Layout, Dominant Color, Edge Histogram and Texture Browsing. The others are highly dependent on these descriptors. Thus, the descriptors used in this study are: Color Layout (CL), Dominant Color (DC), Edge Histogram (EH) and Region Shape (RS). These descriptors are generated in the XML format using the MPEG-7 eXperimentation Model (XM) [15].

After the feature extraction process, BitMatrix and Slim-Tree index structure implementations in [16] and $\mathrm{LMDS}_{FastMap}$ implementation in [4] have been modified in order to compare the images via a distance function. BitMatrix was built by clustering CL, DC, EH and, RS feature dimensions into a pre-defined number of ranges. Similarity measurement for these structures was carried out using the Euclidean Distance measure, which is a metric distance function.

In addition to the indexing mechanism, another critical issue is the appropriate weighting of the low-level features within the distance function. In general, when more than one feature is being processed in similarity searching with weighted distance functions, the same weights are used for the features of all images in the database. However, when comparing two images, any one of the low-level features may be more distinctive than the others, requiring that feature to be associated with a higher weight. Thus, the OWA method [6] is adapted to handle this situation and is compared to the constant weighting schema applied to the Euclidean Distance function. For both Constant Weight and

OWA measures, we have optimized the weights using the Nelder-Mead simplex method [17], with the objective of increasing the area under the precision-recall (P-R) curve. For Constant Euclidean (CW), the optimized weights were:

$w_{CL} = 0.45531, w_{EH} = 0.51084, w_{RS} = 0.021509, w_{DC} = 0.012348$

Highlighting the contributions of the Color Layout and Edge Histogram features and giving only marginal contributions to the Region Shape and Dominant Color features. For OWA, the following optimized weights were obtained:

$w_1 = 0.52754, w_2 = 0.36791, w_3 = 0.053913, w_4 = 0.050638$

P-R results for all multidimensional access methods are shown in Table I and a selection of these results are depicted in Figure 1. Optimized constant weights in feature-based similarity measurement method are found to outperform the OWA approach and among the multidimensional access methods, SlimTree with constant weights performs the best in terms of retrieval accuracy. As embedded space dimension is increased, $\text{LMDS}_{\text{FastMap}}$ accuracy approaches to that of the SlimTree.

Another performance metric evaluated in this study is the CPU time requirement for k nearest neighbor queries. Figure 2 shows the query times for varying values of k in nearest neighbor queries for each of the methods. An embedding dimensionality of d=5 and d=10 are used for $\text{LMDS}_{\text{FastMap}}$.

TABLE I. AREA VALUES UNDER PR CURVES

Access Method	Weighting Approach	
	CW	*OWA*
Sequential Scan	0.4273	0.3906
SlimTree	0.4273	0.3906
BitMatrix	0.3915	0.3609
LMDS (d=5)	0.3607	0.3228
LMDS (d=10)	0.3976	0.3403
LMDS (d=20)	0.4129	0.3751
LMDS (d=30)	0.4191	0.3815
LMDS (d=40)	0.4208	0.3739
LMDS (d=50)	0.4234	0.3812
LMDS (d=100)	0.4196	0.3771

While BitMatrix has similar P-R results to $\text{LMDS}_{\text{FastMap}}$, it requires significantly more time to find the result objects, because of time spent during the clustering process in each query evaluation phase. Moreover, search time values for SlimTree are slightly better than BitMatrix results, but there is still a significant difference between SlimTree and $\text{LMDS}_{\text{FastMap}}$, since SlimTree requires disk page accesses rather frequently.

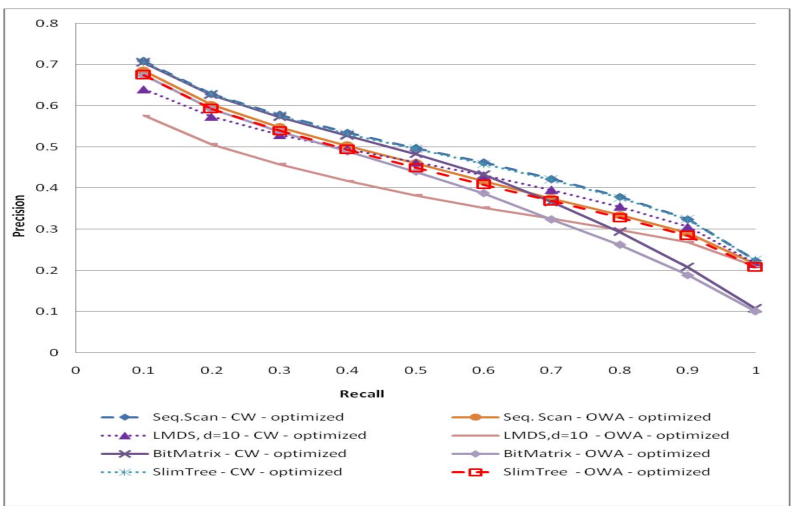

Figure 1. P-R results for both approaches

TABLE II. SEARCH TIME VALUES FOR 10-NN QUERIES USING LMDS

	Weighting approach	
Space dimension	*CW (sec)*	*OWA (sec)*
d=5	0.594	0.547
d=6	0.657	0.703
d=7	0.813	0.828
d=8	0.938	0.953
d=9	1.125	1.172
d=10	1.265	1.248
d=20	2.641	2.532
d=30	3.797	3.812
d=40	5.141	4.969
d=50	6.250	6.063
d=100	11.359	11.500

Table II gives a closer look into the $\text{LMDS}_{\text{FastMap}}$ search time test results. Within the range of dimensions sampled, the search time for k-NN queries increases linearly as the dimensionality of the embedded space increases. Even when the embedded space dimensionality is relatively high (d=100), XP-tree indexing over $\text{LMDS}_{\text{FastMap}}$ embedding is still faster than any other access structure for our dataset. Furthermore, using constant weights or adapting OWA makes no significant difference in retrieval time for $\text{LMDS}_{\text{FastMap}}$.

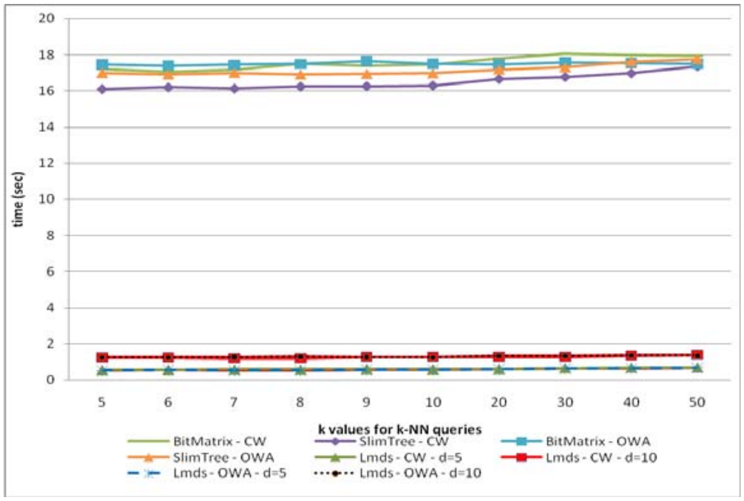

Figure 2. Search time for k-NN queries

IV. CONCLUSION

In this paper, we have focused on distance based indexing methods for image data. We performed search accuracy and CPU time experiments on the Corel image database. This database provided a curated and convenient ground truth against which the accuracy of the retrieval methods can be measured. However, we acknowledge that image similarity is inherently a subjective measure and the classification of images into restrictive groups may not be appropriate for certain applications. Based on our analysis, we recommend that a constant weighted Euclidean distance measure would be sufficient for most applications, provided that the weights are optimized to highlight the most relevant features.

Among the multidimensional scaling methods, the metric indexing structure SlimTree provides the best precision and recall values. Since the weighted Euclidean measure is a metric, SlimTree is guaranteed to give the same results as Sequential Scan. On the other hand, SlimTree requires significantly more time to find the relevant objects to the query object, due to lack of sufficient pruning and thus large number of distance calculations incurred. The clustering based structure BitMatrix has similar accuracy to $LMDS_{FastMap}$ (when embedding dimension is less than ten 10) and similar query times with SlimTree.

To conclude, we have showed that multi-dimensional scaling can reduce the retrieval problem to a spatial-indexing task, where queries can be performed orders of magnitude faster than distance or cluster based indexing methods. The accuracy of the embedded space is shown to be comparable to that of the retrieval performed in the original space. Moreover, we show that constant weighting scheme can perform better than OWA for all access methods proposed in this paper. In future studies, image-registration based similarity measurement methods will be investigated in contrast to the feature-based similarity measure presented in this study.

REFERENCES

[1] V. Gaede, O. Gunther, “Multidimensional Access Methods”, ACM Comp. Surv. V.30, No.2, 1998.

[2] C. Calistru,C. Riberio , G. David , “Multidimensional Descriptor Indexing: Exploring the BitMatrix”, pp. 401-410, CIVR’06, 2006.

[3] C.Traina Jr., A. Traina, B.Seeger, and C. Faloutsos, “Slim-trees: High Performance Metric Trees Minimizing Overlap Between Nodes”, In EDBT 2000, pp. 51--65, Konstanz, Germany, Mar. 2000.

[4] A. Sacan and I. H. Toroslu. Approximate Similarity Search in Genomic Sequence Databases using Landmark-Guided Embedding. SISAP'08, 2008.

[5] “MPEG-7 Overview (ver. 9)”, Int. Org. Stanart., ISO/IEC JTC1/SC29/WG11, May 2003.

[6] R.R. Yager, “On ordered weighted averaging aggregation operators in multi-criteria decision making”, IEEE Trans. SMC. 18, pp. 183-190, 1988.

[7] Corel database, http://www.corel.com.

[8] V. de Silva, J. B. Tenenbaum, “Global versus local methods in nonlinear dimensionality reduction,” in Advances in Neural Information Processing Systems 15, S. T. S. Becker and K. Obermayer, Eds. Cambridge, MA: MIT Press, pp. 705–712, 2003.

[9] V. de Silva, J.B. Tenenbaum “Sparse multidimensional scaling using landmark points”, June 2004.

[10] http://dbs.mathematik.uni-marburg.de

[11] Weka Api, http://www.cs.waikato.ac.nz/ml/weka/

[12] Colt Project, http://dsd.lbl.gov/~hoschek/colt/.

[13] S. Berchtold, D. Keim,H.-P. Kriegel, “ The X-tree: An index structure for high-dimensional data.”, In Proc. of the 22nd Int. Conf. on VLDB, (Bombay) 28–39, 1996.

[14] H. Eidenberger, “How good are the visual MPEG-7 features”, Proc. of the 5th ACM SIGMM Int. WS on Mm. info. retrieval, pp.130-137, Berkeley, 2003.

[15] Mpeg7XM,http://www.lis.ei.tum.de/

[16] E. Açar, S. Arslan, A. Yazıcı and M. Koyuncu, “Slim-Tree and Bitmatrix Index Structures in Image Retrieval System Using MPEG-7 Descriptors,” CBMI'08, 2008.

[17] J.C. Lagarias, J. Reeds, M. H. Wright, and P. E. Wright. “Convergence properties of the Nelder-Mead simplex method in low dimensions.” SIAM Journal of Optimization, 9(1):112–147, 1998.