International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:1, No:4, 2007

Indexing & Searching of Image Data in Multimedia Databases Using Axial Projection

Khalid A. Kaabneh

Abstract—This paper introduces and studies new indexing techniques for content-based queries in images databases. Indexing is the key to providing sophisticated, accurate and fast searches for queries in image data. This research describes a new indexing approach, which depends on linear modeling of signals, using bases for modeling. A basis is a set of chosen images, and modeling an image is a least-squares approximation of the image as a linear combination of the basis images. The coefficients of the basis images are taken together to serve as index for that image.

The paper describes the implementation of the indexing scheme, and presents the findings of our extensive evaluation that was conducted to optimize (1) the choice of the basis matrix (*B*), and (2) the size of the index A (*N*). Furthermore, we compare the performance of our indexing scheme with other schemes. Our results show that our scheme has significantly higher performance.

***Keywords*—**Axial Projection, images, indexing, multimedia database, searching.

I. Introduction

As modern computer technology advances at a rapid pace, multimedia systems have become an important part of our lives. Multimedia systems have been used in areas such as education, training, business, and entertainment. As a result of the extensive need for multimedia data search and the rapid growth in the computational power in computers, the processing of non-textual data such as images and videos have grown tremendously and are expected to have explosive growth in the coming years.

Searching for data, generally, is one of the most fundamental and widely performed operations, and has thus received considerable attention. The complicated nature and immense size of multimedia data make the search problem more difficult, and renders the traditional keyword-based indexing and search techniques inadequate.

An alternative approach for multimedia search is content-based indexing [1]. Given a large database of images, our objective is to derive a scheme for the extraction of suitable features that index the database.

Manuscript received December 5, 2004. This work was supported by the Multimedia Research Group (MMRG) at Amman Arab University for Graduate Studies, Jordan.

K. A. Kaabneh is an assistant professor at the Department of Computer Science, Amman Arab University for Graduate Studies, Amman 11953, Jordan (e-mail: kaabneh@ aau.edu.jo).

In this paper, we develop a new approach to indexing for similarity search, which depends on linear modeling of signals, using a selected basis of some ***M*** images for modeling [2, 3]. Modeling an image is a least-squares approximation of the image as a linear combination of the basis images. This is equivalent to viewing an ***NxM*** image as a vector in an ***NM***-dimensional space and projecting that vector into each basis image-vector of the ***N*** basis images separately [4, 5]. The coefficients of the projected image-vector along the basis images are taken together to serve as an index for that image. This approach is called axial projection [6].

The paper is organized as follows. The next section is overviews of the linear modeling indexing. Section 3 presents our indexing scheme. Section 4 presents the choice of Basis. Our performance evaluation and optimization results are presented in section 5. The last section presents a summary and our conclusions.

II. Linear Modeling Indexing

Modeling is very useful in studying and processing complex systems, and leads to simplified representation of a small number of relevant aspects. Among the most common modeling approaches is data surface fitting [7].

This class of techniques is based on linear modeling of signals, using application dependent bases for modeling. A basis is a set of chosen signals, and modeling a signal is a least-squares approximation of the signal as a linear combination of the basis signals. The coefficients in the linear expansion are taken together as a single index of the signal. The choice of the basis impacts the performance of the corresponding indexing scheme [8].

III. Proposed Indexing Scheme

We have developed a new indexing scheme using surface fitting. The scheme uses a basis (set) of chosen images, and models an image as a least-squares linear combination of the basis images. The coefficients of the basis images are taken together to serve as index for that image.

Although we have a very large database of images, the idea is to select a small number of representative images, which we call that a basis. Let $B_1, B_2, ..., B_N$ be ***N*** basis images, chosen apriori once and for all. The choice of basis will be discussed later.

Geometrically, when we model an image I (as a vector), we project that image vertically or orthogonal on each basis vector B_j separately. Here, the basis vectors are treated as axes. The coefficients of the index $(a_1, a_2, ...a_N)$ are taken to be the cosines of the angles α_j between vector I and axis vector B_j, for j=1,…, **N**.

Specifically,

$$a_j = \cos(\alpha_j) = \frac{I \circ B_j}{\|I\| * \|B_j\|} \quad (1)$$

where (o) stands for the inner product of two vectors I and B_j, and $\|.\|$ stands for the length (=Euclidean norm) of the vector. Note that if I and each B_j are mean-normalized (i.e., made of mean zero by subtracting the mean), then $\cos(\alpha_j)$ = correlation coefficients between I and B_j. That is, the index ***A*** is the vector of correlation coefficients between I and the basis images $B_1, B_2, ..., B_N$. This gives the axial projection scheme a statistical explanation.

In three dimensions, when the basis contains only two image representatives of the database, the situation may be represented by the configuration of Fig. 1, each axis corresponds to a different base B_j.

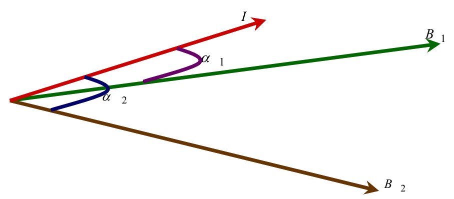

Figure 1: Vector Representation of the Images on Subspace and their angles a_j.

The algorithm to compute the index A is as follows:

1) Convert every image B_j into a one–column vector B_j^c.

2) Mean-normalize vector B_j^c:

$$\hat{B}_j = B_j^c - mean(B_j^c).$$

3) Compute the length of the vector $\hat{B}_j$.

4) Convert image I into a one–column vector I^c.

5) Mean-normalize vector I^c:

$$\hat{I} = I^c - mean(I^c).$$

6) Compute the length of the vector $\hat{I}$.

7) Compute $a_j = corr(I, B_j) = \cos\left(\hat{I}, \hat{B}_j\right) =$

$$a_j = \frac{\hat{I} \circ \hat{B}_j}{\|I\| * \|B_j\|}, \text{ for j=1,..,N.}$$

8) Take $A = (a_1, a_2, ...a_N)$ to be the index of image I.

Note that steps 1 to 3 are computed off line, and not part of the time complexity of computing the index A. The time complexity of computing A is $O(NXY)$ where the images are of size (XxY).

IV. CHOICE OF BASIS.

The main idea is to start selecting an initial basis of M vectors serving as representatives of the image database with a set of training vectors (image database). Then, form M clusters from the set of training vectors: put each training vector I in Cluster j if the j-th initial representative is the closest match to I :

$$d(I, B_j) = \min_{1 \le k \le M} d(I, B_k) \quad (2)$$

where *d(v, w) is* the (Euclidean or MSE) distance between the vectors v and w.

The next step is to restructure the clusters by computing the new basis such that the new representative B_j is the centroid of the recent cluster:

$$B_j = \frac{\sum_{I \in cluster_j} I}{size(cluster_j)} \quad (3)$$

Then compute the new total distortion of the clusters (differences between the training vectors and their centroids):

$$Dist = \sum_{j=1}^{M} \sum_{I \in Cluster_j} d(I, B_j) \quad (4)$$

This process of clustering and determining the new representative is continued until the clusters become stable. The clusters are assumed to have stabilized if the total distortions cease to change much:

$$\left| \frac{Old\_Dist - New\_Dist}{Old\_Dist} \right| \le Tolerance \quad (5)$$

Take the most recent basis to be the final basis.

Choice of Initial Basis:

The following schemes (or any variations of them) could be used to generate the initial clusters (training set):

a. **Random scheme**. The ***N*** initial representatives are randomly selected. Thereafter the clustering algorithm described above is used to compute the final basis.

b. **Splitting scheme**. In this scheme, all the ***N*** data items are initially regarded as a single big cluster. The representative is determined. Then a perturbed value of the representative is computed and the data items are divided into two clusters based on the two representatives. Thereafter the clustering algorithm described above is used to refine the representatives. This process is repeated until ***N*** clusters are reached. (Although this results in ***N*** being a power of 2, by creating only one new representative at every step, or after a suitable number of steps, any value of ***N*** can be handled).

c. **Pairwise scheme.** This is called pair-wise nearest neighbor (PNN) clustering. In this approach, each training vector is formed into a single cluster. Then the pair-wise distances of all data are determined and the closest pair is merged into a single cluster. After merging, a new representative of the cluster is determined. Thereafter the clustering algorithm described above is used to refine the representatives. This process is continued until the number of clusters reaches ***N***.

IV. Experimental Results

We have implemented our proposed indexing scheme, and optimized the following parameters: (1) the choice of basis (***B***), and (2) the appropriate number of coefficients, which is the size of the Basis (***N***). The four different choice of ***B*** are random, splitting, pairwise, and blocking scheme. The five different values of ***N***, which have been considered, are 256, 128, 64, 32, and 16. The optimal value is selected from those.

The performance metrics are the Recall and the search time. The Recall (**R**) is the number of reported hits divided by the total number of actual matches. A related metric is the False Misses Ratio (**FMR**), which is the ratio of unreported matches to the total number of actual matches. Clearly, **FMR = 1 - R**, and therefore, one needs to measure and maximize just the Recall. The False Recall Ratio **(FRR) = F reported / H reported**, where F reported is the number of reported but incorrect matches, and **H** reported is the total number of reported matches. The Effort to completion is defined as the minimum total number of retrieved answers for 100% recall.

Our experimental results are shown in Fig.2, where the Recall is shown as function of the number of selected coefficients (Basis) size **N**, for different choice of **B**. The plots clearly show that Axial Projection is optimal for large values of ***N***. Moreover, the pairwise scheme of the choice of basis performs the best followed by the splitting scheme for the same number of selected coefficients.

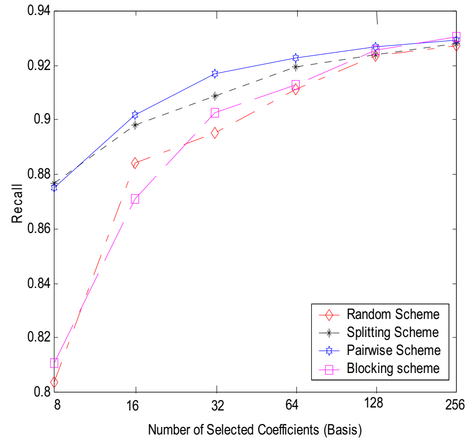

Figure 2: Recall vs. Various Axial Projection Options.

However, there were no false misses in any of the search schemes. We must note, though, that the index search algorithms are expectedly more robust against noise. The larger query coefficient sizes result in better Recall but it increase the search times. It must be noted also that beyond a certain size, there is no remarkable improvement in the Recall. Fig.3 plots the results of the effort to completion, where the effort to completion is shown as a function of the index size ***N*** for different choices of basis ***B***. The plot clearly shows that the bigger the index size, the smaller the number of retrieved images needed to reach the 100% recall for the same choice of basis.

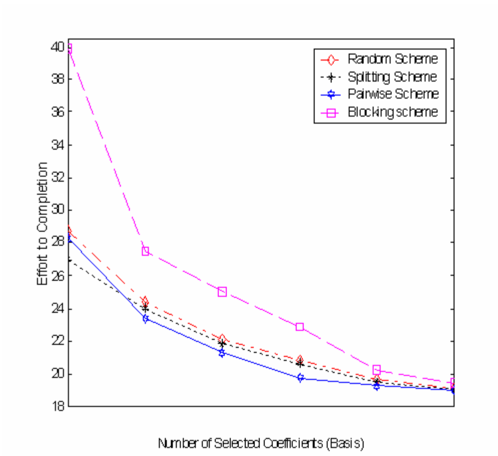

Figure 3: The Effort to Completion vs. Various Axial Projection Options.

To illustrate the use of linear modeling for images indexing and searching, we conducted a number of experiments on a large database of images. The result query is shown in Fig.4.

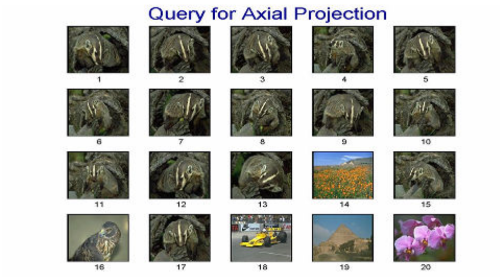

Figure 4: Query Result for Axial Projection Technique.

V. Conclusion

In this paper, we proposed an indexing scheme for image data in multimedia databases based on linear modeling techniques. The results have shown the effectiveness of Axial projection as indices to image data, in terms of search hits and efficiency of searches.

We have implemented the indexing scheme, and conducted extensive evaluation to optimize 1) the choice of the basis matrix (***B***), and (2) the size of the index A (***N***). Furthermore, we compared the performance of our indexing scheme with other schemes. Our results show that our scheme has significantly higher performance in terms of Recall, assuming the same index size. Our results also show that to derive the best performance out the Axial Projection Indexing, the pairwise scheme of the choice of basis is applied. As for the size of the index A (***N***), the optimal size of ***N*** is 64, even though using bigger sizes will increase the performance but will not be of a great impact.

We are considering variations and elaboration of our basic indexing scheme to further improve its performance. This and other indexing approaches will appear in our future work.

References

[1] Candocia, F. M., *"On the Featureless Registration of Differently Exposed Images,"* Proceedings of the 2003 International Conference on Imaging Science, Systems and Technology (CISST '03), Vol. I, pp. 163-169, June 23-26, Las Vegas, Nevada, 2003.

[2] L.G. Brown, “A Survey of Image Registration Techniques,” *ACM Computing Surveys*, Vol. 24, No. 4, pp. 325-376, Dec. 1992.

[3] J.B.A. Maintz and M.A. Viergever, “A Survey of Medical Image Registration,” *Medical Image Analysis*, Vol. 2, No.1, pp. 1-37, 1998.

[4] P.E. Debevec and J. Malik, “Recovering High Dynamic Range Radiance Maps from Photographs,” *Proc. Of SIGGRAPH*, pp. 369-378, 1997.

[5] T. Mitsunaga and S. Nayar, “Radiometric Self Calibration,” *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 374-380, 1999.

[6] J. A. Fessler. Improved PET Quantification Using Penalized Weighted Least-Squares Image Reconstruction. *IEEE Transaction on Medical Imaging*, 1992.

[7] Diaz, A. M., Barros, A. F. and Candocia, F. M., *"Image Registration in Range Using a Constrained Piecewise Linear Model: Analysis and New Results,"* Proceedings of the 2003 International Conference on Imaging Science, Systems and Technology (CISST '03), Vol. I, pp. 152-158, June 23-26, Las Vegas, Nevada, 2003.

[8] A. K. Jain, Fundamentals of Digital Image Processing, Prentice Hall, Englewood Cliffs, NJ, 1989.

Khalid A. Kaabneh received his Bachelor’s degree in Electrical Engineering and Computer Engineering in 1989. In 1991, he earned an MS Degree in Management Information Systems. A Ph.D. was bestowed upon him in 2001 for his research in the area of Multimedia audio watermarking. All degrees were awarded from The George Washington University in Washington DC, USA.
In 2001, he joined the faculty of the School of Information Technology at Mu’tah University, as an Assistant Professor. Currently, he is an Assistant Professor at the Department of Computer Science at Amman Arab University for Graduate Studies.
His research interests include areas like Digital Communications, Wireless Networks, Digital Multimedia, and Copyright Protection. He has supervised numerous Master’s degrees research projects of diverse areas. He co-edited the book *Discrete Structures* (Amman, Jordan: 2004). Dr. Kaabneh has published various research papers in a multitude of local and international journals.