MAPPING BY ADAPTIVE THRESHOLD METHOD FOR DIMENSION REDUCTION OF CONTENT-BASED INDEXING AND RETRIEVAL FEATURES

Esin Guldogan and Moncef Gabbouj

Institute of Signal Processing, Tampere University of Technology,
P.O. Box 553, FIN-33101 Tampere, Finland
E-mail: esin.guldogan@tut.fi, moncef.gabbouj@tut.fi

ABSTRACT

Dimension reduction methods have been commonly used for content-based multimedia indexing and retrieval. In this paper, we investigate the use of a mapping by adaptive threshold (MAT) method for dimension reduction of feature data. The proposed MAT method is implemented and compared to two other well-known dimension reduction methods, namely Principal Component Analysis and Multidimensional Scaling. Experimental studies on image retrieval reveal that the proposed method successfully reduces the dimension of feature vectors without degrading semantic image retrieval performance significantly. Furthermore, its computational complexity is significantly less than the other methods.

1. INTRODUCTION

The amount of digital audio/visual data has been increasing intensively with the improving multimedia technologies and Internet for decades. The generated huge amount has led to multimedia storage, management, and accessing problems. Various methods, algorithms, and systems have been proposed addressing these problems. Content-Based Multimedia Indexing and Retrieval (CBMIR) concept is one of the outcomes of the studies for solving multimedia handling and accessing problems. CBMIR studies have been driven both in commercial and academic institutes, and successful CBMIR systems have been proposed [1], [2], [3]. Despite these successful systems, there is still no perfect unique solution for CBMIR in general.

CBMIR systems utilize so-called low-level features, which represent the content of the corresponding multimedia item. Colour, texture, and shape are the most commonly used data for such features. Feature extraction is often a computationally complex process. Furthermore, combination of low- and high-level features for achieving significantly higher semantic performance increases the complexity of indexing and particularly retrieval. Run-time memory requirement for indexing and retrieval is also significant, and closely related to feature extraction process. Due to high memory and processing power requirements, CBMIR has not been widely used on limited platforms, such as mobile devices or distributed systems. However, the usage and generation of digital multimedia on all kinds of platforms are becoming widespread. Hence, performance optimization of indexing and retrieval plays an important role in practical CBMIR studies. Performance optimization during indexing and retrieval addresses three main groups of problems:

- Processing time and computational complexity,
- Disk and run-time memory space requirements, and
- Semantic retrieval performance.

Memory and processing time problems are becoming more important due to high dimensionality of the feature data and extensive usage of multiple feature data sets. A common approach addressing high dimensionality problem is selecting only the most principal features associated with a multimedia item while preserving the final semantic retrieval performance. There are several methods used in CBMIR for reducing dimensionality [4], [5], [6]. Principal Component Analysis (PCA) is one of the most commonly used methods transforming a number of possibly correlated variables into a smaller number of uncorrelated variables referred to as principal components [7]. It reduces the dimension of the feature space and reveals relationships between objects that facilitate searches by similarity. Multidimensional Scaling (MDS) is another dimension reduction technique projecting high-dimensional representation of feature data into low-dimensional representation [8].

In this paper, we use a mapping by adaptive threshold (MAT) method for dimension reduction of CBMIR features. The proposed MAT and other methods are further described in Section 2. The related experimental results and interpretations are given in Section 3. Finally Section 4 presents the concluding remarks and discussions.

2. DIMENSION REDUCTION METHODS

In CBMIR, dimension reduction methods are mostly utilized for reducing the dimension or size of the feature data. During reduction process, distances between multimedia feature vectors should not be affected significantly for successful retrieval results. It may be even better to increase the distances between the least relevant items while decreasing between the most relevant ones to improve semantic results. In general, the main objective for utilizing a dimension reduction method in CBMIR is decreasing the run-time complexity and memory requirement. Thus, the dimension reduction method itself is expected to have reasonable complexity.

Principal Component Analysis (PCA) and Multidimensional Scaling (MDS) methods have been widely used as dimension reduction methods for various purposes, as well as for CBMIR features. However, these methods modify the original values in feature vectors significantly. The PCA and MDS methods are further described in Section 2.1 and 2.2 respectively. Section 2.3 presents a simple mapping by adaptive threshold based dimension reduction method that does not modify the original values of feature vectors significantly.

2.1 Principal Component Analysis (PCA)

Each multimedia item has a feature data that can be represented using an n-dimensional feature vector X in CBMIR systems. PCA is a well-known technique to map n-dimensional vectors into k-dimensional vectors where $k<<n$. Principal components of original feature vector X can be denoted as:

$$Y_{pca} = E^T X \tag{1}$$

where $n \times k$ matrix E contains k eigenvectors whose i-th column is the i-th eigenvector corresponding to the i-th largest eigenvalue of covariance matrix of the data.

Choosing k largest eigenvalues and their associated eigenvectors yields to minimum least mean square error [9]. However, computing the covariance matrix and its eigenvalues are computationally expensive processes.

2.2 Multidimensional Scaling (MDS)

MDS provides a low-dimensional representation of objects given their high-dimensional representation. Basically, given a matrix of similarities or distances between items, MDS plots the items onto a map such that those items that are perceived to be very close to each other are placed near each other on the map.

MDS refers to a general method that falls into two groups: metric and nonmetric MDS [10], [11]. In metric MDS, the dissimilarities are assumed to be Euclidean distances and similarity is qualitative. Euclidean distance may not be suitable for every data, especially if the components of the data vectors are expressed on an ordinal scale. The method is called nonmetric if similarities are quantitative. Nonmetric MDS uses only the ordinal properties of the dissimilarities. Achieving significant mapping or representation of high-dimensional data with lower dimensionality can be possible by minimizing a certain loss function. In metric MDS, least-square error can be used as a loss function that is defined as:

$$L(x) = \sum_{k \neq l} \left[d(k,l) - d'(k,l) \right]^2 \tag{2}$$

where $d(k,l)$ is the distances between point k and l, and $d'(k,l)$ represents the possible distances between points k and l in low-dimensional space.

Nonmetric MDS uses the following normalized loss function [11]:

$$L(x) = \frac{1}{\sum_{k \neq l} \left[d'(k,l) \right]^2} \sum_{k \neq l} \left[f(d(k,l)) - d'(k,l) \right]^2 \tag{3}$$

where function f minimizing the Eq.3.

MDS is used in CBMIR systems for dimensionality reduction and yields successful results. However, as stated in [12], it is a computationally complex method due to pair wise distance calculation.

2.3 Proposed Mapping By Adaptive Threshold (MAT) Method

The proposed MAT method essentially consists of simple adaptive threshold and non-overlapping window-based mapping functions. In the adaptive thresholding step, one of the feature vector values is selected as the threshold value according to a user-defined scaling factor, and the thresholded feature vector is windowed. User should explicitly specify the scaling factor, which will be used for calculating the new feature vector size. The mapping function assigns one representative value for each window. This method reduces the dimensionality irrespective of correlation among elements of the vector. Thus, it is not recommended for data clustering. Feature vectors are assumed to contain only positive values in this method. For the non-positive feature vectors, a proper function should be used before the thresholding step for transforming negative values, e.g. absolute function. In the mapping step, the original values should be used instead of the transformed values. The following steps represent the proposed MAT method in details:

Step 1: Adaptive Threshold

X represents the original N-dimensional feature vector, and M represents the target dimension, which equals to N/S where S is the user-defined scaling factor.

$$X = \{x_1, x_2, ..., x_N\} \tag{4}$$

The values of vector X are sorted into X' in descending order.

$$X' = \{x_{(1)}, x_{(2)}, ..., x_{(N)}\} \tag{5}$$

where $x_{(i)}$ is the sample of X at i-th rank, and $x_{(i)} > x_{(i+1)}$.

The M-th value in X', $x_{(M)}$ is selected as the threshold, and X'' is constructed by thresholding the original feature vector X.

$$X'' = \left\{ x''_i \mid x''_i = \begin{cases} x_i, if\ x_i \geq x_{(M)} \\ 0\ otherwise \end{cases}, i = 1..N \right\} \tag{6}$$

Step 2: Mapping

The vector X'' is divided into M non-overlapping windows. A simple mapping is performed on these windows to construct the final decimated vector Y. A representative value is assigned for each window, where the representative value of i-th window will be the i-th element of vector Y. Function G finds the representative value given a window of X''.

$$Y = \{y_i \mid y_i = G(\{x''_{(i-1)S+1} ... x''_{iS}\}), i = 1..M\} \tag{7}$$

Some windows may contain only 0 values, in this case the representative value of that window is also 0. G function may be the mean, maximum, or median function. The following pseudo code expresses Step 2 more clearly, when G is the maximum function.

```
Y = Zeros(M) // Final feature vector
for i = 1:N
        // Find window index
        w = round( i / S )
        // Maximum selection
        if ( X''[i] > Y[w] )
                Y[w] = X''[i]
        End if
End for
```

Figure 1 illustrates the original values of a HSV histogram feature vector, and Figure 2 illustrates the result of the proposed MAT method on the same feature vector when G is the maximum function.

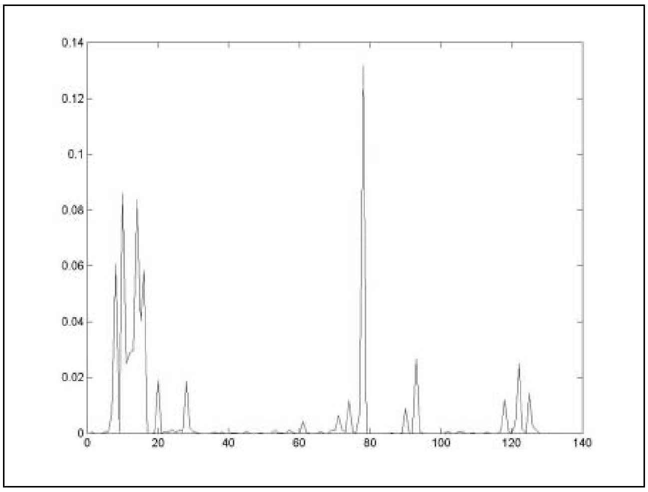

Figure 1: Original HSV feature vector

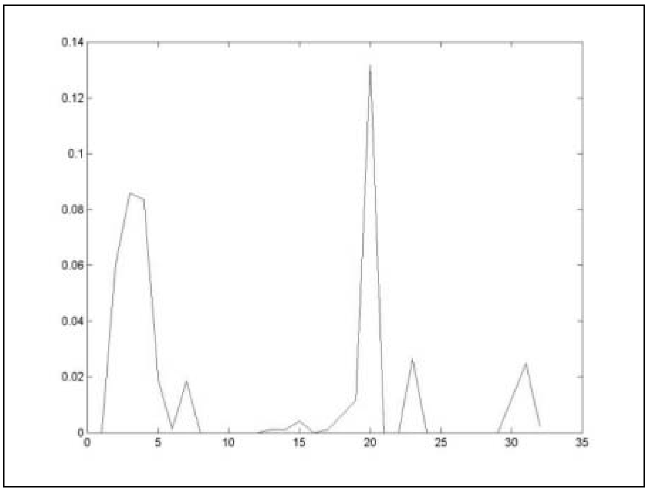

Figure 2: Decimated feature vector using the MAT method

3. EXPERIMENTAL RESULTS

In the experimental studies, Corel image database containing 1000 images is used [13]. HSV, YUV, and RGB color histogram features are extracted from the images using various bins. Histogram feature vectors often have high dimensions leading to high memory and processor usage in CBMIR. The experiments are done on a PC with Intel Pentium IV 2.8 GHz running Microsoft Windows XP operating system. MUVIS Content-Based Indexing and Retrieval system is utilized as the experimental framework. Several randomly selected images are queried separately using PCA, MDS, and the proposed MAT methods for comparison with equal scaling factors. In the experiments, MAT method is used with *G* function as maximum function. Table 1 presents the execution times for each method on one *8x8x8* bins HSV feature vector using scaling factor 4. Considering the given process times, the MAT method is practically more feasible particularly in time-critical cases.

Method	Time consumed (msec)
Principal Component Analysis (PCA)	14266
Multidimensional Scaling (MDS)	102453
MAT Method	1

Table 1: Execution times for each method

Figure 3 sketches the Precision-Recall curves of an example image query using original and decimated feature vectors, where the proposed method with scaling rates 4, 8, and 16 is employed for dimension reduction. In order to provide more reliable comprehension of the performance of the proposed method, Table 2 presents the average retrieval performance results for 50 image queries.

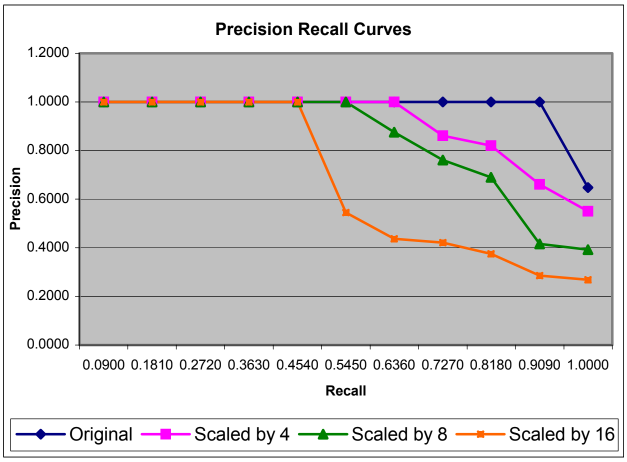

Figure 3: Precision-Recall curves for original feature vector and decimated feature vector by 4, 8, and 16

Employed Feature Vector	Retrieval Accuracy
Original	90%
Scaled by 4	89%
Scaled by 8	83%
Scaled by 16	82%

Table 2: Retrieval performance results with original and decimated feature vectors using the MAT method

The given experimental results imply that the MAT method does not affect image retrieval results significantly. Thus, it is a fast and flexible method that can be used for optimizing processing time and computational complexity of CBMIR systems.

4. CONCLUSIONS

The use of a simple and useful technique referred to as Mapping by Adaptive Threshold for reducing the dimension of feature data is investigated in CBMIR context. Dimension reduction methods contribute in reducing memory consumption and execution time, particularly in retrieval phase. In general, computational complexity of a dimension reduction method itself is more significant in the indexing phase.

The proposed MAT method is integrated into indexing and retrieval scheme for experimental studies. Similarly, PCA and MDS methods are also used in the experiments for comparison. The experimental results reveal the superiority of the proposed method over the other methods in terms of processing time. It is also shown that the MAT method does not affect query performance significantly and it can be employed in time-critical limited systems. However, it may not be suitable for other indexing and retrieval purposes, such as clustering.

Low computational complexity and high semantic performance of the proposed MAT method allows applying further feature data processing methods and complex indexing schemes. Furthermore, depending on the decision function used, the method might not modify the original feature vector values. This feature of the method may be useful in further processing of the feature data for other indexing and retrieval purposes.

Several other dimension reduction methods have been proposed in the literature. The proposed method will be compared to these methods in order to build the overall picture of dimension reduction in CBMIR. Afterwards, further studies may continue with utilizing combination of other methods. In this respect, more experiments and extended comparisons will be performed with the proposed method.

REFERENCES

[1] S. Kiranyaz, K. Caglar, O. Guldogan, and E. Karaoglu, “MUVIS: A Multimedia Browsing, Indexing and Retrieval Framework”, *Proc. of Third International Workshop on Content Based Multimedia Indexing*, CBMI’03, France, September 2003.

[2] A. Pentland, R.W. Picard, S. Sclaroff, “Photobook: tools for content based manipulation of image databases”, *Proc. of SPIE Storage and Retrieval for Image and Video Databases II*, 1994.

[3] S. F. Chang, W. Chen, J. Meng, H. Sundaram and D. Zhong, “VideoQ: An Automated Content Based Video Search System Using Visual Cues”, *Proc. of ACM Multimedia*, Seattle, 1997.

[4] L. Rui, Z. Yongsheng; F. Yonghong, D. Xueqing, “Research on content-based remote sensing image retrieval: the strategy for visual feature selection, extraction, description and similarity measurement”, *Proc. of. International Conferences on Info-tech & Info-net*, ICII 2001, Beijing, 2001.

[5] Linh V. T., Reiner L. “PCA-based representation of color distributions for color based image retrieval”, *Proc. of International Conference Image Processing*, ICIP 2001.

[6] Bingham E. and Mannila H., "Random projection in dimensionality reduction: applications to image and text data", *Proceedings of the 7th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD-2001, August 2001, USA, pp. 245-250.

[7] K. Fukunaga, Introduction to Statistical Pattern Recognition, *Academic Press*, 1990.

[8] Borg I., Groenen P., Modern *Multidimensional Scaling - Theory and Applications* (Springer), 1997.

[9] I.T. Jolliffe. *Principal Component Analysis*, Springer-Verlag, New-York, 1988.

[10] Young, F.W., Multidimensional Scaling, *In: Kotz &Johnson "Encyclopedia of Statistical Sciences*", V. 5, pp 649-658. Wiley & Sons, 1985.

[11] Kruskal, J.B. Nonmetric Multidimensional scaling: A numerical method. *Psychometrika 29*(2): 115-129, 1964.

[12] Beatty M., Manjunath B. S., "Dimensionality reduction using multidimensional scaling for image search, "*Proc. IEEE International Conference on Image Processing*, Santa Barbara, California, October 1997, pp. 835-838.

[13] Wang J. Z., Li J., Wiederhold G., ``SIMPLIcity: Semantics-sensitive Integrated Matching for Picture LIbraries," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol 23, no.9, 2001.