New Compression Techniques for Content-based Retrieval

1. Research Team

Project Leader: Prof. Antonio Ortega, *Electrical Engineering*

Graduate Students: Hua Xie

2. Statement of Project Goals

As international standards (the latest examples being MPEG-4 and JPEG 2000) continue to be developed and provide widely adopted solutions for media communications, some of the most important research problems in multimedia compression are becoming those where pure compression performance is not the main requirement.

Due to the proliferation of multimedia information over the internet, users are confronted with large amounts of content from many sources around the world. Techniques that enable users to efficiently search and exchange information are greatly needed. Content-based retrieval system was proposed to automatically annotate and index the multimedia information by their own audio/visual contents instead of text-based keywords that were manually entered. Most of current content-based retrieval systems are designed in a centralized fashion where feature extraction, indexing and querying are all done in a single database server. This paradigm faces problems of intensive computation and difficulty to scale up.

We argue that we are able to get over this limitation by having a distributed retrieval system where users share the data storage and query computation over the network. Users are able to search and exchange information by transmitting the features, which contain sufficient information for retrieval, to each other through the network. We argue that by compressing the features we are able to reduce both the transmission bandwidth and storage space in a great deal, without losing retrieval performance. Different from traditional compression techniques, which are designed to provide best perceptual quality under given rate constraint, we design novel compression techniques tailored for specific classification purposes.

3. Project Role in Support of IMSC Strategic Plan

Compression is a key component in providing media delivery to the broadest range of terminals and using all types of available transport mechanisms. The new compression technologies described in this project support the IMSC vision by allowing the integration of media delivery into current and future networks (e.g., the Internet) and physical channels (e.g., UWB). The growth in the area of databases requires the development of new compression algorithms that support the additional functionality required by databases where content-based retrieval is the objective. Finally, new compression techniques are also being developed to support storage and retrieval of new types of data, such as that produced by haptic devices.

4. Discussion of Methodology Used

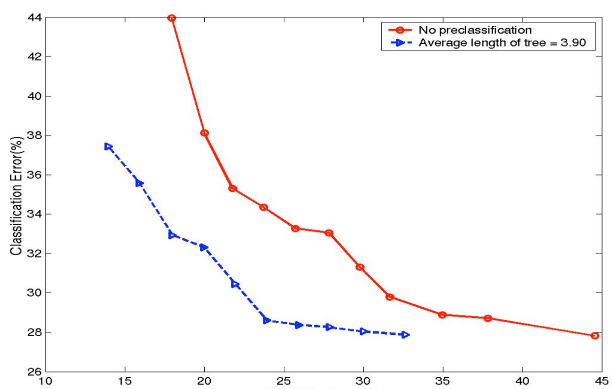
Linear discriminate analysis for transform coding in distributed image classification/retrieval systems. Image retrieval was formulated as an M-ary statistical classification problem. We examined the optimal transform, which can represent the class discrimination information with a sub-space with the lowest dimensionality. We proposed a greedy bit allocation algorithm to minimize the loss in class separability due to quantization. We analyzed the relations between proposed transform coding and Likelihood Ratio Quantization, and developed high rate analysis for certain classes of Gaussian distributions.

Learning of feature relevance from user's feedback using mutual information in content-based retrieval. In the context of content-based image retrieval, there is a gap between the high-level semantics and low-level visual features. User relevance feedback provides a way of getting the user in the loop. In this work, we presented a learning algorithm from the information-theoretic point of view, and designed a feature compression scheme in a distributed retrieval system with relevance feedback.

5. Short Description of Achievements in Previous Years

Classified quantization for compression in distributed image retrieval system. A partial classification is first performed before compressing the data so that we are able to capture the special characteristics of the classes that are relevant to content-based retrieval. The pre-classifier, which is an optimally pruned subtree of the original decision tree, and the quantization parameters for each classes are jointly searched based on a rate-distortion-complexity optimization framework. Substantial improvement in terms of retrieval performance vs. bit rate is achieved using proposed compression scheme as compared to standard encoding.

We performed texture classification with the compressed data and the result is shown in Figure 1. As is clearly shown in the figure, a lower classification error rate was achieved by using a classified encoder than a single encoder. At transmission rate of 20 bits/vector, a reduction of 11% in the classification error rate is achieved by using the classified encoder, with the pre-classification tree length equals to 3.9.



.

Figure 1: classification error vs. rate under different complexity constraints. We see that same classification performance can be achieved at lower rate at the cost of higher complexity.

Experimental results on Corel images are shown in Figure.2. We performed retrieval for K = 20. The quality of the retrieval result is measured by two quantities: Precision and recall.

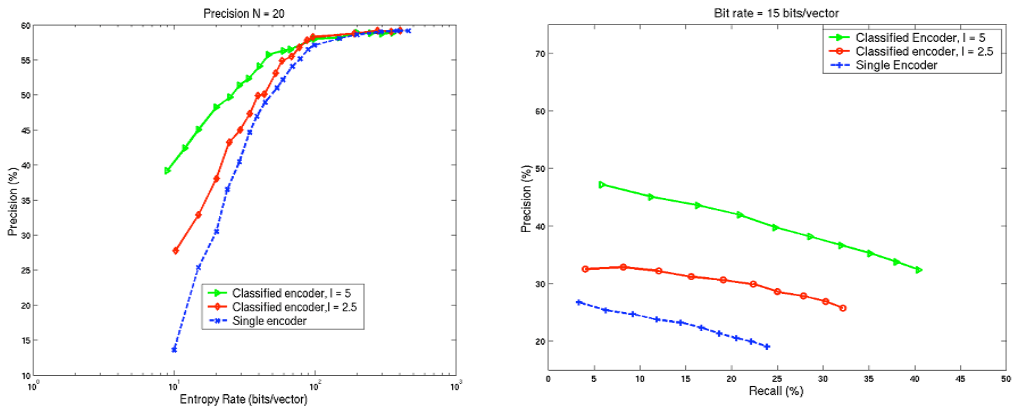


Figure 2: Retrieval performance comparing classified encoder and single encoder.

5a. Detail of Accomplishments During the Past Year

Linear discriminate analysis for transform coding in distributed image classification/retrieval systems. We pose image retrieval as a statistical classification problem and consider the design of a transform coding scheme minimizing the detrimental effect of compression. Different from traditional transform coding which was designed to provide the best reconstruction from compressed data, here the goal is to preserve the separation between the feature vectors generated by different classes. We consider a class separability criterion called scatter measure [2]. It is based on a second-order measure of quality that is defined completely in terms of second-order probabilistic parameters, i.e., means and covariance, of the empirical data. We consider the within-class scatter S_W, which is the scatter of samples around their respective class mean, and the between-class scatter S_B, which is the scatter of the expected vectors around the mixture mean:

$$S_W = \sum_{i=1}^{C} \pi_i E\{(X - M_i)(X - M_i)^t \mid y_i\}$$

$$S_B = \sum_{i=1}^{C} \pi_i (M_i - M)(M_i - M)^t$$

Where M_i and M are respectively, the mean vector for the ith class and the overall mean vector, π_i is the a priori probability. The Karhunen-Loeve method computes the optimal linear transform T_{klt}, which compacts most of the signal energy in lowest dimensional space. Let $S_m = E\{(X - M)(X - M)^t\} = S_W + S_B$ be the total scatter of the samples. Linear discriminate analysis defines the scatter ratio criterion, which is computed as $S = S_W{}^{-1} S_B$. The intuition behind this choice is that to have good class separation, we would like different classes to be as far apart from each other (S_B is large) as possible and, at the same time, samples belonging to the same class to be as closely clustered together (S_W is small) as possible. As an example, suppose that we have a two-class Gaussian distributed source with $x \sim N(M_i, \Sigma_i), i = 0,1$. Let the covariance matrices given by:

$$\Sigma_0 = \Sigma_1 = \begin{bmatrix} \sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \sigma^2 \end{bmatrix} \text{(Gaussian Markov source):} \quad (1)$$

We show in Figure. 3 (a) the principal axis of KLT transform and Fisher discriminate for Gaussian Markov source with $\rho = 0.2$. We can clearly see that the two transforms results in different bases. Figure 3 (b) plots the angle between the principal axis and fisher discriminate as a function of the correlation coefficient ρ. We can clearly see that: the more correlated the dimensions are, the more different the two transforms are from each other.

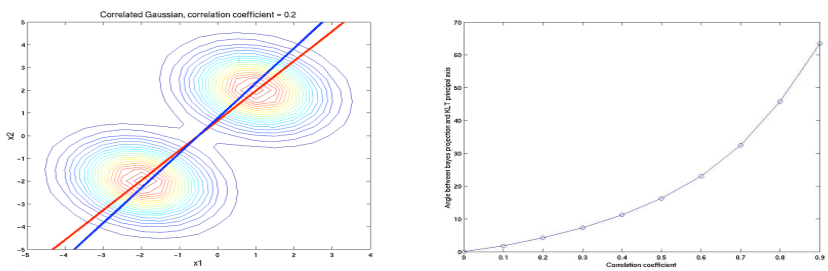


Figure 3: (a) Basis of Linear Discriminant Analysis transform, and KLT for 2-dimensional correlated Gaussian source. (b) Angle between the principal axis and Fisher's discriminate as a function of the correlation coefficient for Gaussian Markov Sources.

We propose a greedy algorithm to perform bit allocation in transform domain such that classification information from all dimensions can be employed in quantizer design. We show that our proposed transform coding scheme achieves much better classification performance than traditional KLT transform coding, based on both real data (Brodatz textures) and synthesized random source. Please refer to [3] for details. We show in Figure 4 (a) the classification performance comparing KLT transform coding and proposed LDA transform coding. We can clearly see that our proposed transform coding scheme achieves much better classification performance than the traditional KLT method. Figure 4 (b) shows the bases of KLT and LDA for this example.

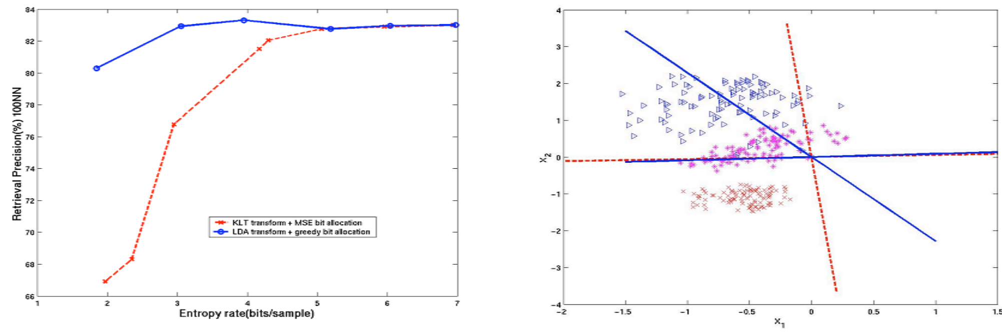


Figure 4: (a) Comparison of the classification performance based on two compression schemes. Solid circle: Proposed LDA transform coding, Dashed star: KLT transform coding. (b) The bases of KLT and LDA for this example.

We show in Figure 5 the Chernoff distance between the two distributions $\hat{f}_0$ and $\hat{f}_1$ of the quantized data using KLT transform coding and our proposed LDA transform coding for synthesized Gaussian Markov source with covariance matrices given by Equation (1).

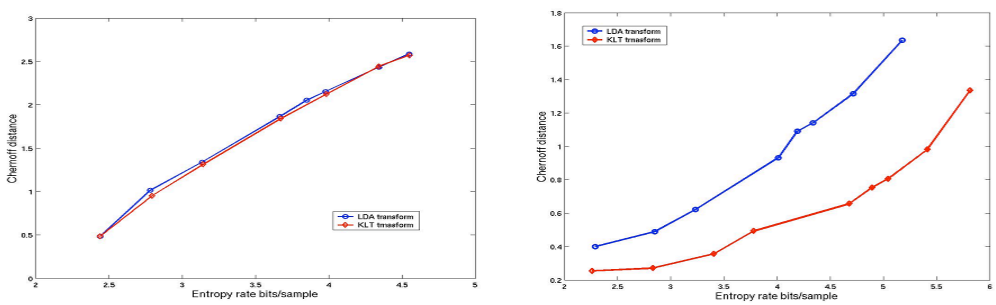


Figure 5: Chernoff distance between the two distributions of compressed data for Gaussian Markovian sources with correlation coefficient (a) 0.2 (b) 0.8.

We can clearly see that proposed algorithm provided much better separation between classes after compression than traditional transform coding, comparing at the same rates.

Learning of feature relevance from user's feedback using mutual information in content-based retrieval. Due to the development of devices for image creating and capturing, both military and civilian equipment generates giga-bytes of images each day. Content-based image retrieval allows users to access the large, unannotated digital image databases. Given a query image, the goal is to find the images that have *similar contents* as the query image. One problem with CBIR is that there exists a gap between high-level semantics and low level features. Relevance feedback, as a natural way of getting the user in the loop, has been shown to provide dramatic performance boost in CBIR systems.

In this work, we study the feature relevance learning from information-theoretic of view, and adapt the distance metric according to the feature relevance. We also examined the optimal feature compression scheme for a distributed CBIR system where the relevance feedback is to be performed at the client.

The intuition on feature relevance in terms of retrieval is that: the most relevant feature should be the one that keeps most information about the class labels. Figure 5 shows an example where adapting the distance metric will return more relevant data to the query.

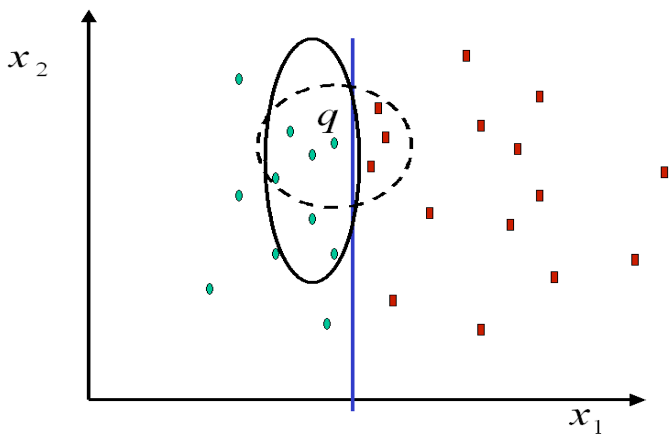


Figure 6: Adapting the distance metric according to feature relevance will increase the precision of Nearest-Neighbor classification.

We use mutual information $I(X_i, Y)$ to measure the relevance of feature X_i:

$$I(X,Y) = \int_x f(x) \sum_y p(y \mid x) \log\left(\frac{p(y \mid x)}{p(y)}\right) dx$$

Transforming into weighting scheme, we have:

$$w_i = \exp(\beta \cdot I(X_i, Y)) / \sum_{k=1}^{F} \exp(\beta \cdot I(X_k, Y))$$

Figure 7 shows an example of retrieved images after certain relevance feedback using proposed algorithm.

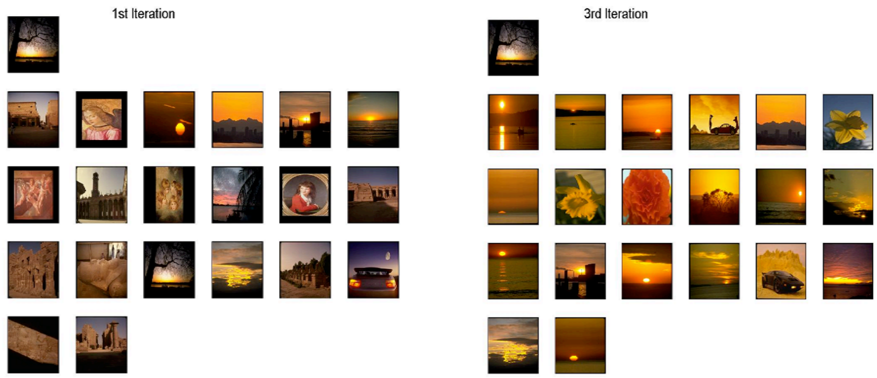


Figure 7: Demonstration of retrieval performance improvement using relevance feedback.

We compared our algorithm with MARS [4] and the result is shown in Figure 8. We can see that our proposed algorithm outperforms MARS scheme.

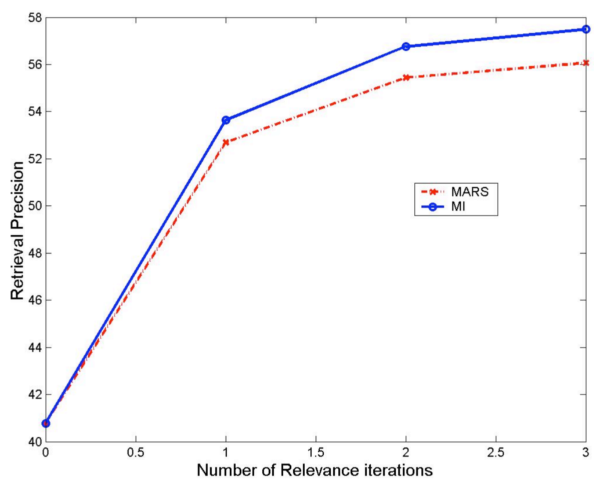


Figure 8: Comparison of retrieval performance between MARS and proposed MI scheme.

6. Other Relevant Work Being Conducted and How this Project is Different

Compression and classification: A drawback in using standard coding techniques is that these have been designed in order to preserve perceptual quality as judged by a human observer/listener. However, when the decoded media is to be used by a classifier (or recognizer) rather than a human, it is possible to achieve better performance (i.e., same recognition rate using fewer coding bits) by designing a compression algorithm that is optimized for the specific classification task to be performed. Here instead of minimizing just the perceptual distortion it is also desirable to minimize the classification error introduced by quantization. Some examples include querying an image database to retrieve an image based on its content or distributed speech recognition, where the recognition is performed using encoded acoustic feature.

Relevance feedback for content-based image retrieval: Instead of heuristically updating the weights or optimizing the distance within the relevant set, we use the information-theoretic measure of mutual information to determine the feature relevance in terms of differentiate relevant images from non-relevant ones. It was shown in [5] that informax minimizes a lower bound on Bayes Error.

7. Plan for the Next Year

New compression techniques for distributed content-based image retrieval

- Design of quantizer that aims at preserving the mutual information of the feature and class labels. When the data are to be compressed and transmitted to the database server for classification, the goal of the encoder is to transmit as much as possible "relevant information" through this "information bottleneck", where the relevant information is the information that the feature provides about the class label.
- Design of feature relevance learning algorithm with multiple classes. One problem with relevance feedback being modeled as a two-class classification was that: "good" images are good in the same way, while "bad" images are bad in their own ways. One remedy to it would be to use a multi-class model instead of a two-class model.
- Application of quantization scheme to compress the database. We use them as compact descriptors of the contents (vector approximation file [6]), and design efficient searching algorithms based on the VA- file.

8. Expected Milestones and Deliverables

Application of Mutual Information for relevance feedback in distributed CBIR system
Application of quantization designs to distributed CBIR.

9. Member Company Benefits

N/A

10. References

[1] H. Xie, A. Ortega, "Entropy- and Complexity- constrained classified quantizer design for distributed image classification", IEEE International workshop on Multimedia Signal Processing, St. Thomas, U.S. Virgin Islands, Dec.2002
[2] Keinosuke Fukunaga, "Introduction to statistical pattern recognition", Academic Press, 1991.
[3] H. Xie, A. Ortega, "Exploration of linear discriminate analysis for transform coding in distributed image classification", in 37th Asilomar Conference on Signals, systems, and Computers. Pacific Grove, CA, November 2003.

[4] Y. Rui and T. Huang, “Optimizing learning in image retrieval”, Proceeding of IEEE int. Conf. On Computer Vision and Pattern Recognition, Jun. 2000.
[5] N. Vasconcelos, “Feature Selection by Maximum Marginal Diversity: optimality and implications for visual recognition”, Proc. of the 2003 IEEE Computer Society Conference on Computer Vision and Pattern recognition.
[6] R. Weber, H. -J. Schek, and S. Blott. “A quantitative analysis and performance study for similarity-search methods in high-dimensional spaces”. In Proceedings of the Int. Conf. on Very Large Data Bases, pages 194--205, New York City, New York, August 1998.