Supporting Content-based Queries over Images in MARS *

Sharad Mehrotra Yong Rui Michael Ortega-Binderberger Thomas S. Huang

1. Introduction

While advances in technology allow us to generate, transmit, and store large amounts of digital images, video, and audio, research in indexing and retrieval of multimedia information is still at its infancy. To address the challenges in building an effective multimedia database system we have built a Multimedia Analysis and Retrieval System (MARS) prototype [1]. This paper summarizes the retrieval subsystem of MARS and its support for content-based queries over image features. Detailed discussion of the algorithms used can be found in [2]. Content-based retrieval techniques have been extensively studied for textual documents in the area of automatic information retrieval [3, 4]. Our objective in MARS is to exploit these existing techniques for content-based retrieval over images and multimedia databases.

2. Image Model Used in MARS

In MARS, an image is represented as a collection of automatically extracted low-level image features (e.g., color, texture, shape and layout), as well as manual text descriptions of the image.

The retrieval performance of an image database is inherently limited by the nature and the quality of the features used to represent the image content. Image features used in MARS are described below. A detailed discussion on the rationale of the chosen feature representations can be found in [5, 1, 6, 7, 8].

Color Features: represented in the HSV space due to its de-correlated and uniform coordinates. It is represented using an 8×8 2D histogram over the HS coordinates and the V coordinate is dropped since it is easily affected by the lighting condition. The histogram intersection measure of distance provides an accurate and efficient measure of (dis)similarity between two images based on their color [9].

Texture Features: represented via a modified version of CCD (*coarseness, contrast*, and *directionality*) developed in [5]. The *Coarseness* measures granularity of the texture (fine vs coarse) and is represented by a coarseness histogram. *Contrast* represents the distribution of luminance in the image. A single number represents this component. *Directionality* is a measure of how "directional" the image is. Similar to coarseness, it is represented by a directionality histogram.

Similarity between two textures is determined using a weighted sum of the Euclidean distance among contrasts and the intersection distances for the other two histogram based components.

Shape Features of an object in an image are represented by its boundary. The boundary of an object is represented using a modified Fourier descriptor (MFD) we developed in [8] and euclidean distance used to measure similarity based on shape of objects.

Layout Features. MARS supports representation of color and texture layouts in an image. To extract the layout features, the image is split into 5×5 subimages. Color and texture features are extracted from each subimage and stored in the database. To compute similarity based on layout, first similarity is computed with respect to each subimage. Then a weighted sum is used to compute the layout distance.

3. Retrieval Models Used In MARS

MARS attempts to exploit extensive research in the information retrieval literature for content-based retrieval over images. A query is graphically constructed by selecting certain images from the collection. The query is interpreted as a Boolean expression over image features and a Boolean retrieval model (adapted for retrieval over images) is used to retrieve a set of images ranked based on the degree of match to the query.

Since the original Boolean Model does not support ranked retrieval, the model is appropriately extended to support a ranked list of answers. First, to facilitate ranking based on a boolean expression on multiple features, similarity measures over all the features are normalized to a common scale between [0,1] where a 0 represents no match and a 1 represents a complete match. Normalization allows us to directly compare similarity measures across different features. Then one of the following two extentions to the basic Boolean Models are used to generate the overall ranking of images based on the boolean query expression.

- **Fuzzy Boolean Retrieval** considers the distance

*This work was supported in part by the Army Research Laboratory under Cooperative Agreement No. DAAL01-96-2-0003 and in part by NSF/DARPA/NASA Digital Library Initiate Program under Cooperative Agreement No. 94-11318.

between the image and the query feature as the degree of membership to the fuzzy set of images that match the query feature. Fuzzy set theory is then used to interpret the Boolean query and the images are ranked based on the their degree of membership in the set.

- **Probabilistic Boolean Retrieval** considers the distance between the image and the query feature to be a measure of probability that the image matches the user's information need. Feature independence is exploited to compute the probability of an image satisfying the query which is used to rank the images.

While the above discussed retrieval models support a ranked retrieval of images, for the approach to be useful, techniques must be developed to retrieve the best N matches efficiently without having to rank each image. Such a technique consists of two steps:

- retrieve images in ranked order based on each feature variable in the query.
- combine the results of the single feature variable queries to generate ranked retrieval for the entire query.

MARS incorporates pipelined approach to query processing to efficiently evaluate the best N matches to a given query.

4. Experimental Evaluation

For our experiments we used a collection of images of ancient African and Peruvian artifacts from the Fowler Museum of Cultural History.

To demonstrate the retrieval quality, we chose 13 typical conceptual queries with known outcomes defined by a human expert; three such queries are "all stools", "stone masks" or "golden pots". This methodology is consistent with that used in the information retrieval community. A set of queries with known outcomes are used to evaluate precision recall curves for each query. The results obtained were very encouraging obtaining up to 60% precision at 50% recall.

In general, our probabilistic model performed better than the fuzzy model. This may be attributed to more information being preserved in the ranking functions used for combination based on *and* and *or* operations.

5. Conclusions

Most existing content-based image retrieval systems also extract low-level image features like color, texture, shape, and structure [5, 1, 6, 9]. These systems, however, support queries on single features separately which we refer to as *simple queries* which limits their usefulness to end users.

In contrast, similar to the approaches taken in information retrieval system, the approach we have taken in developing MARS is to support an "intelligent retrieval" model using which a user can specify their information need to the image database and obtain a ranked list of images. The retrieval model supported is a variation of the Boolean model based on probabilistic and fuzzy interpretation of distances between the image and the query.

While our results are encouraging and illustrate that using IR techniques for content-based retrieval in image databases is a promising approach, much further work needs to be done to to validate the effectiveness of using IR techniques for image retrieval. Some extensions we are currently exploring include:

- Techniques for automatic feature weighting in image databases and incorporating relevance feedback to improve retrieval performance.
- Development of indexing techniques to support efficient ranked retrieval on single features. Currently, our algorithms for Boolean query evaluation assume presence of such an index.

References

[1] T. S. Huang, S. Mehrotra, and K. Ramchandran, "Multimedia analysis and retrieval system (MARS) project," in *Proc of 33rd Annual Clinic on Library Application of Data Processing - Digital Image Access and Retrieval*, 1996.

[2] M. Ortega, Y. Rui, K. Chakrabarti, , S. Mehratra, and T. S. Huang, "Supporting similarity queries in MARS," in *(submitted to)ACM Conf. on Multimedia*, 1997.

[3] G. Salton and M. J. McGill, *Introduction to Modern Information Retrieval*. McGraw-Hill Book Company, 1983.

[4] J. P. Callan, W. B. Croft, and S. M. Harding, "The inquery retrieval system," in *Proc of 3rd Int Conf on Database and Expert System Application*.

[5] M. Flickner, H. Sawhney, W. Niblack, and J. Ashley, "Query by image and video content: The qbic system," *IEEE Computer*, 1995.

[6] M. Miyahara, "Mathematical transform of (r,g,b) color data to munsell (h,s,v) color data," in *SPIE Visual Communications and Image Processing*, vol. 1001.

[7] Y. Rui, A. C. She, and T. S. Huang, "Automated shape segmentation using attraction-based grouping in spatial-color-texture space," in *Proc. IEEE Int. Conf. on Image Proc.*, 1996.

[8] Y. Rui, A. C. She, and T. S. Huang, "Modified fourier descriptors for shape representation – a practical approach," in *Proc of First International Workshop on Image Databases and Multi Media Search*.

[9] J. R. Smith and S.-F. Chang, "Automated binary texture feature sets for image retrieval," in *Proc ICASSP-96*.