??, ??, 1–15 (??)

A Region-Based Representation of Images in MARS

SERGIO D. SERVETTO, YONG RUI, KANNAN RAMCHANDRAN, THOMAS S. HUANG*

{servetto,yrui,kannan,huang}@ifp.uiuc.edu

Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign. 405 N. Matthews St., Urbana, IL 61801

Received ??; Revised ??

Editors: ??

Abstract. We study the problem of representing images within a multimedia Database Management System (DBMS), in order to support fast retrieval operations without compromising storage efficiency. To achieve this goal, we propose new image coding techniques which combine a wavelet representation, embedded coding of the wavelet coefficients, and segmentation of image-domain regions in the wavelet domain. A bitstream is generated in which each image region is encoded independently of other regions, without having to explicitly store information describing the regions. Simulation results show that our proposed algorithms achieve coding performance which compares favorably, both perceptually and objectively, to that achieved using state-of-the-art image/video coding techniques while additionally providing region-based support.

Keywords: Image Databases, Image Compression.

1. Introduction

With advances in the computer technologies and the advent of the World-Wide Web, there has been an explosion in the amount and complexity of digital data being generated, stored, transmitted, analyzed, and accessed. Much of this information is multimedia in nature, including digital images, video, audio, graphics, and text data. Among the many types of media, images are of prime importance. Not only images are the most widely used media type besides text, but they are also one of the most popular means for representing and retrieving video and other multimedia information. This paper is dedicated to efficient storage, transmission and retrieval of images in the context of multimedia databases.

Image retrieval can be traced back to the 1970's, when the keyword annotation approach was the traditional image retrieval paradigm. In this approach, images are first annotated manually by keywords. They can then be retrieved by their corresponding annotations. While it has been very effective for many years, there are two major difficulties with this approach, due to the large amount of manual effort required to develop these annotations: the differences in the interpretation of image contents, and inconsistencies in the keyword assignments generated by different indexers [21, 9, 10]. Furthermore, as the size of image repositories increases, the keyword annotation approach becomes infeasible.

*This work was supported in part by the National Science Foundation under grant NSF MIP-9409587 (RIA-94), by the Army Research Lab under grant DAAL01-96-2-0003, and also by the NSF/DARPA/NASA DLI Program under Cooperative Agreement 94-11318.

To overcome the difficulties of the annotation based approach, an alternative mechanism, Content-Based Image Retrieval (CBIR), has been proposed in the early 1990's. Besides using human-assigned keywords, CBIR systems use the visual content of the images as the image index (e.g. color, texture, and shape features). This greatly alleviates the difficulties of the pure annotation based approach, since the feature extraction process can be automated, while ensuring that the image content is always "consistent". Since its advent, CBIR has attracted significant research attention, including government [14, 15], industry [5, 7, 10], and academia [13, 17, 18, 23, 35]. In fact, ISO/IEC has launched a new work item, MPEG-7 [1, 2, 3], to define a standard Multimedia Content Description Interface. Many CBIR systems, both commercial [5, 7, 21, 9, 10] and academic [13, 17, 18, 23, 35], have been developed recently.

Despite this extensive research effort, there remain many challenges to be addressed before a successful image retrieval system becomes more viable and practical:

- Most existing systems only support queries based on entire images. However, more often than not, it is image *regions* and not the whole image that contain salient visual features and important objects. How to support region-based queries over the image database is one of the key issues in CBIR.
- Most of the current research efforts focus on the retrieval aspect, while the storage and transmission aspects (i.e., the *systems* aspects) of image databases have been relatively less studied. However, since images constitute storage-intensive media, efficient storage techniques are of paramount importance. Furthermore, the explosion of Internet activity has placed heavy demands on the querying and access of remote databases as well as local ones. Given the relatively low transmission bandwidth of the Internet, efficiency in the representation of images becomes a major factor affecting the overall system performance.

To address some of these challenging research issues involved in multimedia databases, the MARS (Multimedia Analysis and Retrieval System) project was started during the Spring of 1995 at the University of Illinois [13, 22, 26, 28, 29]. MARS supports retrieval of image, video, and audio data. A brief description of MARS is presented in Section 2.

The main contribution offered in this paper is the development of a new representation of images, to support access to individual image regions directly in the compressed domain. We extend the still image coding algorithms of [32], to provide functionality required by MARS. We do so by combining the right set of tools in the design of algorithms capable of generating compressed bitstreams with the desirable properties discussed above. Two salient features of the proposed representation are (a) that it results in a negligible rate/distortion performance degradation, when compared to the compressed bitstream generated by similar state-of-the-art techniques which do not support regions; and (b) that it has a "unified" architecture, characterized by a fairly low computational complexity, in which we do not need to deal with region maps and image data separately. The inputs to our new proposed algorithms are the image and a segmentation region map, computed by a subsystem in MARS. Our proposed encoders partition the set of all wavelet coefficients of the input image into subsets associated with each image region, and then encode these sets independently of each other. Fully embedded bitstreams are produced for each set, that allow multiresolution display on a *per-region basis*.

The rest of the paper is organized as follows. Section 2 gives a brief description of the system architecture of MARS, and highlights the important role of the proposed region based storage module in this system. In Section 3 we discuss our proposed wavelet based region storage technique in detail. Extensive experimental results over various types of images of the proposed coders are given in Section 4. Concluding remarks and future research directions are given in Section 5.

2. An Overview of MARS

In this section only a brief description of MARS is presented. For a detailed description, the reader is referred to [13, 20]. Figure 1 shows the main components of the system, described below.

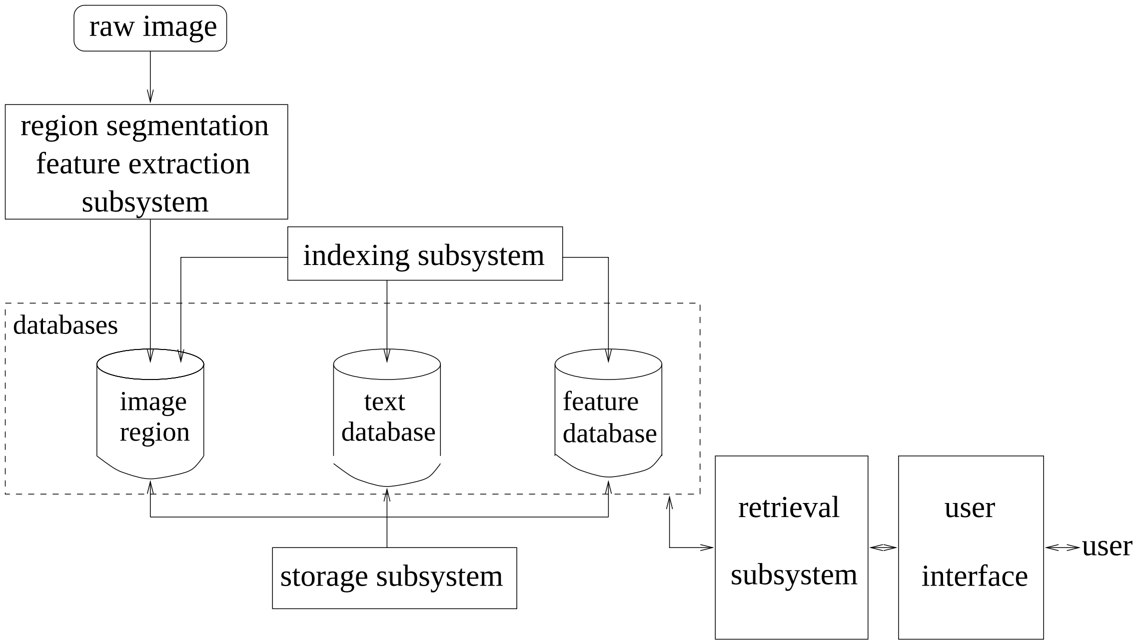


Fig. 1. The system architecture

User Interface. This is the interface where the user can interact with the system. The current implementation is written in Java applets and accessible over the Internet. The user interface allows users to graphically pose content-based queries as well as traditional text-based queries.

Retrieval Subsystem. It takes the query specified at the user interface, evaluates the query using the information stored in all the three databases, and returns to the user images that are best matches to the input query. The query language supported allows users to pose complex queries that are composed using low-level image features as well as textual descriptions [20]. Another unique feature of the MARS retrieval subsystem is that it supports *relevance feedback* [26, 28]. Relevance feedback is the process of automatically adjusting an existing query using the information fed back by the user about the relevance of previously retrieved objects such that the adjusted query is a better approximation to the information need of the user [6, 31]. This approach greatly reduces the user's effort of composing a query and captures the user's information need more precisely.

Indexing Subsystem. For large image collections, it is highly desirable to match queries without searching the entire image collection. To achieve this, multi-dimensional indexing techniques need to be utilized. The current MARS-image system uses a clustering based indexing approach to facilitate the fast search [25].

Region Segmentation Subsystem. As discussed in the introduction, image *regions* may contain more important information than the entire image does. In order to extract features from image regions, image segmentation has to be performed first. The MARS-image system uses a color-texture-spatial based grouping technique to segment image regions [29]. The output of this subsystem is an image region map consisting of multiple image regions. An example segmentation is shown in Figure 2.

Image segmentation has made considerable progress in the past few years [12, 11, 34]. For specific applications, such as GIS systems, reasonably good results have been achieved [16]. However, in general unconstrained domains, image segmentation is still an open problem in Computer Vision and Im-

(a)

(b)

Fig. 2. Example of a segmentation of an image: (a) an original image; (b) a segmentation map.

age Understanding. The MARS Region Segmentation Subsystem has made no attempt at solving this problem. Our goal is not to segment out high-level image *objects*, but rather to identify low-level homogeneous image *regions*, since these contain all the information needed by our system to perform query evaluation.

Feature Extraction Subsystem. MARS supports both global (whole image) features as well as local (image region) features. The features used in the system are color, texture, and shape [20]. Specifically, we use color histogram and color moments as the color feature representations [36]; coarseness-contrast-directionality, wavelet based texture, and co-occurrence matrix texture as the texture representations [26]; and modified Fourier descriptors as the shape feature representation [27].

Storage Subsystem. Not only does MARS support efficient storage of the entire image, it also supports efficient storage of individual image regions. In the rest of the paper we will describe the proposed region based storage technique in detail. Note that efficient storage is important to both local disk storage and to transmitting content over the network to support remote queries.

3. Region-Based Representation of Images

3.1. *Storage of Images in DBMS's*

Classical DBMS's provide very limited support for the storage of image data. The main complication with the integration of image data into existing DBMS's lies in the fact that operations such as testing for equality and order relations, essential to the evaluation of queries, are meaningless when applied raw image data. The main goal of MARS is to extend classical DB functionality to other media types as well.

If images are to be stored in multimedia DBMS's, all the *systems* issues related to the efficiency with which data can be retrieved from physical storage gain dramatic importance. This is so because images, even in compressed form, involve significantly larger amounts of data than those typically managed by DBMS's, consisting mostly of numbers and strings of characters. Furthermore, with the popularity of computer networks, DMBSs may manage data distributed over a number of interconnected machines, in which case, in assessing the performance of such systems, delays due to network transmission have to be added to delays due to physical retrieval from disk. Due to all these reasons, the way in which an image is represented internally (i.e., how an image is compressed and physically stored in the DBMS) is a most important issue.

In order to provide low access times, the amount of data to be effectively transferred between the

(a) (b) (c) (d)

Fig. 3. Sample reconstructions obtained from a single bitstream, by specifying different decoding parameters: (a) an original image (Luciano); (b) a segmentation map; (c) a reconstruction obtained by selecting bits corresponding to Luciano's face only, until the available bitrate is exhausted; (d) another reconstruction, now obtained by selecting only background bits. In between the two extremes given by (c) and (d) many different intermediate reconstructions are possible.

image storage subsystem and the image display client must be kept to a minimum. To accomplish this goal, we propose a new image representation in which (a) we identify sub-bitstreams of the compressed bitstream corresponding to arbitrarily-shaped regions in the image domain, and (b) we can decode these sub-bitstreams independently of each other without needing to decode the whole image to full resolution. An example of the sought functionality is depicted in Fig. 3.

We would like to make the following remarks:

- The goal of reducing the amount of I/O needed to retrieve images implies that compression performance should not be sacrificed in order to provide access to individual image regions directly in the compressed bitstream.
- This representation offers advantages in terms of computational complexity: it is a poor strategy of CPU usage to decode an entire image, to maybe retrieve only a small subimage of interest, for which a low resolution version would have probably been good enough.
- From a coding perspective, the problem of providing access to independent regions comes basically in two "flavors": either the region is already known to the decoder (in this case, the user formulating a query), or it is not. Always from a coding perspective, if a particular region is already known to the decoder, our problem is nearly trivial, since it only involves sending an appropriate subset of wavelet coefficients. In this paper, we focus instead on the case in which the regions are not known to the decoder, and have to be encoded somehow. Later in this section we present two alternative solutions to this problem: one in which the region is explicitly sent to the decoder as side information, and another one in which the regions are *implied* by the subsets of wavelet coefficients sent. In Section 4, we present results on the performance of each of these solutions.

3.2. Region-Based Coding Using Wavelets

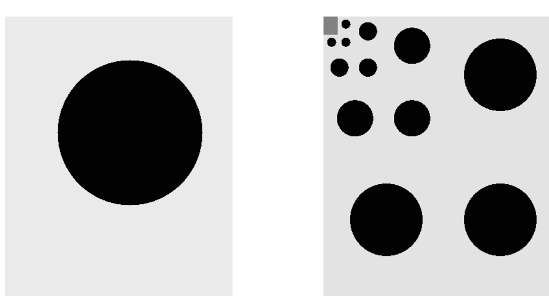
Wavelet techniques have received significant attention in the last few years in a variety of areas related to Signal Processing (Vetterli and Kovačević [38] provide an excellent tutorial introduction to wavelets). In the context of image data compression this was, to a large extent, due to the high rate/distortion performance achieved by Zerotree based algorithms [30, 33, 39]. A novel alternative representation to the standard Zerotrees was proposed in [32], based on morphological operators. This new representation efficiently captures *arbitrarily* shaped sets of subband coefficients, identified to be the most important ones to encode accurately.

Specifically for region-based coding, wavelets offer a definite advantage over other transform domain techniques. Based on the joint space-frequency localization property of the wavelet transform [38], coefficients in the wavelet domain can be uniquely associated with image regions in the spatial domain. Although classical transforms used for image compression (e.g., the Discrete Cosine Transform (DCT) [4]) do not have the joint space-frequency localization property for signal energy, such localization can be forced by partitioning the image into blocks, and applying the transform within each block (e.g., block DCTs). However, from a compression standpoint, this artificial approach has been found to perform less satisfactorily than wavelets, which provide the localization property naturally as part of their construction. As a result, meaningful properties of the data in the spatial domain (e.g., that certain pixels belong to some image region, that this group of pixels is located to the left of those, etc.) can be easily mapped to the wavelet domain, as shown in Fig. 4.

The wavelet representation has been used in the development of most of the highest performance image coding algorithms [30, 32, 33, 39], and recently also to video coding algorithms [40]. Therefore, it should be possible to design an image representation in which regions are coded independently of each other, while still retaining the high rate/distortion performance of state-of-the-art image coding algorithms, and without incurring in an excessively high computational cost.

3.3. Basic Reference System

In order to encode arbitrarily shaped regions of image data, we first partition wavelet coefficients into various sets corresponding to each of the image regions in the spatial domain, and then encode each of these sets independently using a previously developed morphological image coding algorithm [32]. Observe that a penalty in coding performance is to be paid by taking this approach, since there may exist statistical dependencies among coefficients in different sets that could be used for further compression. In order to determine how much coding performance degrades due to our particular way of encoding image regions, in Section 4 these variations are compared against the performance of the basic coder used to build them. A brief summary of the key components of that coder is presented in this subsection.



(a) (b)

Fig. 4. To illustrate how the spatial region maps have an equivalent in the wavelet domain: (a) a sample region map; (b) the equivalent map in the wavelet domain.

3.3.1. Morphological Coding of Sets of Wavelet Coefficients A fundamental building block for all constructions presented in this paper is the fully embedded still image coder presented in [32], based on morphology and its applications to multidimensional signal processing problems [19]. The standard architecture for transform coding methods consists of a linear decorrelating transform, followed by scalar quantization of the transform coefficients, and final entropy coding of the quantized coefficients.

Experimental results have been reported [24, 32], establishing the presence within image subbands of at least two sets of coefficients having dissimilar statistical properties. It was suggested in [24], and effectively shown in [32], that morphology may provide an efficient means of identifying and encoding such sets of coefficients. That work presents one particular algorithm to compute and encode an image-dependent subband classification map, in a uniquely decodable way, making use of only a minimal amount of side information.

3.3.2. Quantization Each coefficient is quantized with a sequence of embedded quantizers $\{Q_0 \dots Q_{n-1}\}$ [37]:

$$Q_i(x) = \begin{cases} \mathbf{z} & |x| < T/2^{i+1} \\ \mathbf{p} & T/2^{i+1} \le x < T/2^i \\ \mathbf{n} & -T/2^i < x \le -T/2^{i+1} \end{cases}$$

for $i = 0 \dots n-1$, and for some parameter T (typically T is the largest magnitude in the data stream to be encoded). By successively applying these quantizers to each wavelet coefficient, a symbol stream is generated with the property that as

n is allowed to increase, a sequence of reconstructions is generated that converges to the original image, and such that any initial substream corresponds to some lower resolution encoding.

3.3.3. Entropy Coding Unlike the case of optimal quantization for a single target bitrate, the design of optimal embedded quantizers is a problem that cannot be solved in general. Equitz and Cover [8] showed that optimality can only be achieved when the source has a certain Markov property. By optimality, we refer to the ability of an encoder to achieve the theoretical R-D performance bound for the source being coded.

In this paper, for practical reasons, we take the approach of fixing beforehand the quantizers to be used, independently of the image to encode. As a result, the performance of this embedded coder is entirely determined by the efficiency with which the resulting symbol stream is entropy coded. To achieve maximum efficiency, conditional arithmetic coding based multiple probability tables is used on each resolution level [32].

3.4. Region-Based Image Coding with Explicit Side Information

In this case, image regions are encoded by first encoding a segmentation map defining the regions, and then encoding subband data assuming the map is known. After the map is known, the encoding algorithm takes a particularly simple form in this case: just apply the standard morphological coder as many times as there are objects in the image, but each time considering only coefficients drawn from a single object.

Efficiency in encoding the segmentation map is critical to the success of this approach, because of two main reasons: (a) this map will have to be decoded each time an object in the image is accessed, and therefore computational complexity of the decoder becomes an issue; and (b) this map is side information, which eats into the total bit budget. We entropy code the map as one more independent bitstream (the other ones being the image regions). Pseudocode for this algorithm is shown in Table 1.

For simplicity, this description is given only for a binary segmentation: the generalization to multiple objects is conceptually trivial, and there is no coding performance degradation associated with increasing the number of objects (other than the increase in description complexity for the segmentation map). Note that the DC subband is encoded separately from the objects comprising the image. The reason for doing this is to provide a coarse resolution approximation to the entire image, and in this way provide refinements only for those objects of interest to the user. Alternatively, one could segment the purely lowpass subband too, and treat it as just as any other subband for coding. Which of the two techniques is more appropriate would be determined by specific applications. In MARS, we feel that providing some coarse approximation to the undesired objects is a better solution than providing just a "black" image for these objects.

There are many possible ways in which the segmentation map can be entropy coded (step 1). If a specific application has *a priori* knowledge about the shape and/or location of objects present in typical images it deals with, then that information can be used to reduce the number of bits necessary to encode the segmentation map. In general, such specific prior information is not available, and therefore a technique is called for that can guarantee high coding performance under all circumstances. After experimenting with a variety of techniques (including adaptive and context based arithmetic coding), the Lempel-Ziv algorithm was chosen as the one consistently giving the best results.

3.5. Region-Based Image Coding without Explicit Side Information

The main difference between this case and the case in which image regions are sent explicitly to the decoder as side information is that here the information describing image regions is not available at the decoder. To deal with this problem, given the original image and the map of regions, the encoder creates one image per object in the map, and then encodes each of them separately and independently. Pseudocode for this algorithm is shown in Table 2.

As in the previous case, without loss of generality the algorithm is described only for a binary

Table 1. Pseudocode Description for the Encoder, assuming an Uninformed Decoder with Explicit Side Information

Input:

An image, and a binary segmentation map of image pixels.

Output:

Four compressed bitstreams.

Algorithm:

1. Entropy code the segmentation map. This is the first bitstream.
2. Take the wavelet transform of the image.
3. For each image subband, appropriately subsample the segmentation map, to obtain a segmentation of the wavelet coefficients for that subband.
4. Encode the DC subband. This is the second bitstream.
5. Apply the fully embedded morphological coder to the wavelet coefficients; do not scan the entire rectangular array, encode only those coefficients whose label is "1" in the binary map. This produces the third bitstream.
6. Repeat step 4, but now sending only coefficients labeled "0" in the binary map. This produces the fourth bitstream.

Table 2. Pseudocode Description for the Encoder, assuming an Uninformed Decoder and no Side Information

Input:

An image, and a binary segmentation map of image pixels.

Output:

Three compressed bitstreams.

Algorithm:

1. Take the wavelet transform of the image.
2. For each image subband, appropriately subsample the classification map, to obtain a classification for that subband of the wavelet coefficients.
3. Define two new wavelet fields: one in which coefficients labeled "0" are kept intact and coefficients labeled "1" are set to zero, and another one in which coefficients labeled "0" are set to zero and coefficients labeled "1" are kept intact.
4. Encode the DC subband. This is the first bitstream.
5. Apply the standard fully embedded morphological coder [32] to each of these images. These are the other two bitstreams.

partition. Note the following fundamental fact about this algorithm. In a data compression application, a significant number of bits is required only to resolve "unexpected" events. Although the amount of data to encode is apparently doubled by this algorithm, the data introduced to decouple image regions is highly redundant: big clusters of zeros, highly predictable both within and across scales in the wavelet representation. Furthermore the morphological coder, with its ability to efficiently encode arbitrarily shaped sets of coefficients, provides an excellent method for coding the sets resulting from classification of coefficients. It will be shown later that under fairly general conditions, this coding technique results in better rate/distortion performance than explicitly storing shape information. For decoding, given a total bitrate available, and given how much of it should be spent on each object, this algorithm simply executes the embedded morphological decoder until the bitrate is exhausted on each object.

4. Experimental Results

In this section, we provide extensive simulation results, to show that the penalty paid for providing region-based coding is negligible in terms of coding performance, compared to the performance achievable when no regions are supported. C source code that implements each of the proposed coders can be obtained at `http://www.ifp.uiuc.edu/~servetto/research/`. The example images used in this article are used with permission from the Fowler Museum of Cultural History at the University of California–Los Angeles. These images were part of an image database delivery project called the Museum Educational Site Licensing Project (MESL), sponsored by the Getty Information Institute, containing 286 images. The goal of the MESL project was to test the licensing and delivery of digital images and meta data from seven U. S. museums to seven U. S. universities. The University of Illinois was

(a) (b)

Fig. 5. A typical image in our database: (a) an original image; (b) the regions resulting from identifying one of the objects present in the original.

Table 3. Rate/Distortion performance comparison for the proposed algorithms.

Rate (bpp) (# of bits)	0.0625 (10208)	0.125 (20416)	0.25 (40832)	0.50 (81664)	1.00 (163328)
Standard Morph	23.29	25.65	28.30	32.03	37.60
Regions, No Side Inf.	23.05	25.39	27.98	31.72	37.26
Regions, Explicit Side Inf.	8.80	22.87	26.95	31.14	36.89
JPEG	N/A	N/A	25.67	29.69	34.69

selected as a participant in the MESL project. A typical image in the database is shown in Fig. 5(a), with the segmentation shown in Fig. 5(b).

We present the results of an experiment testing both coders considered in Section 3. The experiment consists of measuring the maximum PSNR value that can be achieved at different bitrates for the bitstream supporting region decoding, and compare that number against the PSNR value achieved by the morphological coder at equal bitrates. Note that because each image region may receive a different number of bits, there will be one distribution of bits among regions that will result in maximum PSNR of the global reconstruction. This maximum PSNR number is the one compared against the no segmentation case, to determine how performance degrades due to providing region-based coding.

The wavelet used is the 10-18 biorthogonal wavelet of Daubechies. The image size is 352x464, and the wavelet tree is grown to depth 4. In all cases, the performance of the standard JPEG algorithm is reported (without any form of region support), since this is the image representation previously used by MARS. The JPEG coder used is the one that comes with the standard XV utility, version 3.00. Typical coding results are shown

Fig. 6. Sample reconstructions of the still image coder, case 2: (a) the original image (1306624 bits, 8bpp), compressed by factors of (b) 16:1 (81664 bits, 0.5bpp), and (c) 47:1 (27766 bits, 0.17bpp). For a perceptual quality comparison, (d) shows a JPEG reconstructed image at 0.17bpp.

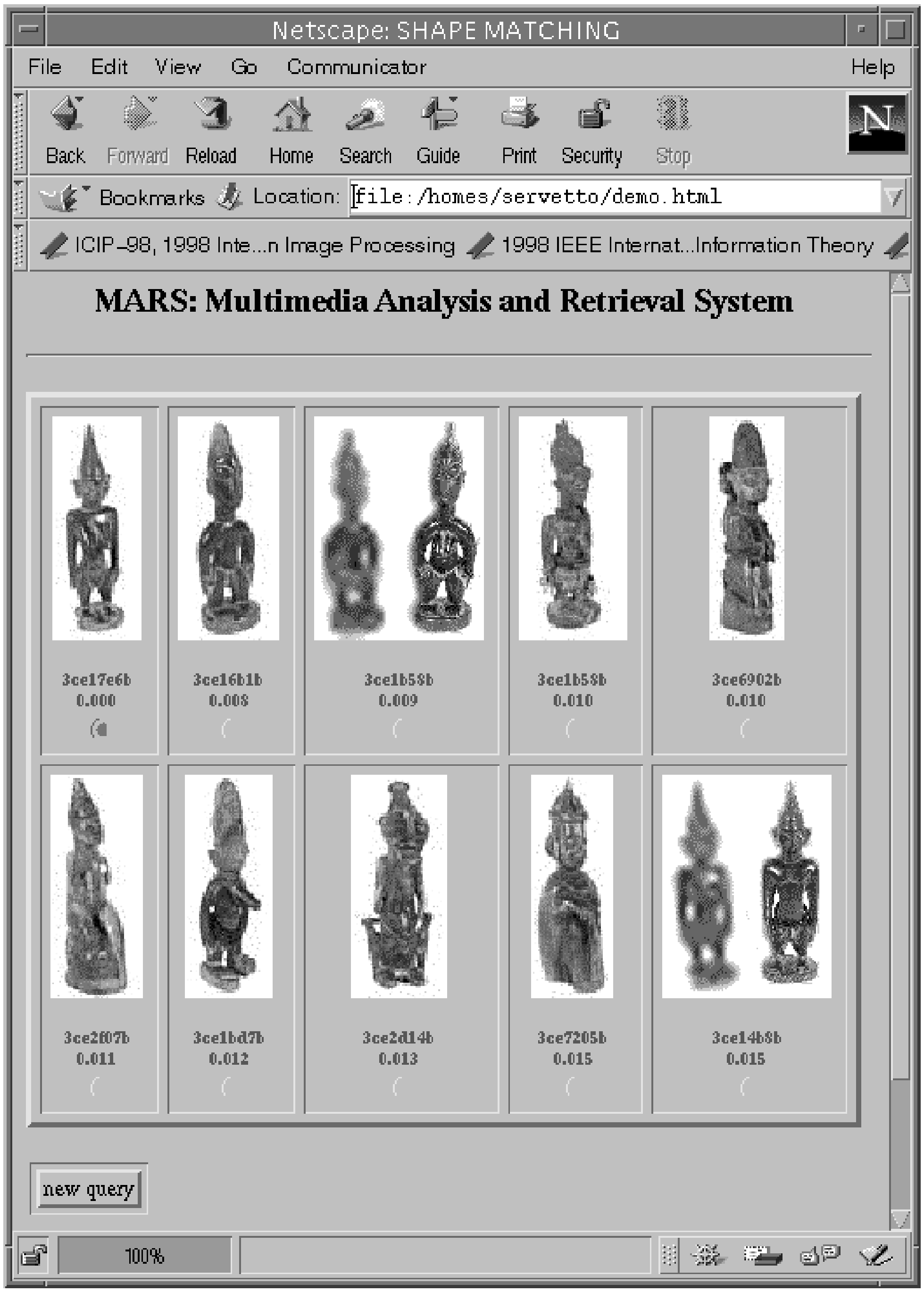


Fig. 7. A typical result of evaluating a query.

in Table 3, for the image and binary segmentation shown in Fig. 5.

Note that when the segmentation map is explicitly included in the bitstream as side information, the resulting coding performance degrades significantly at low rates. This occurs because the number of bits devoted to side information remains fixed, it does not scale with the total bit budget appropriately. As a result, at low rates there are simply not enough bits to encode the data, all of them go to encode side information. Based on these observations, we conclude that the case of no side information renders the best coding performance always, deviating approx. 0.25-0.35dB from the performance of the standard morph coder for this image. In a perceptual sense, the images are indistinguishable, as can be seen in the sample reconstructions shown in Fig. 6.

To illustrate how our new proposed image representation is integrated into MARS, in Fig. 7 we show a typical result of evaluating a query. In a separate window, the user provided the image in the upper-left corner, and specified that was interested in retrieving images with regions of a similar shape. MARS ranked all the images in the database according to a shape similarity measure, and displayed the 10 best matches. Observe how, in the case of images containing more than one significant region, only the region of interest to the user is displayed at high resolution. This feature results in significant transmission time savings, by being able to avoid sending regions not of interest to the user.

5. Conclusions

In this paper, we addressed the problem of efficiency in the internal representation of image regions in a multimedia DBMS. We suggest that a desirable property of any image representation for this application is that of supporting retrieval of arbitrarily shaped image regions directly from the compressed bitstream, and we developed new coding methods to support such functionality.

The image coding methods proposed in this paper combine the compression efficiency of state-of-the-art "pure" compression algorithms, with the flexibility of the bitstream syntax required by a DMBS system like MARS, to support region-based processing. This is achieved by combining the right set of compression and database systems tools. The main features of the proposed algorithms are (a) that compared to when only support for rectangular images is provided, they have a negligible rate/distortion performance degradation; and (b) that they produce perceptually indistinguishable images from those obtained by equivalent coders that do not provide support for regions.

References

1. MPEG-7 applications document. *ISO/IEC JTC1/SC29/WG11 N1922, MPEG97*, Oct 1997.
2. MPEG-7: Context and objectives (v.5). *ISO/IEC JTC1/SC29/WG11 N1920, MPEG97*, Oct 1997.
3. Third draft of MPEG-7 requirements. *ISO/IEC JTC1/SC29/WG11 N1921, MPEG97*, Oct 1997.
4. N. Ahmed, T. Natarajan, and K. R. Rao. Discrete Cosine Transform. *IEEE Transactions on Computers*, 23:88–93, January 1974.
5. Jeffrey R. Bach, Charles Fuller, Amarnath Gupta, Arun Hampapur, Bradley Horowitz, Rich Humphrey, Ramesh Jain, and Chiao fe Shu. The Virage image search engine: An open framework for image management. In *Proc. SPIE Storage and Retrieval for Image and Video Databases*.
6. Chris Buckley and Gerard Salton. Optimization of relevance feedback weights. In *Proc. of SIGIR'95*, 1995.
7. James Dowe. Content-based retrieval in multimedia imaging. In *Proc. SPIE Storage and Retrieval for Image and Video Databases*, 1993.
8. W. Equitz and T. Cover. Successive Refinement of Information. *IEEE Transactions on Information Theory*, 37(2):269–275, March 1991.
9. C. Faloutsos, M. Flickner, W. Niblack, D. Petkovic, W. Equitz, and R. Barber. Efficient and effective querying by image content. Technical report, IBM Research Report, 1993.
10. M. Flickner, Harpreet Sawhney, Wayne Niblack, Jonathan Ashley, Q. Huang, Byron Dom, Monika Gorkani, Jim Hafine, Denis Lee, Dragutin Petkovic, David Steele, and Peter Yanker. Query by image and video content: The QBIC system. *IEEE Computer*, 1995.
11. T. Gevers and V. K. Kajcovski. Image segmentation by directed region subdivision. In *Proc. IEEE Int. Conf. on Image Proc.*, 1994.
12. Michael Hansen and William Higgins. Watershed-driven relaxation labeling for image segmentation. In *Proc. IEEE Int. Conf. on Image Proc.*, 1994.
13. Thomas S. Huang, Sharad Mehrotra, and Kannan 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.
14. R. Jain. Workshop report: NSF workshop on visual information management systems. In *Proc.*

SPIE Storage and Retrieval for Image and Video Databases, 1993.
15. R. Jain, A. Pentland, and D. Petkovic. NSF-ARPA workshop on visual information management systems. Cambridge, MA, June 1995.
16. W. Y. Ma and B. S. Manjunath. Edge flow: a framework of boundary detection and image segmentation. In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, 1997.
17. W. Y. Ma and B. S. Manjunath. Netra: A toolbox for navigating large image databases. In *Proc. IEEE Int. Conf. on Image Proc.*, 1997.
18. M. K. Mandal, T Aboulnasr, and S. Panchanathan. Image indexing using moments and wavelets. *IEEE Transactions on Consumer Electronics*, 42(3):557–565, Aug 1996.
19. P. Maragos and R. Schafer. Morphological Systems for Multidimensional Signal Processing. *Proceedings of the IEEE*, 78(4):690–710, April 1990.
20. Sharad Mehrotra, Kaushik Chakrabarti, Mike Ortega, Yong Rui, and Thomas S. Huang. Multimedia analysis and retrieval system. In *Proc. of The 3rd Int. Workshop on Information Retrieval Systems*, 1997.
21. W. Niblack, R. Barber, and et al. The QBIC project: Querying images by content using color, texture and shape. In *Proc. SPIE Storage and Retrieval for Image and Video Databases*, Feb 1994.
22. Michael Ortega, Yong Rui, Kaushik Chakrabarti, Sharad Mehrotra, and Thomas S. Huang. Supporting similarity queries in MARS. In *Proc. of ACM Conf. on Multimedia*, 1997.
23. A. Pentland, R. W. Picard, and S. Sclaroff. Photobook: Content-based manipulation of image databases. *International Journal of Computer Vision*, 1996.
24. K. Ramchandran and M. Orchard. An Investigation of Wavelet-Based Image Coding Using an Entropy-Constrained Quantization Framework. *IEEE Transactions on Signal Processing*, 46(2):342–353, February 1998.
25. Yong Rui, Kaushik Chakrabarti, Sharad Mehrotra, Yunxin Zhao, and Thomas S. Huang. Dynamic clustering for optimal retrieval in high dimensional multimedia databases. In *TR-MARS-10-97*, 1997.
26. Yong Rui, Thomas S. Huang, and Sharad Mehrotra. Relevance feedback techniques in interactive content-based image retrieval. In *Proc. of IS&T SPIE Storage and Retrieval of Images/Video Databases VI, EI'98*, 1998.
27. Yong Rui, Thomas S. Huang, Sharad Mehrotra, and Michael Ortega. Automatic matching tool selection using relevance feedback in MARS. In *Proc. of 2nd Int. Conf. on Visual Information Systems*, 1997.
28. Yong Rui, Thomas S. Huang, Sharad Mehrotra, and Michael Ortega. A relevance feedback architecture in content-based multimedia information retrieval systems. In *Proc of IEEE Workshop on Content-based Access of Image and Video Libraries, in conjunction with IEEE CVPR '97*, 1997.
29. Yong Rui, Alfred C. She, and Thomas S. Huang. Automated shape segmentation using attraction-based grouping in spatial-color-texture space. In *Proc. IEEE Int. Conf. on Image Proc.*, 1996.
30. A. Said and W. Pearlman. A New, Fast, and Efficient Image Codec Based on Set Partitioning in Hierarchical Trees. *IEEE Transactions on Circuits and Systems for Video Technology*, 6(3):243–250, June 1996.
31. Gerard Salton and Michael J. McGill. *Introduction to Modern Information Retrieval*. McGraw-Hill Book Company, 1983.
32. S. Servetto, K. Ramchandran, and M. Orchard. Image Coding based on a Morphological Representation of Wavelet Data. *Submitted to the IEEE Transactions on Image Processing*, December 1996.
33. J. Shapiro. Embedded Image Coding Using Zerotrees of Wavelet Coefficients. *IEEE Transactions on Signal Processing*, 41(12):3445–3462, December 1993.
34. Alfred C. She and Thomas S. Huang. Segmentation of road scenes using color and fractal-based texture classification. In *Proc ICIP*, Austin, Nov. 1994.
35. J. R. Smith and S.-F. Chang. Visually searching the web for content. *IEEE Multimedia Magazine*, 4(3):12–20, Summer 1997. also Columbia U. CU/CTR Technical Report 459-96-25.
36. Michael Swain and Dana Ballard. Color indexing. *International Journal of Computer Vision*, 7(1), 1991.
37. D. Taubman and A. Zakhor. Multirate 3-D Subband Coding of Video. *IEEE Transactions on Image Processing*, 3(5):572–588, September 1994.
38. M. Vetterli and J. Kovačević. *Wavelets and Subband Coding*. Prentice Hall, Englewoods Cliffs, NJ, 1st edition, 1995.
39. Z. Xiong, K. Ramchandran, and M. Orchard. Space-Frequency Quantization for Wavelet Image Coding. *IEEE Transactions on Image Processing*, 6(5):677–693, May 1997.
40. X. Yang and K. Ramchandran. Hierarchical Backward Motion Compensation for Wavelet Video Coding Using Optimized Interpolation Filters. In *Proceedings of the IEEE International Conference on Image Processing*, Santa Barbara, CA, 1997.

Sergio D. Servetto received the Licenciatura en Informática from Universidad Nacional de La Plata (UNLP), Argentina, in 1992. He is currently working towards the Ph.D. degree in Computer Science, at the University of Illinois at Urbana-Champaign (UIUC).

Between 1991 and 1994 he was a permanent Member of the Technical Staff at the Computer Research and Advanced Applications Group (CRAAG) of IBM Argentina. Since 1994 he has been a Graduate Research Assistant at the Image Formation and Processing Group of the Beckman Institute at UIUC. During 1992 he held undergraduate and during 1993 graduate teaching positions in the Department of Computer Science at UNLP, and he was a Teaching Assistant

for two semesters during 1996 and 1997 in the Department of Electrical and Computer Engineering at UIUC. During the summer of 1997 he was a summer intern at the Multimedia Communications Research Department of Bell Laboratories, in Murray Hill, NJ. During the summer of 1998 he will be a summer intern at the Information Sciences Research Department of AT&T Labs, in Florham Park, NJ.

His research interests include image/video modeling and compression problems, the problem of robust transmission of compressed image/video data over unreliable channels, and the algorithmic and computational complexity aspects of problems in communications and coding theory.

Yong Rui received the B.S. degree from the Southeast University, P. R. China in 1991 and the M.S. degree from Tsinghua University, P. R. China in 1994, both in Electrical Engineering. He is currently finishing the Ph.D. degree in Electrical and Computer Engineering at the University of Illinois at Urbana-Champaign. Since 1995 he has been a Graduate Research Assistant at the Image Formation and Processing Group of the Beckman Institute for Advance Science and Technology at UIUC. During the summer of 1998 he is a summer intern at the Vision Technology Group of Microsoft Research, in Redmond, WA.

His research interests include multimedia information retrieval, multimedia signal processing, computer vision and artificial intelligence. He has published over 20 technical papers in the above areas.

He is a Huitong University Fellowship recipient 1989-1990, a Guanghua University Fellowship recipient 1992-1993, and a CSE Engineering College Fellowship recipient 1996-1998.

Kannan Ramchandran received his B.S. degree from the City College of New York, and his M.S. and Ph.D. from Columbia University, all in Electrical Engineering in 1982, 1984, and 1993 respectively. From 1984 to 1990, he was a Member of the Technical Staff at AT&T Bell Labs in the telecommunications R&D area in optical fiber carrier systems. From 1990 to 1993, he was a Graduate Research Assistant at the Center for Telecommunications Research at Columbia University. Since 1993, he has been with the University of Illinois at Urbana-Champaign, where he is currently an Assistant Professor in the Electrical and Computer Engineering Department, and a Research Assistant Professor at the Beckman Institute and the Coordinated Science Laboratory at Illinois.

Dr. Ramchandran was the recipient of the 1993 Elaihu I. Jury Award at Columbia University for the best doctoral thesis in the area of systems, signal processing, or communications. He received an NSF Research Initiation Award in 1994, an Army Research Office Young Investigator Award in 1996, an NSF CAREER Award in 1997, and an Office of Naval Research Young Investigator Award in 1997. He received the 1996 Senior Best Paper award from the IEEE Signal Processing Society for a paper with Martin Vetterli that appeared in the IEEE Transactions on Image Processing in April 1993. He is a member of the IEEE IMDSP technical committee, and serves as an Associate Editor for the IEEE Transactions on Image Processing.

His research interests include image and video compression, multirate signal and image processing and wavelets, and joint source-channel coding for robust image and video communications over noisy channels and heterogeneous networks.

Thomas S. Huang received his B.S. Degree in Electrical Engineering from National Taiwan University, Taipei, Taiwan, China; and his M.S. and Sc.D. Degrees in Electrical Engineering from the Massachusetts Institute of Technology, Cambridge, Massachusetts. He was on the Faculty of the Department of Electrical Engineering at MIT from 1963 to 1973; and on the Faculty of the School of Electrical Engineering and Director of its Laboratory for Information and Signal Processing at Purdue University from 1973 to 1980. In 1980, he joined the University of Illinois at Urbana-Champaign, where he is now William L. Everitt Distinguished Professor of Electrical and Computer Engineering, and Research Professor at the Coordinated Science Laboratory, and Head of the Image Formation and Processing Group at the Beckman Institute for Advanced Science and Technology.

During his sabbatical leaves, Dr. Huang has worked at the MIT Lincoln Laboratory, the IBM Thomas J. Watson Research Center, and the Rheinishes Landes

Museum in Bonn, West Germany, and held visiting Professor positions at the Swiss Institutes of Technology in Zurich and Lausanne, University of Hannover in West Germany, INRS-Telecommunications of the University of Quebec in Montreal, Canada and University of Tokyo, Japan. He has served as a consultant to numerous industrial firms and government agencies both in the U.S. and abroad.

Dr. Huang's professional interests lie in the broad area of information technology, especially the transmission and processing of multidimensional signals. He has published 12 books, and over 300 papers in Network Theory, Digital Filtering, Image Processing, and Computer Vision. He is a Fellow of the International Association of Pattern Recognition, IEEE, and the Optical Society of American; and has received a Guggenheim Fellowship, an A.V. Humboldt Foundation Senior U.S. Scientist Award, and a Fellowship from the Japan Association for the Promotion of Science. He received the IEEE Acoustics, Speech, and Signal Processing Society's Technical Achievement Award in 1987, and the Society Award in 1991. He is a Founding Editor of the International Journal Computer Vision, Graphics, and Image Processing; and Editor of the Springer Series in Information Sciences, published by Springer Verlag.