

# Multilevel Wavelet Feature Statistics for Efficient Retrieval, Transmission, and Display of Medical Images by Hybrid Encoding

**Shuyu Yang**

*Department of Electrical and Computer Engineering, Texas Tech University, Lubbock, TX 79409-3102, USA  
Email: shu.yang@ttu.edu*

**Sunanda Mitra**

*Department of Electrical and Computer Engineering, Texas Tech University, Lubbock, TX 79409-3102, USA  
Email: sunanda.mitra@coe.ttu.edu*

**Enrique Corona**

*Department of Electrical and Computer Engineering, Texas Tech University, Lubbock, TX 79409-3102, USA  
Email: ecorona@ttacs.ttu.edu*

**Brian Nutter**

*Department of Electrical and Computer Engineering, Texas Tech University, Lubbock, TX 79409-3102, USA  
Email: brian.nutter@coe.ttu.edu*

**D. J. Lee**

*Department of Electrical and Computer Engineering, Brigham Young University, Provo, UT 84602, USA  
Email: djlee@ee.byu.edu*

*Received 31 March 2002 and in revised form 25 October 2002*

Many common modalities of medical images acquire high-resolution and multispectral images, which are subsequently processed, visualized, and transmitted by subsampling. These subsampled images compromise resolution for processing ability, thus risking loss of significant diagnostic information. A hybrid multiresolution vector quantizer (HMQV) has been developed exploiting the statistical characteristics of the features in a multiresolution wavelet-transformed domain. The global codebook generated by HMQV, using a combination of multiresolution vector quantization and residual scalar encoding, retains edge information better and avoids significant blurring observed in reconstructed medical images by other well-known encoding schemes at low bit rates. Two specific image modalities, namely, X-ray radiographic and magnetic resonance imaging (MRI), have been considered as examples. The ability of HMQV in reconstructing high-fidelity images at low bit rates makes it particularly desirable for medical image encoding and fast transmission of 3D medical images generated from multiview stereo pairs for visual communications.

**Keywords and phrases:** high fidelity hybrid encoding, global codebook, low bit rate, multilevel wavelet feature statistics, efficient retrieval of high-resolution medical images.

## 1. INTRODUCTION

Large volumes of digitized radiographic images accumulated in hospitals and educational institutes pose a challenge in image database management, requiring high fidelity and image modality-specific compression approaches. Such level of image management necessitates a system that provides easy access and high fidelity reconstruction. The use of image compression for fast medical image retrieval is a debatable subject

since high compression ratios usually introduce critical information loss that might impede accurate diagnosis. However, requirements for image quality also differ depending on applications. It is therefore desirable to construct a flexible image management system that can cater to the specific needs of its users. The system should address important issues such as user-preferred image resolution and scale and transmission time and method (progressive or nonprogressive transmission), as well as possess a user friendly interface.

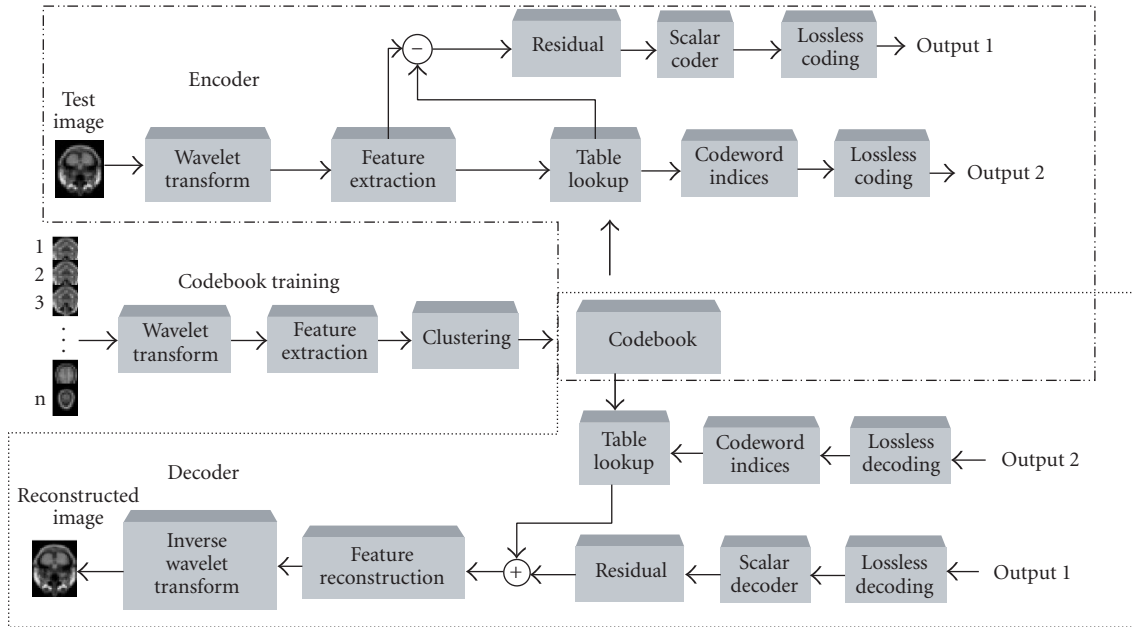


FIGURE 1: A block diagram of the HMVQ coding scheme.

Such system's applications are broad in nature and include telemedicine, video conferencing, and distance education, to name a few [1, 2].

Content-based retrieval of specific images from large image databases is a challenging research area relevant to many types of image archives encountered in medical, remote sensing, and hyperspectral imagery. In general, image features must be extracted to facilitate indexing and content-based retrieval procedures. When multiscale vectors are used for codebook training using the Euclidean distance as a distortion measure, distortions from each coefficient of the vector are equally weighed, thus, the contribution to the distortion depends on the coefficients themselves instead of their orders. This principle has been proven successful in scalar coding methods such as the embedded zerotree wavelet (EZW) coding [3] and the set partitioning in hierarchical trees (SPIHT) [4]. In EZW and SPIHT, many bits have to be used in distinguishing significant coefficients and coding their locations. The use of multiscale vectors [5, 6, 7, 8, 9] can further improve performance by saving valuable bits used in coding the locations of important coefficients since the location information has already been embedded in the vectors and their order.

Traditionally, vectors are generated by grouping neighboring wavelet coefficients within the same subband and orientation; square blocks are usually used for this purpose. The size of the block (i.e., vector dimension) is usually chosen randomly or as a result of bit-allocation optimization. The resulting multiresolution codebooks [10] fail to form efficient global codebooks for large medical image data sets. The hybrid multiscale vector quantization (HMVQ) scheme described in this paper, on the other hand, generates multidimensional vectors across multiresolution levels,

thus eliminating the problem of building codebooks for all subimages at each level. In addition, analysis of the magnitude distribution of the multiscale vectors has led to the novel scheme of HMVQ, having an embedded residual scalar quantization within the global codebook. Preliminary results of HMVQ have been presented in [7, 8, 9], showing excellent performance for good quality reconstruction of natural and medical images. However, a codebook designed for a specific application is desirable to obtain high fidelity image reconstruction at low bit rates. This paper presents the analysis and criteria of designing such codebooks (HMVQ) in detail with a novel wavelet feature statistics-based hybrid encoding, including vector quantization and residual scalar encoding. Results obtained from three specific 2D medical image data sets are included with discussions on the advantages of HMVQ in encoding and fast transmission of 3D medical images.

We have organized this paper by stating the necessity of designing low bit rate yet high fidelity encoder/decoder for efficient archiving and transmission of large medical image data sets in Section 1. Section 2 presents a detailed description of analysis and design of HMVQ. Section 3 presents the preliminary results of high fidelity reconstruction of two different image modalities. Section 4 addresses the advantages of extending HMVQ to encoding 3D images generated from stereo pairs. Section 5 discusses future research and conclusions.

## 2. ANALYSES AND DESIGN OF HMVQ

Figure 1 shows the complete block diagram of the HMVQ-based encoder/decoder system. The image in the spatial domain is first transformed into the wavelet domain to remove

the statistical redundancy among image pixels. Codebooks designed in the transform domain are believed to be closer to optimal than those designed in the spatial domain, because the transformed coefficients have better defined distributions than image pixel distributions [10, 11].

### 2.1. Multiscale feature extraction

Traditionally, vectors in the wavelet domain are generated by grouping neighboring wavelet coefficients within the same subband and orientation in the same way as in the spatial domain. Vector dimensions vary and depend on the outcome of the adopted bit allocation scheme. For example, in [10, 11], bit allocation is obtained based on rate distortion optimization as a function of subband and orientation. The total distortion rate function  $D_T(R_T)$  is given by

$$D_T(R_T) = \frac{1}{2^{2M}} D_M^{SQ}(R_M^{SQ}) + \sum_{m=1}^M \frac{1}{2^{2m}} \sum_{d=1}^3 D_{m,d}(R_{m,d}), \quad (1)$$

where  $D_M^{SQ}(R_M^{SQ})$  represents the subimage of the lowest resolution,  $D_{m,d}(R_{m,d})$  represents the average distortion resulting from encoding the subimage  $(m, d)$  at  $(R_{m,d})$  bits per pixel,  $M$  is the total number of scale, and  $d$  represents three orientations. The total distortion rate function  $D_T(R_T)$  is minimized subject to the total rate  $R_T$ , where  $R_T$  is defined as

$$R_T = \frac{1}{2^{2M}} R_M^{SQ} + \sum_{m=1}^M \frac{1}{2^{2m}} \sum_{d=1}^3 R_{m,d}. \quad (2)$$

The optimized rate at a certain scale  $m$  and orientation  $d$  is then given by

$$\begin{aligned} R_{m,d_{\text{opt}}} &= \frac{4^M R_T - R_M^{SQ}}{4^M - 1} \\ &+ \frac{1}{r} \log_2 \left[ \frac{C_{m,d}(k, r)}{\left[ \prod_{m'=1}^M \left[ \prod_{d'=1}^3 [C_{m',d'}(k, r)]^{1/4m'} \right]^{4^M} \right]^{4^M/4^M - 1}} \right]. \end{aligned} \quad (3)$$

Generally, when Euclidean distance is used as the distortion measure,  $r = 2$ . Then the lower bound is defined by the coefficient  $c(k, 2)$  of vector dimension  $k$ , and is given by

$$c(k, 2) \geq \frac{1}{(k+2)\pi} \Gamma\left(1 + \frac{k}{2}\right), \quad (4)$$

where  $\Gamma(x)$  is the Gamma function.

As a result, this vector extraction method produces vectors of different dimensions at different scales and orientations. Consequently, multiresolution codebooks, which consist of subcodebooks of different dimensions and sizes, are needed. Although the use of subcodebooks makes the vector-codeword matching process faster, the resulting vector dimension and codebook size become image-size dependent.

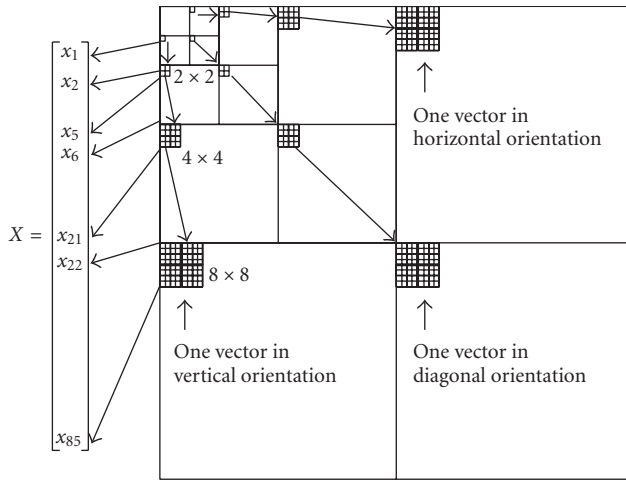
Therefore, the latter type of vector extraction methods is difficult to use for training and generating universal codebooks.

On the other hand, motivated by the success of the hierarchical scalar encoding of wavelet transform coefficients, such as the EZW algorithm and SPIHT, several attempts have been made to adopt a similar methodology to discard insignificant vectors (or zerotrees) as a preprocessing step before the actual vector quantization is performed, using traditional vector extraction methods. In [12], the set-partitioning approach in SPIHT is used to partially order the vectors of wavelet coefficients by their vector magnitudes, followed by a multistage or tree-structured vector quantization for successive refinement. In [13], 21-dimensional vectors are generated by cascading vectors from lower scale to higher scales in the same orientation in a 3-level wavelet transform. Coefficients 1, 4, and 16 from the 3rd, 2nd, and 1st level bands of the same orientation are sequenced to form the desired vectors. If the magnitudes of all the elements of such a vector are less than a threshold, the vector is considered to be a zerotree and not coded. After all zerotrees are designated, the remaining coefficients are reorganized into lower-dimensional vectors, and then vector quantized.

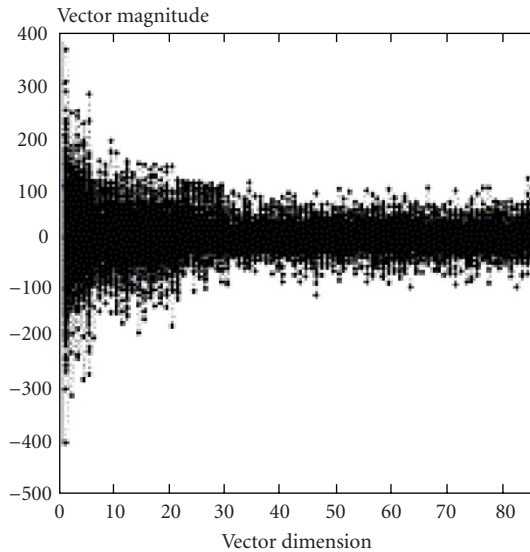
Our approach of vector extraction resembles only the first stage of generating vectors similar to [13] but quite different in the way it is organized as explained below. Firstly, instead of using the multiscale vectors just for insignificant coefficient rejection, we use the entire multiscale vectors as sample vectors for codebook training. Secondly, the dimension of the vector is not limited to 21. Depending on the level of wavelet transform and the complexity of the quantizer, it can be varied. Our new way of forming sample vectors takes both dependencies into consideration. Vectors are formed by stacking blocks of wavelet coefficients at different scales at the same orientation location. Since the scale size decreases as the decomposition level goes up, block size at lower level is twice the size of that of its adjacent higher level. The same procedure is used to extract feature vectors for all three orientations. The dimension of the vector is fixed once the decomposition level is chosen.

In our approach, multiscale feature vectors are extracted from the wavelet coefficients such that both interscale and intrascale redundancy can be exploited in vector quantization. Figure 2a illustrates how an 85-dimensional vector is extracted from a 4-level wavelet transformed image. Coefficients 1, 4, 16, and 64 from the fourth, third, second, and first level subbands of the same orientation are sequenced. The use of multiscale vectors for vector quantization has several advantages over the use of vectors formed from traditional rectangular blocks. The new multiscale vectors are image-size independent, retain image features, and exploit intra- and interscale redundancy, and the resulting codebook is scalable (i.e., higher-dimensional codebooks contain all codewords for lower-dimensional ones).

The major advantage of using such multiscale vector generation scheme is that we are able to capture image features from the coarser version to finer version within one vector, thus making it image-size independent. This common feature is illustrated in Figure 2b, where a number of vectors



(a)



(b)

FIGURE 2: (a) An example of multiscale vector extraction. (b) Distribution of multiscale vector magnitudes.

from different images are plotted together to illustrate the relationship between vector magnitudes with vector dimensions. Thus, when vectors are trained into a codebook, the codebook incorporates both image features and wavelet coefficient properties. In addition, both intrascale and interscale redundancy among wavelet coefficients can be efficiently exploited since the vector contains coefficients inside the subbands and across the subbands. Based on the same principle, human perceptual models can be embedded into the optimization process [14].

## 2.2. HMQ including residual scalar encoding [8, 9]

### Residual encoding

All vector quantization schemes result in somewhat blurring in the reconstructed image, especially when the codebook

size is reduced to meet practical processing speed and storage requirements. Detail features such as edges can be lost, particularly, at low bit rates. It is therefore desirable to find an approach to compensate for the lost details. To accomplish such a goal, a second-step residual scalar coding is used in our approach after the vector quantization of the multiscale vectors. The residual represents the details lost during vector quantization. Because multiscale vectors preserve the scale structure of the wavelet coefficients, zerotree-based coding algorithms such as EZW and SPIHT can be used for residual coding. When the codebook is well designed, the residual contains only a small number of large magnitude elements. Therefore, only a few large magnitude elements have to be coded, saving a large number of bits.

### Possibility of generating universal codebooks

If any image information can be described by a common distribution and a clustering algorithm that achieves the global minimum for this type of distribution is used to design a codebook, such a codebook can be referred to as a universal codebook [11, 15]. When a simple coding scheme, such as the one described in [16], is used, a universal codebook for all types of images is difficult to generate. The problem of generating a universal codebook can be addressed in two ways. Firstly, regardless of the source characteristics, an efficient codebook generation algorithm must be used to produce global codebooks with reasonable computational complexity. Roughly speaking, there are two most popular techniques for codebook generation. One way is to use pattern recognition techniques to generate codebooks with a large amount of training data and seek a minimum distortion codebook for the data [17, 18]. By using training data sets, the codebook can be optimized for the data type. Clustering algorithms are usually used for codebook training. However, well-structured lattice codebooks have also been designed [19], in which the centroids are predefined once the type of lattice is selected. Secondly, the ability to characterize image information by a common distribution is needed. Since it is obvious that this cannot be accomplished in the spatial domain, image coefficients in the transformed domain should be considered. However, for vector quantization, we are seeking an approach that can use a limited number of vectors to represent the vast variety of image features as shown in Figure 2b.

### Vector quantization in the wavelet domain

It has already been demonstrated that image wavelet coefficients possess the most valuable property of having a distribution similar to a generalized Gaussian distribution [10, 11] for every subband. If the coefficients are adequately decorrelated such that the vectors extracted from the coefficients can be approximated as i.i.d. generalized Gaussian distributed, then the gain in reduction of distortion by vector quantization is higher than Gaussian and uniform sources. Because of such predictable coefficient distributions and theoretically high distortion reduction, image vector quantization in the wavelet domain is believed to be able to achieve a better



performance than in other domains and can be a starting ground for building a universal codebook.

However, the choice of clustering algorithm has a significant effect on codebook generation by vector quantization. The LBG algorithm [20], ever since it came to existence in 1980, it has been the most popularly used clustering algorithm for vector quantization codebook training because of its simplicity and adequate performance. However, its shortcoming of being easily trapped in local minima is also well known. The recently developed deterministic annealing (DA) [21] algorithm is believed to reach the global minimum despite lacking theoretical support. Our investigation of LBG, DA, and AFLC [22] reveals various difficulties and advantages associated with each of them in their application to vector quantization [7, 23]. We came to the conclusion that when the source distribution is symmetric and rotationally invariant around the origin, DA comes closer to the global optimum than the other two. Otherwise, LBG gives the most consistent performance. Fortunately, we can observe that wavelet coefficients are approximately symmetric and rotationally invariant to the origin, thus, DA is the best choice for accurate codebook training. However, DA is also computational intensive. Therefore, algorithm selection is a compromise that depends on available resources.

### 3. RESULTS

The performance of HMQV was tested with two different medical image modalities, MRI and X-ray radiographic data. Separate codebooks were formed for each modality to have high fidelity reconstruction at low bit rate by keeping the codebook size small.

#### 3.1. MRI data

The first set of training data we used is a group of slices (slice 1 to slice 31) from a 3D simulated MR image of a human brain <http://www.bic.mni.mcgill.ca/brainweb>. This set of images is an MR simulation of  $T_1$ -weighted, zero noise level, zero intensity nonuniformity, 1-mm thick, and 8 bits per pixel (bpp) normal human brain with voxels of  $181 \times 217 \times 181$  ( $X \times Y \times Z$ ) when it is at a 1-mm isotropic voxel grid in Talairach space. Thus, the training images are reasonably different because of the span from top of the brain to the lower part of the brain despite belonging to the same class.

Figure 3 shows some of the images from the training set. A few slices inside the group, for example, slice 6, slice 12, and so forth, are randomly chosen and excluded from the training set and later used as test images. A codebook of size 256 is used. Reconstructed images comparing the HMQV and SPIHT are shown in Figure 4. The results show that HMQV preserves more detail information than SPIHT. This is more evident in Figure 8 where Canny edge detection operation has been performed on Figure 4b and Figure 4e. Numerical comparison on peak signal-to-noise ratio (PSNR) versus bit rate (PSNR(R)) is summarized in Figure 7.

#### 3.2. X-ray radiographic data

When the targeted images belong to the same category, a special codebook can be generated to improve the performance of HMQV. To obtain a codebook of reasonable size, a training set must be selected. Two training sets were chosen from the cervical and lumbar spine X-ray images collected by NHANES II [24, 25]. The original images were 12 bpp with size of 2487 by 2048. To aid processing, the images are converted to 8 bpp. For experimental purposes, parts of the images that contained important information were cropped, resulting in training images of size 2048 by 1024. A codebook containing 256 multiscale codewords is generated for lumbar image encoding. Similarly, another codebook is obtained for the cervical spine images, which are also 8 bpp 1024 by 1024 gray scale images. The test images, which are outside the training set, are used to demonstrate the quality of the reconstructed images at different bit rates. Figure 5 presents the lumbar and cervical spine test images, all displayed at a ratio of 1 to 256 of their original sizes. Because it is not practical to show the reconstructed images in their original sizes here, a region of interest in the spine area is shown in Figure 6, with an edge detection comparison in Figure 8. Here, better edge preservation of HMQV codec over SPIHT codec can be clearly observed. The overall PSNR versus bit rate performance of the HMQV codec is compared to that of SPIHT in Figure 7a for lumbar images and Figure 7b for cervical spine images.

#### Quantitative evaluation of HMQV performance

The effectiveness of HMQV in terms of quantitative measures such as the PSNR is demonstrated for medical as well as standard images in Figure 8. For standard images, 85-dimensional vectors from a set of 28 images, most of which are from the USC standard image database and some are taken from the author's own database, are generated to design a codebook for standard images. A codebook size of 256 is used in this experiment. The well-known Lena (8 bpp), which is outside the training set, is used as the test image [23]. In Figure 7d, PSNR versus bit rate curves resulting from HMQV is compared with that of SPIHT as well as another well-known multiresolution vector quantizer [10]. HMQV outperforms both. In Figure 8, edges detected on sections of the reconstructed cervical spine and Lena images further demonstrate better detail retaining capability of HMQV over SPIHT even at a very low bit rate.

#### 3.3. HMQV in management of 3D medical images

Evaluation of deformation in 3D shape may provide significant diagnostic aid in early detection and follow-up of a disease such as glaucoma by changes observed in the optic disc volume by quantitative measures [26, 27].

Figure 9 shows how such quantitative measures can be obtained from stereoscopic fundus images taken in an ophthalmology clinic by computing the disparity map [26, 27, 28, 29]. However, storage of such 3D images in addition to the stereo pairs of large patient population necessitates the use of a high fidelity encoding scheme. Any 2D

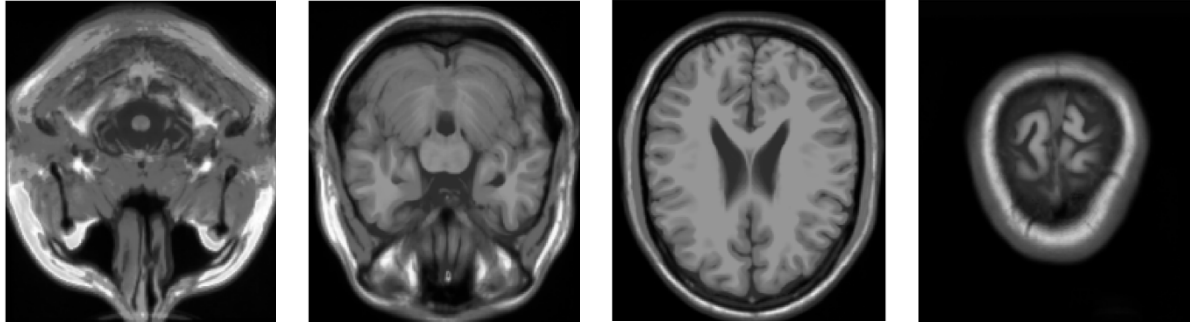


FIGURE 3: Some images from the training set showing widely different contents.

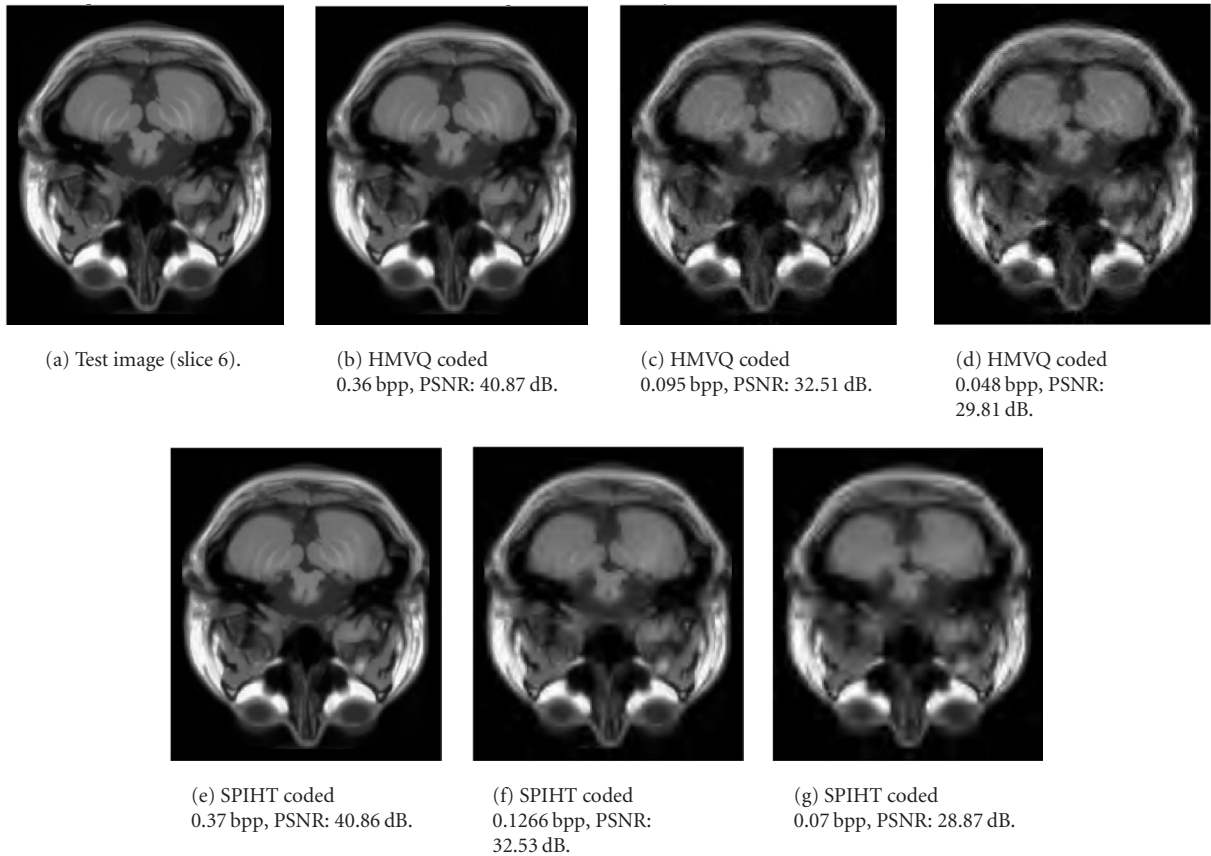


FIGURE 4: Comparison of reconstructed images by HMVQ and SPIHT.

encoding scheme is equally applicable to 3D images by encoding the 2D disparity map in a multiview system capable of 3D rendering [30]. Figure 10 shows a schematic diagram of how HMVQ can be incorporated into a multiview system, thus reducing the bit stream to be transmitted for efficient retrieval of 3D shapes.

#### 4. DISCUSSIONS

The results of applying HMVQ to generate codebooks for different image modalities demonstrate improved performance

of HMVQ over SPIHT in high fidelity reconstruction at low bit rates. We also demonstrate that HMVQ codec gives better PSNR versus bit rate performance (Figure 7) on different types of images over scalar quantizer SPIHT as well as vector quantizer (Figure 7d). Perceptually, reconstructed images from HMVQ also have better detail preservation than those from SPIHT, as shown in Figure 8, where more edges can be detected in HMVQ-reconstructed images than in SPIHT-reconstructed images. We have presented an example where 3D surface of retinal structures can be recovered and displayed from a stereo pair under some constraints.



(a) Lumbar test image.

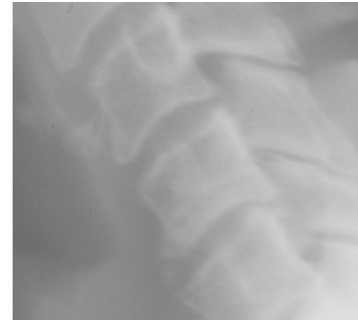


(b) Cervical spine test image.

FIGURE 5: The test images.

However, such a 3D surface recovery is an ill-posed problem and cannot be recovered exactly. Reconstruction and display of natural scenes involve intensive computation to process multiview data necessary to avoid occlusion and pose tremendous difficulty for on-chip processing and efficient communications networking [31]. High-fidelity novel encoding techniques are, therefore, essential to reduce computational cost and overall processing time [1].

Another example of such medical image management application is the digitally archived 17,000 cervical and lumbar spine images at the National Library of Medicine [24]. These images were collected in the second National Health and Nutrition Examination Survey (NHANES II), and they contain instances of both normal and abnormal spine features of interest to researchers in osteoarthritis. These images are currently accessible to the public by the Web-based Medical Information Retrieval System (WebMIRS) [25], in a spatial resolution reduced by a factor of 4 both horizontally and vertically. This simple subsampling method has the significant disadvantage of degrading visual quality considerably. Alternative methods using lossy compression such as vector quantization [32, 33] are known to have improved SNR and can potentially override this loss of visual quality while simultaneously decreasing the file size. However, developing global codebook for large databases is an extremely difficult task and no such codebook is available currently. Preliminary results of the performance of a proposed system using HMVQ for content-based retrieval and high-fidelity reconstruction for both lumbar and cervical X-ray images from this large database have been presented recently [8].



(a) A section of cervical spine from the original test image.



(b) HMVQ reconstructed image section. Bit rate: 0.024 bpp and PSNR: 44.57.



(c) SPIHT reconstructed image section. Bit rate: 0.045 bpp and PSNR: 39.99.

FIGURE 6: Reconstructed images of cervical spine from HMVQ and SPIHT.

Once the user decodes the transmitted image data, the images are usually displayed on a 2D display monitor. Human binocular vision, however, perceives 3D shapes exploiting the disparity of the corresponding pixels in the images [34]. Multiview high-resolution autostereoscopic images provide significant improvement in visual information transmission and display, and may form an integral part of future communication systems with applications in a number of areas such as telemedicine [1, 2]. Some preliminary work in multiview including autostereoscopic video compression is already in progress in the digital layered MVP

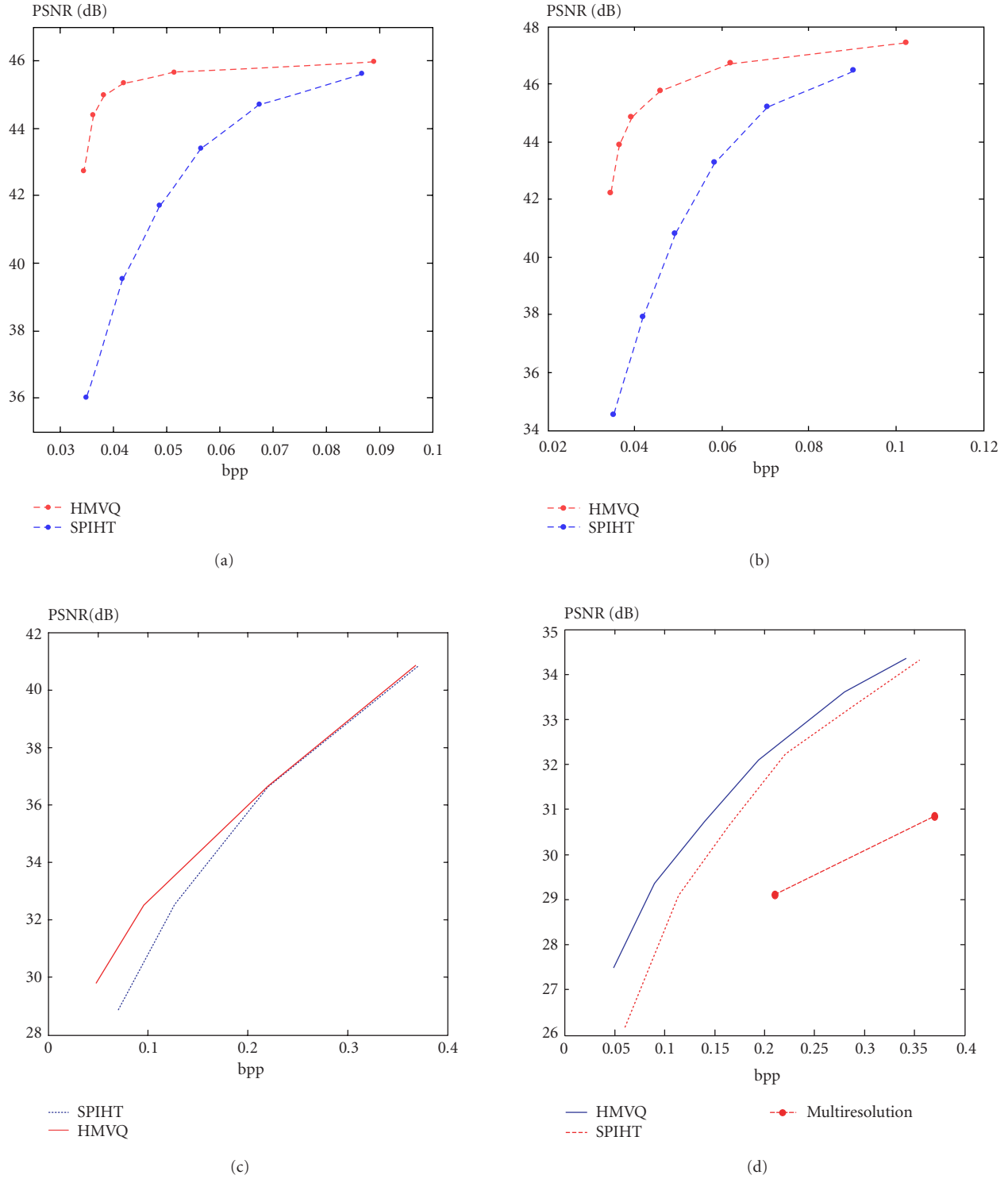


FIGURE 7: Comparison of reconstructed image quality in terms of PSNR. Clockwise: lumbar spine, cervical spine, Lena, and MR images.

(multiview profile) mode of the MPEG-2 standard. However, further research in algorithmic development for high fidelity video compression is needed where human binocular vision characteristics can be exploited to reduce transmission costs [1].

Efficient digital design of such communication systems is extremely challenging and requires innovative ideas in developing algorithms for 3D reconstruction and display of the 3D objects embedded in an image which can be processed by specialized DSPs. We have presented the concept of





FIGURE 8: Comparison of edge preservation on the sections of cervical spine, Lena, and MRI images.

a multiview digital autostereoscopic system including signal processing modules for efficient extraction of depth, color, and texture information for high resolution 3D display of embedded objects in image sequences acquired from medical as well as natural environments.

## 5. CONCLUSIONS

We have demonstrated the ability of a hybrid encoding scheme such as HMVQ in yielding superior performance

over a well-known current encoding scheme, namely, SPIHT, both quantitatively and perceptually in encoding some medical images even at low bit rates. Although intensive researches and analyses on the use of wavelets in image coding have already been reported [11], difficulties still exist in generating an efficient global codebook by vector quantization as evident by the popularity of SPIHT, a wavelet-based scalar quantization method for image encoding. Future success and acceptance of a hybrid coding, using a combination of vector and scalar encoding as in HMVQ for medical image encod-

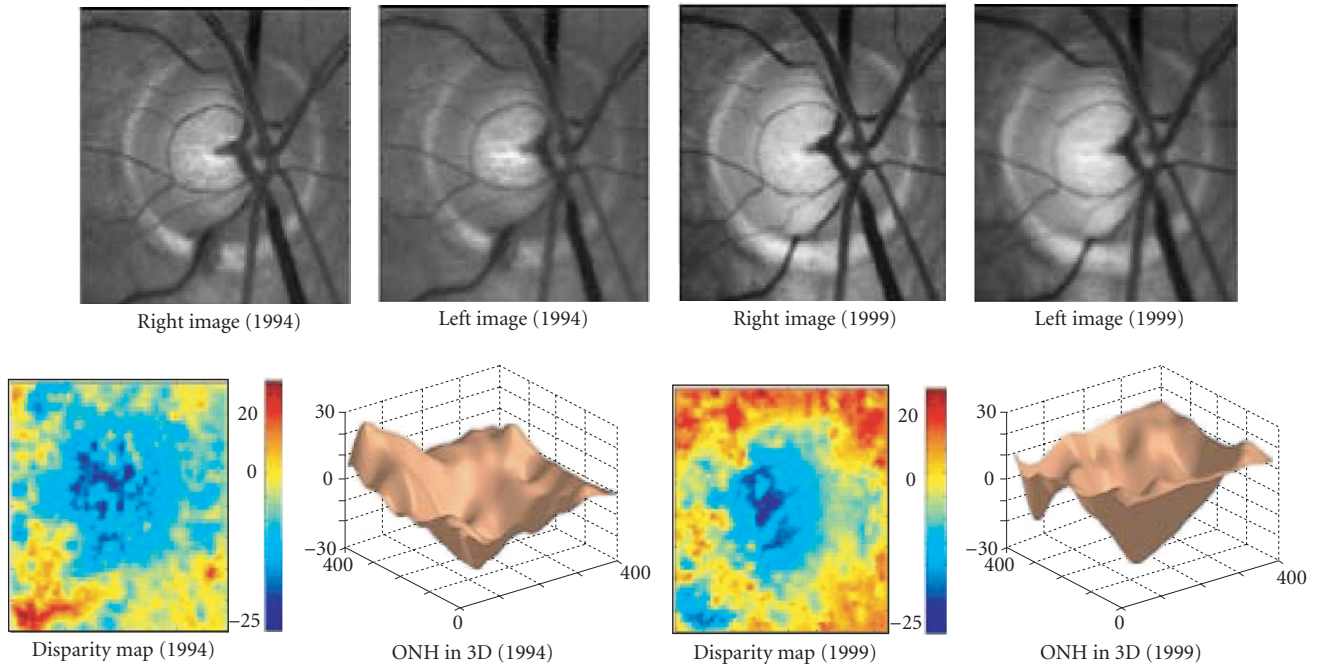


FIGURE 9: Fundus images of a glaucoma patient shown on the top left were taken in 1994. Images of the same eye of the same patient taken in 1999 are shown on the top right. The corresponding disparity matrices and depth representations are shown on the bottom.

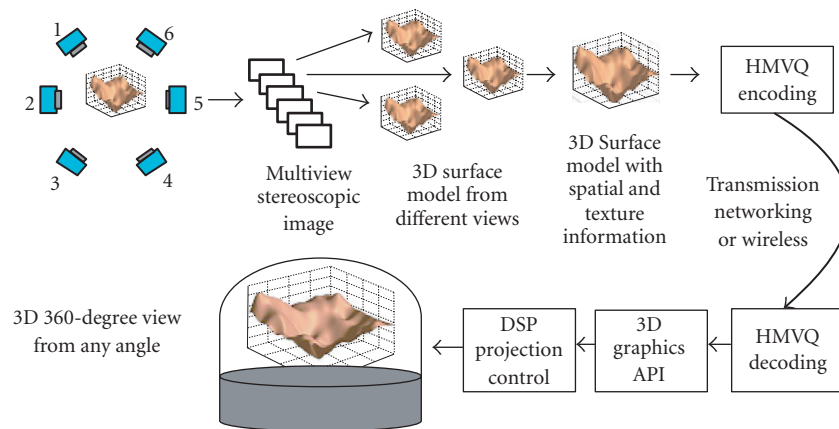


FIGURE 10: A schematic diagram of a multiview 3D digital stereoscopic video communication system.

ing, depend on designing and cascading a lossless encoder module for general classes of medical images as shown in Figure 1. Our current results do not include the lossless module, thus indicating potential improvement in performance of HMVQ when the design of such a module is completed. At present, we have such a lossless module only for a limited class of X-ray images showing definite improvement in performance in reconstructing such images with high fidelity.

An optimal adaptive wavelet filter technique has also been developed to minimize the energy in the high-frequency subbands and thus maximizing the energy in the

low-frequency subband of images decomposed by wavelet transforms. A wavelet-transformed image can thus be represented using only one-fourth of the data required for the entire image without introducing perceptible distortion [31, 35, 36]. The filter design itself involves a nonlinear, non-convex adaptive optimization under specific constraints to achieve an image representation, which can be efficiently implemented in a compact DSP-based system as shown in Figure 10. Such systems could be of potential benefit to fast transmission of large 2D and 3D medical image data sets while retaining high fidelity.

## ACKNOWLEDGMENTS

This research work has been partially supported by funds from the Advanced Technology Program (ATP) (Grant # 003644-0280-ATP) of the state of Texas, and the National Science Foundation (NSF) Grant EIA-9980296. The authors gratefully acknowledge the National Library of Medicine for the X-ray images and Dr. Peter Soliz of Kestrel Corporation and Dr. Young H. Kwon of the University of Iowa Hospitals and Clinics for the fundus images.

## REFERENCES

- [1] J. Konrad, "Visual communications of tomorrow: natural, efficient and flexible," *IEEE Communications Magazine*, vol. 39, no. 1, pp. 126–133, 2001.
- [2] J.-R. Ohm and K. Muller, "Incomplete 3-D-multiview representation of video objects," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 9, no. 2, pp. 389–400, 1999, Special Issue on Synthetic Natural Hybrid Coding.
- [3] J. M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Trans. Signal Processing*, vol. 41, no. 12, pp. 3445–3462, 1993.
- [4] A. Said and W. A. Pearlman, "A new, fast, and efficient image codec based on set partitioning in hierarchical trees," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 6, no. 3, pp. 243–250, 1996.
- [5] S. Mitra, S. Yang, and V. Kustov, "Wavelet-based vector quantization for high-fidelity compression and fast transmission of medical images," *Journal of Digital Imaging*, vol. 11, no. 4, (suppl. 2), pp. 24–30, 1998.
- [6] S. Mitra and S. Yang, "High fidelity adaptive vector quantization at very low bit rates for progressive transmission of radiographic images," *Journal of Electronic Imaging*, vol. 8, no. 1, pp. 23–35, 1999.
- [7] S. Yang and S. Mitra, "Content based vector coder for efficient retrieval of information," in *BISC International Workshop on Fuzzy Logic and the Internet (FLINT 2001)*, University of California, Berkeley, Calif, USA, August 2001.
- [8] S. Yang and S. Mitra, "Efficient storage and management of radiographic images using a novel wavelet based multiscale vector quantizer," in *SPIE Medical Imaging Symposium*, San Diego, Calif, USA, February 2002.
- [9] S. Yang and S. Mitra, "Statistical and adaptive approaches for segmentation and vector source encoding of medical images," in *SPIE Medical Imaging Symposium*, San Diego, Calif, USA, February 2002.
- [10] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, "Image coding using wavelet transform," *IEEE Trans. Image Processing*, vol. 1, no. 2, pp. 205–220, 1992.
- [11] M. Barlaud, Ed., *Wavelets in Image Communication*, Elsevier Science, Amsterdam, The Netherlands, 1994.
- [12] D. F. Lyons, D. L. Neuhoff, and D. Hui, "Reduced storage tree-structured vector quantization," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, vol. 5, pp. 602–605, Minneapolis, Minn, USA, April 1993.
- [13] D. Mukherjee and S. Mitra, "Vector set partitioning with classified successive refinement VQ for embedded wavelet image coding," in *Proc. IEEE Int. Symp. Circuits and Systems*, pp. 25–28, Monterey, Calif, USA, June 1998.
- [14] R. E. Van Dyck and S. A. Rajala, "Subband/VQ coding of color images with perceptually optimal bit allocation," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 4, no. 1, pp. 68–82, 1994.
- [15] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*, Kluwer Academic, Boston, Mass, USA, 1992.
- [16] E. Vidal, "An algorithm for finding nearest neighbors in (approximately) constant average time complexity," *Pattern Recognition Letters*, vol. 4, pp. 145–147, 1986.
- [17] M. R. Soleymani and S. D. Morgera, "A fast MMSE encoding algorithm for vector quantization," *IEEE Trans. Communications*, vol. 37, pp. 656–659, 1989.
- [18] S. C. Tai, C. C. Lai, and Y. C. Lin, "Two fast nearest neighbor searching algorithms for image vector quantization," *IEEE Trans. Communications*, vol. 44, no. 12, pp. 1623–1628, 1996.
- [19] X. Wu and L. Guan, "Acceleration of the LBG algorithm," *IEEE Trans. Communications*, vol. 42, pp. 1518–1523, 1994.
- [20] Y. Linde, A. Buzo, and R. M. Gray, "An algorithm for vector quantization design," *IEEE Trans. Communications*, vol. 28, no. 4, pp. 84–95, 1980.
- [21] K. Rose, "Deterministic annealing for clustering, compression, classification, regression, and related optimization problems," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2210–2239, 1998.
- [22] S. C. Newton, S. Pemmaraju, and S. Mitra, "Adaptive fuzzy leader clustering of complex data sets in pattern recognition," *IEEE Transactions on Neural Networks*, vol. 3, no. 5, pp. 794–800, 1992.
- [23] S. Yang, *Performance analysis from rate distortion theory of wavelet domain vector quantization encoding*, Ph.D. thesis, Texas Tech University, Texas, USA, May 2002.
- [24] National Library of Medicine, *X-ray image archive*, <ftp://ceb.nlm.nih.gov>.
- [25] L. R. Long, S. R. Pillemar, R. C. Lawrence, et al., "WebMIRS : Web-based Medical Information Retrieval System," in *SPIE Proceedings*, vol. 3312, pp. 392–403, San Jose, Calif, USA, January 1998.
- [26] E. Corona, S. Mitra, M. Wilson, and P. Soliz, "Digital stereo optic disc image analyzer for monitoring progression of glaucoma," in *SPIE Medical Imaging Symposium*, pp. 82–93, San Diego, Calif, USA, February 2002.
- [27] E. Corona, S. Mitra, M. Wilson, T. Krile, Y. H. Kwon, and P. Soliz, "Digital stereo image analyzer for generating automated 3-D measures of optic disc deformation in glaucoma," *IEEE Trans. on Medical Imaging*, vol. 21, no. 10, pp. 1244–1253, 2002, Special Issue on Image Analysis in Drug Discovery and Clinical Trials.
- [28] J. M. Ramirez, S. Mitra, and J. Morales, "Visualization of the three dimensional topography of the optic nerve head through a passive stereo vision model," *Journal of Electronic Imaging*, vol. 8, no. 1, pp. 92–97, 1999.
- [29] D. J. Lee, S. Mitra, and T. Krile, "Analysis of sequential complex images using feature extraction and 2-D cepstrum technique," *Journal of Optical Society of America {A}*, vol. 6, pp. 863–870, 1989, Feature Issue on Pattern Recognition and Image Understanding.
- [30] J.-R. Ohm, "Encoding and reconstruction of multiview video objects," *IEEE Signal Processing Magazine*, vol. 16, no. 3, pp. 47–54, 1999.
- [31] S. Mitra, V. Kustov, P. Srinivasan, and S. Shishkin, "Real time adaptive PR-QMF bank design for image coding using interior-point algorithm," in *9th IEEE DSP (DSP 2000) Workshop*, Waldemar Ranch Resort, Hunt, Tex, USA, October 2000.
- [32] C. E. Shannon, "Coding theorems for a discrete source with a fidelity criterion," in *IRE National Convention Record*, vol. 7, part 4, pp. 142–163, New York, NY, USA, March 1959.



- [33] T. Berger, *Rate Distortion Theory*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1971.
- [34] D. Marr, *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*, W. H. Freeman and Company, San Francisco, Calif, USA, 1982.
- [35] V. Kustov, *Adaptive wavelet filter design for digital signal processing systems*, Ph.D. thesis, Texas Tech University, Texas, USA, December 2000.
- [36] S. Mitra, S. Yang, R. Kumar, and B. Nutter, "An optimized hybrid vector quantization for efficient source encoding," in *45th Midwest Symposium on Circuits and Systems (MWSCAS)*, Tulsa, Okla, USA, August 2002.

**Shuyu Yang** obtained her B.S. degree in communications engineering from the Department of Electrical Engineering of Shanghai JiaoTong University in China in 1992. She worked as a network Engineer, designing optical transmission network in the Design Institute of Guangzhou Telecommunications Bureau (now China Telecom., Guangzhou) from 1992 to 1994. From 1994 to 1996, she was the Manager of system engineering in Guangzhou Telecom., Information System Engineering Co. Ltd., China Telecom, Guangzhou. Dr. Yang obtained her M.S. and Ph.D. degrees in electrical engineering from Texas Tech University in 1999 and 2002, respectively. She is currently undergoing postdoctoral training in Computer Vision and Image Analysis Laboratory at the Department of Electrical and Computer Engineering at Texas Tech University. Dr. Yang's research interests include image compression, image segmentation, medical image analysis, wavelet analysis, pattern recognition, neural network, and fuzzy logic. She has over 20 scientific publications.



**Sunanda Mitra**, a professor in the Department of Electrical and Computer Engineering, at Texas Tech University (TTU), received her B.S. and M.S. degrees in physics from Calcutta University, India in 1955 and 1957, respectively, and her Ph.D. in Physics from Marburg University, Germany in 1966. Since 1988, she has been the Director of the Computer Vision and Image Analysis Laboratory in the Department of Electrical and Computer Engineering. Prior to taking the faculty position at TTU in 1984, Dr. Sunanda Mitra has worked as a research scientist at TTU, TTU Health Sciences Center, and as a visiting faculty at the Mount Sinai School of Medicine in New York. Dr. Mitra's specialization includes medical image segmentation and analysis, data compression, 3D modeling from stereo vision and pattern recognition. Prof. Mitra served on the Board of Scientific Counselors of the National Library of Medicine at the National Institutes of Health (USA) from 1997-2001. She has chaired the Technical Committee of Computational Medicine of the IEEE (Institute of Electrical and Electronics Engineers) Computer Society. She is also on the program committee of the International Medical Imaging Symposium on "Image Processing" sponsored by the SPIE (The International Society for Optical Engineering).



**Enrique Corona** received his M.S.E.E. degree from Texas Tech University in 2002. He worked under the guidance of professor Sunanda Mi, where he developed other projects including a 3D model of the optic nerve head for human eye, pattern classifiers, video and still image compression simulations among others. He obtained his B.S.E.E. from the Universidad de las Américas, Puebla, Mexico in 1999. His main research interests are digital signal processing of one and two dimensions as well as fuzzy and classic control, pattern recognition, and digital systems.



**Brian Nutter** is an Associate Professor at the Electrical and Computer Engineering Department, Texas Tech University. He received his B.S.E.E. and Ph.D. degrees from Texas Tech University in 1987 and 1990, respectively. Dr. Nutter worked as a software/electronics Designer, Manager, and Consultant to a variety of rapid prototyping companies, including 3D Systems (90-92), Soligen (92-98), and MEMgen (consulting). His work provided key technologies in data representation, motion control, and system design. Dr. Nutter worked as Vice President of Engineering at WillowBrook Technologies, a manufacturer of digital telephone systems, from 1998 to 2002. He provided the technical direction for a very innovative telephony solution. Dr. Nutter was a member of the startup team for both WillowBrook and Soligen. His interests include telecommunications, networks, signal and image processing, rapid prototyping, and real-time embedded systems.



**D. J. Lee** received his B.S.E.E. from National Taiwan University of Science and Technology in 1984, the M.S. and Ph.D. degrees in Electrical Engineering from Texas Tech University in 1985 and 1990, respectively, and an MBA degree from Shenandoah University in 1999. Dr. Lee is currently an Associate Professor in the Department of Electrical and Computer Engineering at Brigham Young University (BYU), Provo, Utah. Prior to joining BYU in 2001, he served in the machine vision industry for over eleven years. His work experience includes, staff scientist at Innovision Corporation in Madison, Wisconsin from 1990 to 1995, senior system engineer at Texas Instruments in Dallas, Texas from 1995 to 1996, and R&D manager and V.P. of R&D at AGRI-TECH from 1996 to 2000. He joined Robotic Vision System Inc. in 2000 as the Director of Vision Technology and was responsible of designing the state-of-the-art high speed wafer inspection systems. He has designed and built over 40 real-time machine vision systems and products for various industries including automotive, pharmaceutical, semiconductor, surveillance, and military, etc. His current research work focus is in 3D reconstruction, medical image analysis, object tracking, shape-analysis, and shape-based pattern recognition.

