Wavelet-Based Communication of Medical Image Sequences

K. E. Zachariadis$^2$, N. V. Boulgouris$^2$
N. Thomos$^{1,2}$, G. Triantafyllidis$^{1,2}$, and M. G. Strintzis$^{1,2}$

$^1$Information Processing Laboratory
Electrical and Computer Engineering Department
Aristotle University of Thessaloniki
Thessaloniki 54006, Greece

$^2$Informatics and Telematics Institute
1st Km Thermi-Panorama Rd.
Thessaloniki 57001, Greece

e-mail: \{kezachar, nbly, nthomos, gstritian, strintz\}@iti.gr

Abstract

In this paper we present a novel communication scheme for medical image sequences. The proposed scheme is based on a novel layered wavelet coder which is very efficient for the compression and transmission of medical image sequences over heterogeneous networks (i.e., ISDN, ATM, Satellite Link, UMTS). Whenever the available bandwidth is permissive, visually lossless quality of the reconstructed sequence can be achieved. Experiments on medical image sequences yield very promising results and show the efficiency of the proposed method.

I. INTRODUCTION

Telemedicine systems enable experts in main hospital facilities to examine distant patients. In most telemedicine applications, medical doctors have a limited amount of data to evaluate the gravity of the clinical symptoms and to make a decision regarding the patient’s medical condition. The limitations imposed to the amount of data that are made available to physicians are due to the unavailability of efficient systems for the compression and transmission of large volumes of medical data. The problem becomes more apparent in the transmission of medical image sequences, such as ultrasound image sequences and angiograms, due to the requirement for huge bandwidth in order to transmit them with perfect quality.

Medical image compression has attracted significant attention and a variety of compression systems have been proposed [1]. The new DICOM standard [2] for lossy compression adopted the recently finalized JPEG2000 standard [3], which overcomes most of the limitations (compression at low bit-rates or transmission in noisy environments) of the baseline JPEG. The new standard is a fully scalable (in spatial resolution and quality) still picture compression system. It is based on a subband decomposition and an efficient wavelet coefficient coding technique called Embedded Block Coding with Optimized Truncation (EBCOT) [4], which accommodates the need for increasing resolution and fidelity by generating a fully progressive bitstream.

Unlike medical image compression, medical image sequence compression is not a mature research area. The achievable compression ratio with an intra-image compression scheme (i.e., coding every frame in the sequence as if it is a still image) is limited (at maximum 20:1 for MR images). The coding efficiency can be improved by exploiting the temporal redundancy between successive images in a sequence. The most effective way for redundancy removal is motion estimation/compensation, a well-known technique from DCT-based video coding schemes as MPEG-1/2.

The best performing image sequence coders employ the wavelet transform [5] which can achieve very good information compression. In [6] a wavelet-based video coder was proposed using 3D wavelet transform. A similar scheme was recently implemented in [7] in the framework of a medical image sequence system. In this paper, we propose a novel medical image sequence coder based on the 2D wavelet transform. Our coder is endowed with the property of Fine Granular Scalability (FGS) [8], as depicted in Fig. 1. The bitstream comprises of two layers: the base and the enhancement layer. Due to the FGS property, the enhancement layer can be truncated and decoded at any length. Practically, our coder guarantees a minimum acceptable video quality, over heterogeneous networks, while offering visually lossless quality when a high bitrate channel is available. Experiments conducted on medical image sequences yield very promising results and show the suitability of the proposed method for medical applications.
The paper is arranged as follows. In Section II, the proposed medical image sequence coding methodology is presented. In Section III, the advantages of the FGS wavelet coder for medical applications are explained. Experimental results are reported in Section IV and finally, conclusions are drawn in Section V.

![Diagram of FGS scalability structure at the encoder and decoder.](image)

**Fig. 1.** FGS scalability structure at the encoder and decoder.

### II. Video Source Coding

#### A. Motion Compensated Prediction

The proposed system for the generation of embedded bitstreams is depicted in Fig. 3. The first frame in each GOP (Group of Pictures, Fig. 2) is intra-coded (i.e., it is treated like a still image) using block-based wavelet coding. The correlation between consecutive frames is subsequently removed using Overlapped Block Motion Compensation (OBMC) [9]. The reference frames used to calculate motion vectors are the original frames in order to ensure that the highest possible precision is achieved in the estimation of the motion vectors. Motion vectors are losslessly coded using the techniques in [10].

![Diagram of GOP (M frames).](image)

**Fig. 2.** Structure of a Group Of Pictures (GOP) in the proposed coding scheme.

Using the previously estimated half-pixel accurate motion vectors, the procedure for the generation of the stream for the inter frames continues as follows: initially, the first inter-frame (noted as P) is compensated. A version of the intra-frame (first frame in a GOP, noted as I), reconstructed using only the base layer of the coded bitstream, is used as reference for the compensation process. The prediction error is derived by subtracting the compensated prediction from the original inter-frame. The prediction error is wavelet transformed and coded into layers. A version of the error frame is reconstructed using only the base layer of the coded bitstream and then added to the compensated frame. The resulting inter-frame (instead of the original), reconstructed using only the base layer, will serve as the reference frame for the compensation of the next inter-frame. The same procedure is iterated until all frames in a GOP are treated.

Using the methodology described above, the proposed wavelet image sequence coder generates a layered bitstream. The bitstream is composed of a base layer, which is used for the derivation of the reference frame in
the motion compensation process, and a complementary enhancement layer that is used to improve the quality and can be truncated at any point. This method has been adopted to ensure that the motion compensation process performed at the encoder can be identically replicated at the decoder even if only the base layer is received. Otherwise, in case the decoder (Fig. 4) is unable to use the same reference frames, errors would accumulate in the decoded video sequence causing a distortion usually termed as drift [11]. With the proposed methodology, the possibility of facing drift at the decoder is eliminated and thus, a reconstructed sequence of high quality is obtained even if only the base layer is received. In our coder the minimum possible quality, which corresponds to the base layer, is equal to a half ISDN line (64 Kbps).

The wavelet coefficients are coded using a simple bitplane encoder which employs the context models in [4]. In the ensuing section, the complete wavelet coding method, used for both intra and inter frames, is described.
B. Block-based wavelet coding of motion compensation residuals

The intra-frame and the motion compensated residuals are decomposed using a wavelet transform based on the 9-7 bi-orthogonal filter bank [5]. The transmission of information is done in a bitplane-wise manner starting from the Most Significant Bit (MSB) to the Least Significant Bit (LSB). Within each bitplane, subbands are encoded in a predefined scanning order from the lowest to the highest resolution (as shown in Fig. 5).

![Fig. 5. Scanning order of subbands.](image)

For each subband, first the coefficients whose most significant bit lie in the bitplane currently coded are identified by comparing them to a threshold \( T = 2^n \) where \( n \) is the index of the bitplane that is being coded. If a coefficient becomes significant then its sign is coded. This process is called significance identification. Similarly, the refinement layer is defined as the one containing the \( n \)-th bitplane of coefficients found significant in previous passes. In order to achieve greater coding efficiency, the significant identification pass is divided in two layers as in [12]: the layer containing predicted significant coefficients and the layer containing predicted insignificant coefficients. The transmission order of the layers are predicted significant, refinement and predicted insignificant.

The number of symbols that have to be coded is reduced if a single bit is coded for each subband indicating if all coefficients in a subband are insignificant. In case a coefficient is found to be significant, the subband is split into smaller code-blocks. The same process is applied to the code-blocks. The smallest permitted code-block size is equal to the lowest frequency subband size.

The symbol stream generated as described above is coded using adaptive arithmetic codes. The context modeling strategy in [4] is followed for the coding of all layers. The sign of significant coefficients is entropy coded using a single arithmetic model. The final bitstream is a succession of layers which very often are independent to each other. This is schematically described in Fig. 6.

![Fig. 6. The bitstream output by the present Layered Image Coder is a succession of layers.](image)

III. ADVANTAGES OF THE PROPOSED CODER FOR MEDICAL APPLICATIONS

Telemedicine can greatly benefit from advanced tools for scalable video coding. Consider the scenario of a physician receiving an echogram sequence which must be examined from a distant location. The sequence may be transmitted via communication channels with different bandwidths, such as a terrestrial cellular network, a satellite link or even a fiber-optic. In a real-time system which captures and encodes an echogram, it is practically infeasible (due to delay and processing power constraints) to encode it in different rates to serve all possible communication links. Due to FGS, the present image sequence coder is able to offer scalability
and bandwidth adaptivity. We assume that all communication links have at least 64 Kbps bandwidth, so the base layer is always received. Provided that the base layer is received, our scheme does not suffer from temporal quality degradation due to drift [11], since only the base layer is needed for the exact replication of the motion compensated prediction that took place during encoding. Furthermore, any additional portion of the enhancement layer can be decoded at arbitrary length and not at some specific rates, as in other layered video coding schemes. Whenever the available rate is enough, the proposed coder can guarantee near lossless performance.

Additionally, wavelet block-based schemes are inherently more robust to bit errors, that may occur in the bitstreams, in comparison to zero-tree [11] like methods. This is due to the fact that the block-based approach allows the independent coding of blocks of wavelet coefficients and therefore facilitates the localization of corrupted streams.

IV. EXPERIMENTAL RESULTS

The proposed video coding scheme was experimentally evaluated for the transmission of a typical echogram image sequence (resolution 400 x 400, 25 frames per second). Specifically, one GOP was used that comprised 10 (M = 10) frames. For each frame, the Peak-Signal-to-Noise-Ratio is used as a measure of the reconstruction quality (in dB):$$\text{PSNR} = 10 \log_{10} \frac{255^2}{MSE}$$
where the Mean Square Error is equal to:

$$MSE = \frac{1}{N} \sum_{i=1}^{N} (x_i - \hat{x}_i)^2$$

where \(N\) is the number of pixels in the image and \(x_i, \hat{x}_i\) are the pixel intensities in the initial and reconstructed image respectively. Consequently, the average quality calculated over the entire GOP is given by

$$\overline{\text{PSNR}} = \frac{1}{M} \sum_{m=1}^{M} \text{PSNR}_m$$  \hspace{1cm} (1)

where \(\text{PSNR}_m\) is the \(\text{PSNR}\) value for the \(m\)-th reconstructed frame.

We compare our video coding technique with a coder based on the JPEG2000 image coding standard [13]. Specifically, every frame in the sequence is compressed independently using the JPEG2000 coder (no motion compensation is deployed). The test parameters were chosen so as to lead to a fair comparison (same wavelet transform, same number of decomposition levels, progression by layer). Results are reported in Table I for several transmission rates. The base layer bandwidth, which corresponds to the minimum possible reconstructed quality, is set to half the bandwidth provided by an ISDN line (64 Kbps). Table I shows a comparison between the proposed FGS Wavelet Video Coder and the Independent JPEG2000, termed FGS-WVC and Ind. JPEG2000 respectively in the table. The results are in terms of mean \(\text{PSNR}\), defined by Eq. (1), over a GOP. For the case of Independent JPEG2000 coding, we equally divide the available bytes over the frames of the GOP. As seen in Table I, our scheme outperforms the Independent JPEG2000 coding for all test rates. In addition, in Fig. 8 we compare the original and the reconstructed images using both schemes. Apparently, our scheme offers better visual quality even at low bit rates (i.e. 128 Kbps). Furthermore, due to its FGS property our scheme supports homogeneity in quality over time. This is seen in Fig. 7, .

V. CONCLUSIONS AND FUTURE WORK

A novel medical image sequence compression scheme was proposed, based on the wavelet transform. Our medical image sequence coder is endowed with Fine Granular Scalability (FGS) properties. The bit-stream comprises of two layers: the base and the enhancement layer. FGS guarantees a minimum acceptable video quality, over heterogeneous (different capacities) networks (i.e. ISDN, ATM, Satellite Link) and at the same
time achieves visually lossless quality whenever there is a high-rate channel (1-2 Mbps). Experiments on image sequences yielded very promising results and showed the prospective of the method.

The present scheme can be improved in a number of ways. Specifically, in order to gain additional coding efficiency and thus better performance in lower bit-rates, the scheme we propose does not utilize integer wavelets. In future implementations integer wavelets will be employed allowing perfect reconstruction of the original sequence. Additionally, channel coding techniques will be incorporated in order to achieve more error-resilient transmission over unreliable channels. Furthermore, a ROI mask could be employed in order to provide the functionality of decoding a certain area of the sequence at a better quality.

VI. ACKNOWLEDGEMENT

This work was supported by the EU IST project “OTELO: Integrated, end-to-end, mobile tele-echography system”.

REFERENCES


Fig. 8. Reconstructed images for transmission of the 8th frame of the echogram sequence in a channel of rate 128 Kbps: (a) Original frame, (b) Frame reconstructed using FGS-WVC (33.89 dB), (c) Frame reconstructed using Ind. JPEG2000 (32.73 dB)


---

Note: The text is not fully legible due to the quality of the image, but it appears to be discussing image compression techniques and their applications in various contexts.
