Motion Estimation and Wavelet Transform in Angiogram Video Coding

Min-Jen Tsai, B.K.T Ho, and John D. Villasenor

Electrical Engineering Department
University of California, Los Angeles

Abstract - Angiogram video sequence compression based on the wavelet representation and fast motion estimation is presented in this paper. The characteristic motion of angiograms is investigated, leading to a statistical result that is analyzed for fast motion estimation in order to get good interframe prediction. Wavelet transform decomposes the video frame into a set of sub-frames in different resolutions corresponding to different frequency scales. Weighted scale and uniform scalar quantization are applied to the subband coefficients according to their perceptual importance to the reconstructed image. Arithmetic coding is also used to achieve high coding efficiency for either base frame or error frame. The proposed algorithm not only achieves low bit-rate coding but also maintains high fidelity which is essential for medical images.

I. INTRODUCTION

The discrete wavelet transform (DWT) has recently received considerable attention in the context of image processing because it realizes good coding performance. Theoretically, its energy compaction properties outperform other conventional methods [1]. It also offers a flexible multi-resolution image representation and can be designed to incorporate the properties of the human visual system for efficient image compression. It localizes the visually "relevant" information and inherent parallelizability is suitable for efficient VLSI layout.

The advantages of the discrete wavelet transform include reduction of the artifacts introduced by Fourier-based spectral methods and elimination of the distortion arising from data blocking. There are two types of degradation existing in conventional orthogonal transform coding. Those are blocking effect and "mosquito" noise. The degradation, blocking effect, is due to the fact that their transform bases are discontinuous at block boundaries. The mosquito noise (or called, "corona effect, halo effect"), comes from quantization error of higher frequency components due to object edges which spans across the block in the transform domain. The use of wavelet transforms can reduce both types of degradation because wavelet transform bases for higher frequency are shorter than for lower frequency, therefore higher frequency noise does not spread out apart from the object edges.

From the medical image diagnostic point of view, maintaining a reliable quality in compressed medical image is the first priority. Wavelet transform has successfully reduced these two types of degradation. We will investigate the application in angiogram images and show its performance.

In this paper, we use normalized mean square error (NMSE) as the optimality criterion. Even NMSE is not optimal from a perceptual point of view since the human eye is the end receiver of the effect of the quantization error in the reconstructed images. NMSE still allows fast computation and offers a simple and efficient index when we want to make the image quality comparison.

This paper is organized as follows: section 2 describes the framework of wavelet transform and the coding scheme. Section 3 presents the fast motion estimation algorithm. Section 4 will show the simulation results.

II. Wavelet Transform and Coding Scheme

A. Wavelet Transform

The utility of wavelet decomposition of the image for compression lies in the ability of the wavelet transform to concentrate a large percentage of total image energy in the low-frequency terms. The ability to concentrate power in the low-frequency terms depends largely on the wavelet coefficient set used.

Wavelet sets can be selected to span the frequency spectrum in many different ways. In this paper, we use a set of biorthonormal bases with compactly supported wavelets developed in [2]. Compact support implies a finite length for analysis and synthesis filters. There is a compromise between the degree of compactness and the degree of regularity. The wavelet function becomes more regular as the number of taps in filter increases, which results in more computations. The coefficient sets from [2] are used in this work since they show a sufficient energy concentration for the low-frequency wavelets.

The 2-D separable wavelet decomposition can be implemented first in column and then in row independently. The decomposed image forms a pyramid structure up to three levels (Fig. 1). We have decomposed the angiogram (Fig. 2) with the filters and the transformed image is in Fig. 3.

B. Quantization

After the transformation is complete, we have a collection of subimages. In quantizing those coefficients, various approaches have been proposed in the literature that explore their statistical characteristics and perceptual relevance, as well as the existing inter/intra band correlation [5,6,7].

Several different techniques involving vector or scalar quantization are used to encode wavelet coefficients in order to achieve higher compression. According to [3,4], vector

0-7803-1487-5/94$04.00©1994IEEE

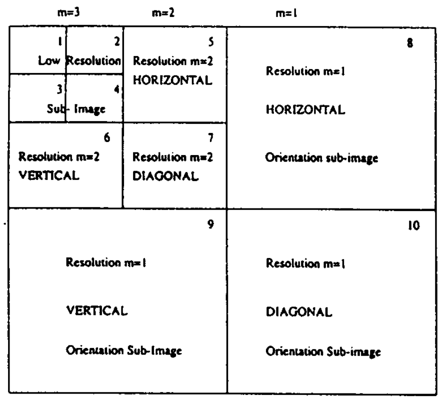


Fig. 1. The image wavelet representation

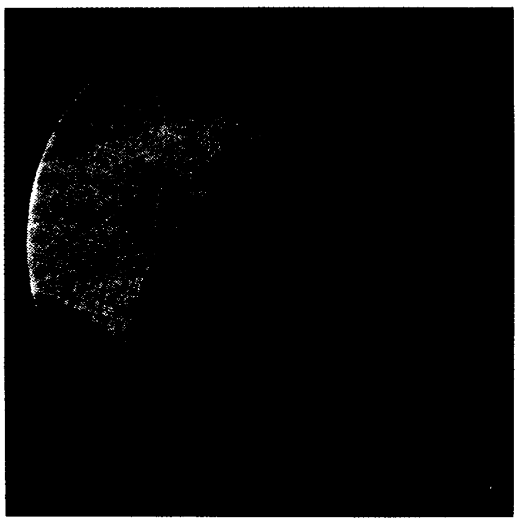


Fig. 2. Angiogram image

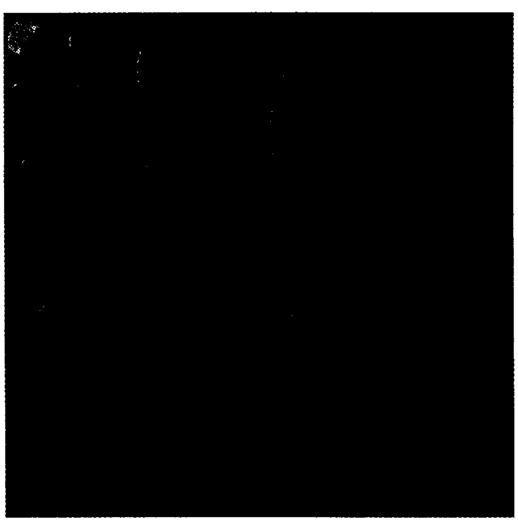


Fig. 3. Wavelet represetation on three levels of image Fig 2.

quantization(VQ) attains an asymptotic lower bound distortion than scalar quantization(SQ). However, the complexity of VQ grows exponentially with the block size. The restricted codebook space cannot reproduce all possible patterns.

For scalar quantization, it was shown that uniform quantization allows the optimum entropy coding of coefficients [8]. This leads us to invent a simple, efficient SQ scheme based on the hardware consideration as follows.

First, we need to perform normalization for the base frame (the intra-pictures which is not predicted for video sequence) and error frame(the difference image between original image and the predicted image which comes from motion compensation, see Section III) which determines the maximum absolute value from the subimages within one frame as the divisor. Dividing all the coefficients by this value, the converted coefficients will be distributed from -1 to 1. Since most of the image data format is 24 bits, we will multiply those converted coefficients by 2^{23} -1 (1 bit for the sign bit) and get the rescaled coefficients. Then, we rescale those subimages and digitize them from floating point format to integer format. Restricting the data format to 24 bits will make our scheme compatible and easy to implement in the hardware.

The next step will be to perform the uniform quantization for all bands. We throw out the less significant bits and round the remaining value. For example, value $6_{decimal}$ (0110_{binary}) bitwise rounds to $8_{decimal}$ (1000_{binary}) if we skip last three least significant bits. We can see the statistic result of the relationship between the number of bits discarded (shift bits) and its corresponding NMSE for base frame and error frame from Table I,II.

C. Coding Scheme

The resulting quantized coefficients can be coded in a number of ways. The traditional method is to employ a Huffman code, whereas the more recent method is to use one of the newer entropy or arithmetic coders. The advantage of the latter methods is that they require no prior analysis of the data set to determine bit allocation [9], and achieve the theoretical entropy bound to compression efficiency.

Table III, IV show the relationship between the number of discarded bits (shift bits) and the compression ratio by using arithmetic coding for the base frame and the error frame. It indicates the number of least significant bits we can discard with respect to the desired compression ratio.

III. Fast Motion Estimation

Generally, good performance for image sequence coding relies on the use of interframe coding techniques which

Shift Bits	NMSE(%)
23	47.761826
22	16.140457
21	3.952601
20	1.232896
19	0.414296
18	0.183515
17	0.091701
16	0.041668
15	0.021960
14	0.010639
13	0.003381
12	0.000872

TABLE I. Base Frame: shift bits to NMSE

Shift Bits	NMSE(%)
23	97.294205
22	91.120781
21	76.569809
20	56.729385
19	36.157211
18	16.801336
17	4.966285
16	1.260331
15	0.318008
14	0.080273
13	0.020216
12	0.005047

TABLE II. Error Frame: shift bit to NMSE

Shift Bits	NMSE(%)
23	517.05
22	323.24
21	218.09
20	165.39
19	124.12
18	88.59
17	55.27
16	32.31
15	17.50
14	8.25
13	4.49
12	2.98

TABLE III. Base frame: shift bits to compression ratio

Shift Bits	NMSE(%)
23	1191.56
22	819.20
21	358.57
20	105.41
19	28.99
18	9.63
17	4.97
16	3.27
15	2.43
14	1.92
13	1.58
12	1.34

TABLE IV. Error Frame: shift bits to compression ratio

exploit both spatial and temporal correlations among signals. Then the resulting error images are compressed and transmitted along with the corresponding displacement vectors.

There are several fast motion estimation method based on pixel subsampling for fast computation [10]. All of these methods try to reduce the number of pixels used in the calculation during the block matching search. Fig. 4 shows the graph of how to choose those pixels in each block. Where the "Decimation" method will subsample every other pixel within the 16x16 block, the "Star" method calculate only those pixels on the diagonal line, the "Avg 2x2" will only use the upper left 2x2 block (only four pixels) to do the block matching calculation.

By performing motion estimation, we can get the plot of NMSE from different subsampling methods(see Fig. 5). Assuming we can perfectly reconstruct the predicted image for further use, the prediction phase follows: original frame 0 predicts frame 1, original frame 1 predicts frame 2, ... The value after each subsampling name is the error measure which is the ratio defined by the area of the difference between the test search method and exhaustive search divided by the area of the

XOXOXOXO	XOOOOOOX	XXOOOOOO
OXOXOXOX	OXOOOOXO	XXOOOOOO
XOXOXOXO	OOXOOXOO	OOOOOOOO
OXOXOXOX	OOOXXOOO	OOOOOOOO
XOXOXOXO	OOOXXOOO	OOOOOOOO
OXOXOXOX	OOXOOXOO	OOOOOOOO
XOXOXOXO	OXOOOOXO	OOOOOOOO
OXOXOXOX	XOOOOOOX	OOOOOOOO
Decimation	Star	Avg 2x2

Fig. 4. Subsampling by choosing 'x' location pixels for decimation, star and avg 2x2 schemes

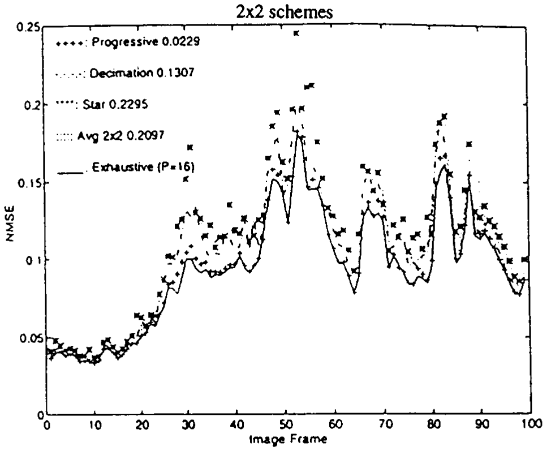


Fig. 5. NMSE in different subsampling scheme

exhaustive search. Fig. 5 indicates that subsampling was not unable to achieve a qualified quality for angiogram video sequence.

Fig. 6 shows NMSE from the exhaustive search for different search ranges. Reduction of the search range will definitely reduce the computation burden but may fail to find the best matched block existing outside the search area. Generally speaking, the motion in an angiogram is comparatively fast. Increasing the search range will increase the chances of finding a better match for motion vectors..

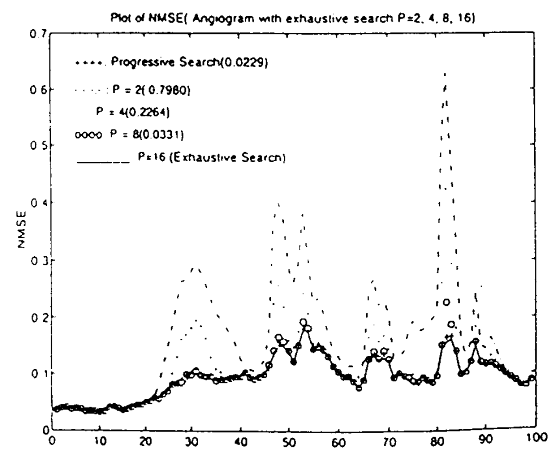


Fig. 6. NMSE in different search range

For the search step, we should point out that the procedure to find the best matched motion vector should be searched radially from the origin instead of raster scanning order(Fig. 7). We want to minimize the vector length.

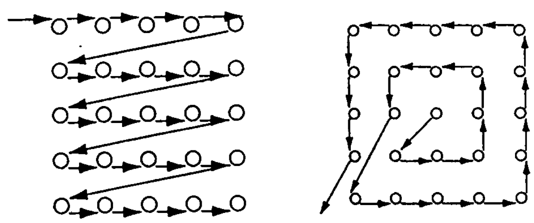


Fig. 7. (a) Raster scanning search (b) Radial loop search

We propose the residue progressive search method to reduce the computation expense and maintain trustful image quality. First, we obtain the subtraction image for adjacent frames. If there is a 100 frame video sequence, we will get 99 frame subtraction images. Analyze the energy residue for each 16x16 block with respect to the best matched motion vector from exhaustive search whose search range is from -16 to 16. We can get an energy residue to search range table(for instance, if the best matched motion vector is [5, -8], the search range for that block is 8).

We assume the energy residue will reflect the movement of image content. From the statistical result, we empirically pick the threshold for five levels: 0, 1, 2, 4 and 8. According to the residue for each block, we set an initial search range for each block depending on which level it falls into. During the search, we will dynamically adjust the search range depending on whether the mean absolute difference decreases or not. If it doesn't decrease, we switch to the other block to do search. If it decreases, we extend one more radial search each time until twice the initial setting is reached. This interactive function will guarantee that the search minimizes the error with minimum expense. Even the threshold setting is empirical and intuitive, as long as we know the property of the image, it does give us prior information for our prediction. Fig 5. shows this search strategy has very low error measure value among other fast search algorithms.

This scheme could be summarized as follows:

1. Set threshold
2. Decide the initial search range = P for each block
3. Search each block:
   - 3.1 if $P = 0$, stop.
   - 3.2 if $P \geq 0$, then
     for i=0 to P
     find the minimum mean absolute difference D.
   - 3.3 try i = P+1,
     find the minimum mean absolute difference
     - 3.3.1 if $D' \geq D$, stop.
     - 3.3.2 else, while($D' \leq D$ & $i \leq 2P$),
       modify D', increase i by 1;
       otherwise, stop.
4. search next block .

IV. Simulation Results

Fig. 8 shows the NMSE for the video sequence using the algorithms described above. Every tenth frame, we resend a base frame. Forward prediction for motion estimation is performed. We have tried to maintain a low, constant NMSE level for the image sequence. Perceptually, NMSE below 0.05% is almost indistinguishable from the original images. Since the compression ratio is high(30:1), it is applicable and acceptable for implementation.

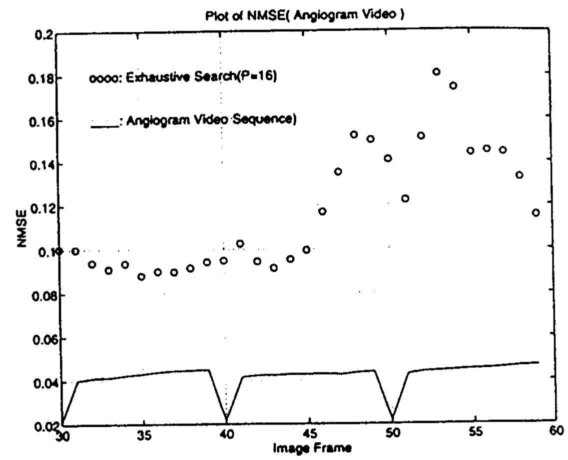


Fig. 8. NMSE of Angiogram video Sequence

In terms of computation saving, we can choose the divisors for base frame and error frame as fixed numbers which can reduce the time to find the max value among the subbands. The number of shift bits for base frame and error frame could also be set since their statistic distribution for individual subband are very close. From our test for angiograms, the divisor for base frame is 16384, shift bit is 12 and the divisor for error frame is 512, shift bit is 18. The compression ratio is still more than 30.

A block diagram of the proposed codec is shown in Fig 9.

V. Conclusions

This paper describes a unified framework to medical image coding techniques including the wavelet transform, normalization, uniform quantization, arithmetic coding and efficient motion estimation. Advantages of the proposed mechanism over virtually all known techniques is its simplicity, efficiency and compatibility. Low bit-rate coding for maintaining high fidelity of the medical image is achieved.

REFERENCES

[1] Daubechies, I., "Orthonormal Bases of Compactly Supported Wavelets," Comm. Pure Appl. Math. 411:909-996(1988).

[2] M. Antonini,"Image Coding Using Wavelet Transform", *IEEE Trans. Image Processing*, vol.1, No. 2, pp205-220, Apr. 1992.

[3] P. Zador, "Asymptotic quantization error of continuoous signals and their quantization dimension," *IEEE Trans. Inform. Theory*, vol. IT-28, pp. 139-149,1982.

[4] J.H. Conway, N.J.A. Sloane, "A lower bound on the average error of vector quantizers", *IEEE Trans. Inform. Theory, vol.* IT-31, pp. 106-109, Jan. 1985.

[5] P.H. Westerink, D.E. Boekee, J. Biemond, and J.W. Woods, "Subband Coding of Image Coding Using Vector Quantization" *IEEE Trans. Computers* 36:713-719, June 1988

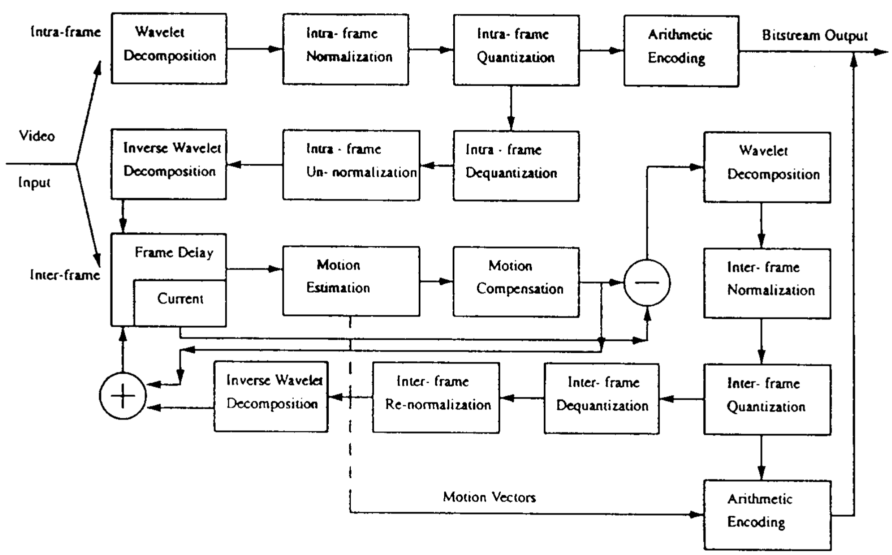
[6] H. Gharavi and A. Tabatabai, "Subband Coding of Monochrome and Color Images," *IEEE Trans. Circuits and Systèms* 35:207-214, February 1988.

[7] Y.H. Kim and J.W. Modestino, "Adaptive Entropy coded Subband Coding of Images " *IEEE Trans. Image Processing*, 1(1):31-48, Junuray 1992.

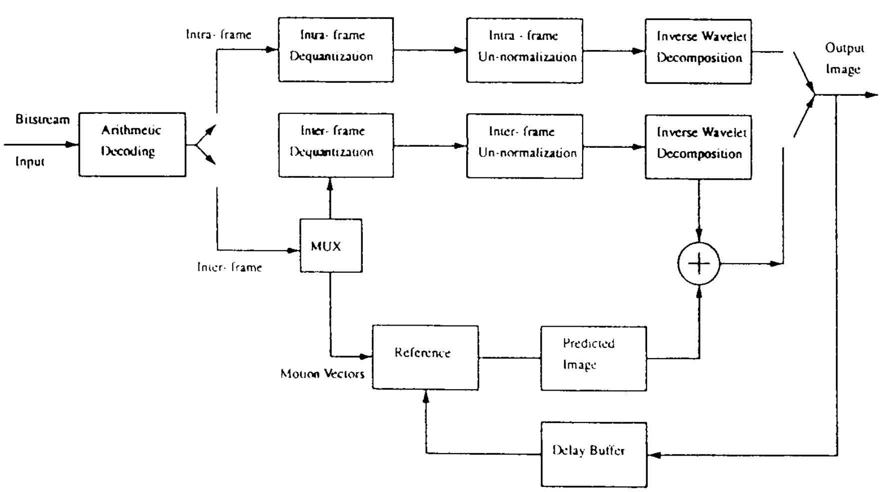
[8] T. Berger, "Optimum Quantizers and Permutation Codes," *IEEE Trans. Information Theory* 18(6), November 1972.

[9] Ian H. Wittern, Radford M. Neal, and John G. Cleary, "Arithmetic Coding for Data Compression" *Communication of the ACM*, Junuray 1987, Volume 30, Number 6.

[10] Andre Zaccarin, Bede Liu, "Fast Algorithm for Block Motion Estimation " *ICASSP*, 1992, Volume 3:449-452.



Schematic Block Diagram of the Encoding Process



Schematic Block Diagram of the Decoding Process

Fig 9.. Encoder and Decoder