Wavelet Based Image Sequence Compression

*A. Aksay *, G. Bozdagi *, M. B. Akhan ***

* Department of Electrical and Electronics Engineering,
Middle East Technical University, Ankara, Turkey
** Department of Electrical and Electronics Engineering
University of Hertfordshire, Hatfield, Herts, UK

ABSTRACT

The idea behind wavelet compression lies in selectively transmitting and reconstructing only those spatial frequencies that are most significant to the eye. Due to this fact, wavelet compression has the best compression rate for 2-D images. However if a sequence of images is to be compressed, temporal correlation must also be utilized to decrease the bit-rate while keeping the visual quality. The aim of this work is to compare several algorithms to code a sequence of images using wavelet transform and evaluate their performances in terms of bit-rate, SNR, visual quality and computational time.

I. INTRODUCTION

The literature on video coding using wavelet transform can be divided into 2 categories: 2-1/2 D and 3 D transforms. 3 D transforms assume the video signal as a 3 D signal in x, y and t and perform transformation in 3 D [1]. These techniques bring complexity and also do not give promising results. In 2-1/2 D wavelet schemes, motion information either in the time or wavelet domain is combined with the 2 D wavelet transform. For time domain motion estimation techniques [2], wavelet transform is applied to motion compensated residue images. This technique distorts the perceptual quality and coding efficiency because of the sharp edges found in the motion compensated residue images. In order to overcome this problem, overlapped block matching techniques can be employed [3], [4]. By this technique, sharp edges found in the residue image disappear, since the algorithm is somewhat averaging the possible candidates for each pixel. Motion estimation in wavelet domain is also possible [5], [6] . In this case, previous and current images are wavelet transformed and in that domain motion estimation and difference is employed. The third possibility is to extract the time-domain motion vectors in wavelet domain. First problem in this scheme is that due to subsampling in wavelet transform. Since the transformation is shift variant, correct time domain motion vectors can not be found in case they are not power of 2. To make the transform shift invariant, complex wavelet transform is introduced with the cost of redundancy of 4:1 in wavelet coefficients [7], [8]. Thus, this transform enables one to find real valued motion vectors. However, in the case of video compression, the main aim is to decrease the bit-rate while keeping the visual quality high. Due to this fact, motion vectors do not have to match with the correct ones if they decrease the bit-rate. In this work, we do not experiment on the complex wavelet transform due to this fact.

The rest of the paper is organized as follows. Section 2 describes the techniques that are the going to be compared, namely: predictive coding, motion compensated coding and wavelet domain motion compensated coding in detail.
Section 3 gives the full coder-decoder structure. Experimental results are given in Section 4. Finally Section 5 presents the conclusions.

II. 2-1/2 D WAVELET SCHEMES

In 2-1/2 D wavelet schemes, motion information either in the time or wavelet domain is combined with the 2 D wavelet transform. The easiest way to achieve this is to assume the motion vector as 0 and use the temporal differences of consecutive frames as inputs to wavelet transform. The method is called as **predictive coding**.

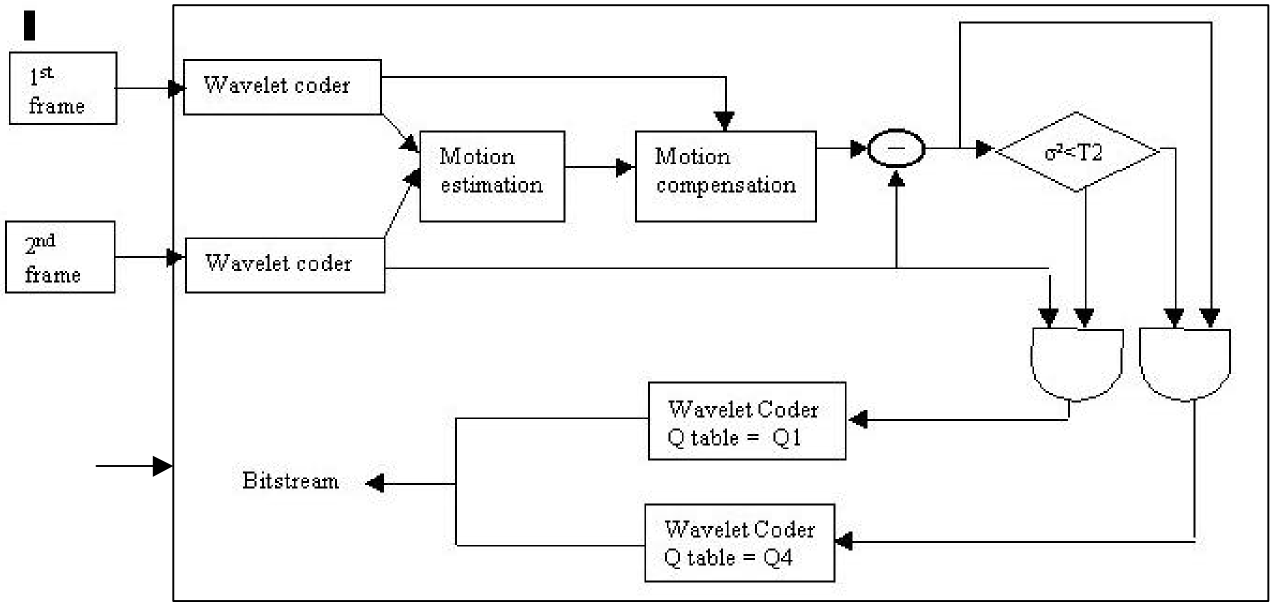
Since only the differences are coded after the initial frame, one of the disadvantages of this method is the error propagation. To overcome this, in our experiments, we send an original frame at each T frame. T is adaptive and is based on the energy within the difference image. Since the image content of the difference frame is completely different than the original frames, a separate quantization table is used in each case.

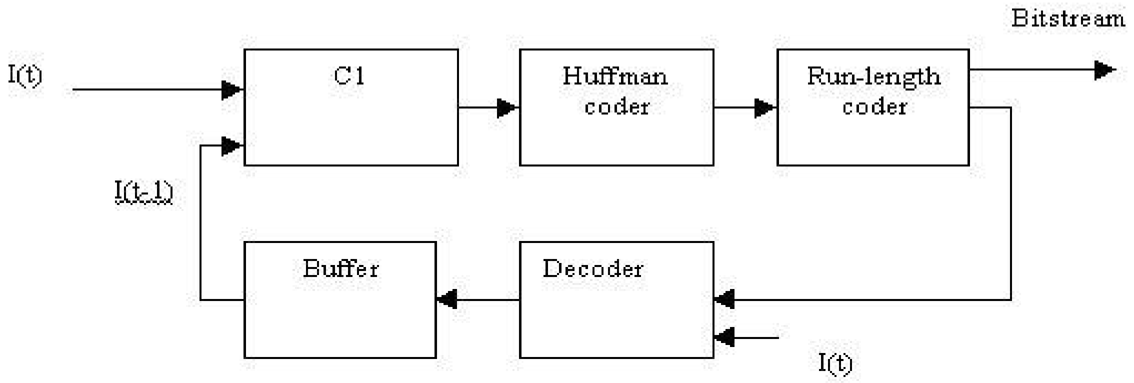
Motion compensated coding: In this method, the current frame is estimated from the previous frame in time domain. Hierarchical methods, exhaustive or 3-level search techniques can be used for motion estimation. Although exhaustive search is the most computationally extensive method, in order to get the best estimation, we used this technique in our experiments. Once the estimation is done, the difference of the motion compensated and the original current frame is coded using wavelet transform.

Since only the differences are coded after the initial frame, one of the disadvantages of this method is the error propagation. To overcome this, in our experiments, we send an original frame at each T frame. T is adaptive and is based on the energy within the difference image. Since the image content of the difference frame is completely different than the original frames, a separate quantization table is used in each case.

Motion compensated coding: In this method, the current frame is estimated from the previous frame in time domain. Hierarchical methods, exhaustive or 3-level search techniques can be used for motion estimation. Although exhaustive search is the most computationally extensive method, in order to get the best estimation, we used this technique in our experiments. Once the estimation is done, the difference of the motion compensated and the original current frame is coded using wavelet transform. Again a different quantization table is used for the difference and original frame coding. This method produces lower energy difference signals due to compensation and yield better coding than predictive coding as shown by the experiments.

Wavelet domain motion compensated coding: In the motion compensated wavelet coding technique, the system doesn't utilize the advantages of the wavelet transform, since the motion estimation is done before taking the transform. To utilize the advantages of the transform, we perform a hierarchical motion estimation in the wavelet domain. The motion estimate of each scale is used as a starting point for the higher scale. After the compensation the energy of diagonal quarter is used as a fall back decision mechanism. If the energy is above a threshold than the system automatically sends the original image instead of the difference image. A detailed block diagram of the motion compensated wavelet coding used in the experiments is shown in Fig. 1.





III. CODER DECODER DESIGN

In order to compare the results of the 3 algorithms full coder-decoder is designed. Daubechies wavelets with 4-taps are used since they are easy to implement. The quantization tables are chosen based on the characteristics of the image to be coded. The values used for the difference images is given in Table 1. For each subband there is a bin-width number, by which all the coefficients in that subband are quantized with. As higher frequencies are not visually so much important, bin-width of higher frequencies are lower numbers. The Huffman and run-length coding is applied as the last step of the coder. For decoding the reserve operation is performed. The full structure of the coder is given in Fig. 2.

IV. EXPERIMENTAL RESULTS

In the experiments, *"Calendar"* and *"Miss America"* sequences are used. The results from simulations of 3 algorithms have been obtained and compared.

For motion estimation in wavelet domain, we have tried two search algorithms. In both algorithms, a block size of 4x4 is used. The first algorithm is exhaustive search with a search window of variable size with no initial guess. In the upper subband, the search size is set to be 4 pixels. When going from one subband to the next level, the search size is multiplied by two. Second algorithm is an exhaustive search with a search window of one-pixel depth. At each subband, the initial guesses from previous subbands are used after appropriate scaling. At the upper level, we still have a search size of 4 pixel.

In time domain motion estimation, exhaustive search with a block size of 16x16 is used. When we use smaller block size, energy of residue image in time domain is reduced whereas energy of the residue image in wavelet domain is increased as the boundaries of the transform and the boundaries of motion estimation does not align.

When entropy of motion vectors are calculated, we have two different values for the motion vectors of wavelet domain. Since for each block we use the vectors from the previous subband and refine only by one pixel, we only need the refinement vector as motion information. So in our calculation, we use the refinement vector as the actual motion vector. The entropy of the motion vectors is shown in Table 2 for different algorithms.

Based on the complete coder-decoder design, Table 3 shows the average bitrate obtained for *Miss America* sequence over 30 frames using two different quantizatin tables. In Fig. 3 the original and the reconstructed 8th frame based on the wavelet based decoder are shown. Table 4 shows the bit-rate for the same experimental set-up for *Calendar* sequence.

Subband	Bin-Width
HH1	0xB3
HL1	0xd7
LH1	0xd7
HH2	0xdc
HL2	0x107
LH2	0x107
HH3	0x114
HL3	0x147
LH3	0x14a
HH4	0x16b
HL4	0x16b
LH4	0x16b
LL4	0x170

Table 1. Quantization table for high-frequency images.

	Entropy of motion vectors
Wavelet domain (av Q)	2.96
Time domain (av Q)	6.11
Wavelet domain (max Q)	3.01
Time domain (max Q)	5.70

Table 2. The entropy of motion vectors for the *Miss America* sequence.

estimation type	quant	Bpp	snr
predictive	av	3.2	39.7
time domain	av	2.22	40.04
wavelet domain	av	2.83	40.36
predictive	max	5.2	44.04
time domain	max	4.01	44.17
wavelet domain	max	4.75	45

Table 3. Average bitrate for *Miss America* sequence over 30 frames.

Fig. 3. The original and reconstructed 8th frame.

estimation type	quant	bpp	snr
predictive	av	0.95	41.6
time domain	Av	0.9	41.72
wavelet domain	Av	0.9	42.21
predictive	Max	2.44	44.8
time domain	Max	2.33	44.97
wavelet domain	Max	2.27	45.69

Table 4. Average bitrate for *Calender* sequence over 310 frames.

V. CONCLUSIONS

Our results show that by using motion estimation in wavelet domain, we increase the SNR with a slight increase in the bitrate. When the computational times are compared the wavelet domain algorithm is at least 10 times faster than the time domain algorithm which makes it favorable for real time applications.

VI. REFERENCES

1. D. Taubman and A. Zakhor, “Multirate 3-D subband coding of video,” IEEE Trans. Image Processing, vol. 3, pp. 572-588, Sept. 1988.
2. Z. Xiong, K. Ramchandran, M. T. Orchard and Y.-Q. Zhang, "A Comparative Study of DCT- and Wavelet-Based Image Coding," IEEE Trans. Circuits Syst. Video Technol., vol. 9, pp. 692-695, Aug. 1999.
3. D. Marpe and H. L. Cycon, "Very Low Bit-Rate Video Coding Using Wavelet-Based Techniques," IEEE Trans. Circuits Syst. Video Technol., vol. 9, pp. 85-94, Feb. 1999.
4. M. Ohta and S. Nogaki, "Hybrid Picture Coding with Wavelet Transform and Overlapped Motion-Compensated Interframe Prediction Coding," IEEE Trans. Signal Proc., vol. 41, pp 3416-3424, Dec.1993.
5. Y.-Q. Zhang and S. Zafar, “Motion-compensated wavelet transform coding for color video compression,” IEEE Trans. Circuits Syst. Video Technol., vol. 2, pp. 285-296, Sept. 1992.
6. Q. Wang and M. Ghanbari, "Scalable Coding of Very High Resolution Video Using the Virtual Zerotree," IEEE Trans. Circuits Syst. Video Technol., vol. 7, pp. 719-727, Oct. 1997.
7. J. Magarey and N. Kingsbury, "Motion Estimation using Complex Wavelets," SPIE Vol. 2825, pp. 674-685.
8. J. F. A. Magarey and N. G. Kingsbury, "Motion Estimation using Complex Wavelets," Technical Report CUED/F-INFENG/TR.226, Cambridge University Engineering Department, August 1995.