International Workshop on Video Processing and Multimedia Communications, June 28-30, 2000.

INTRA-FRAME WAVELET VIDEO CODING

Dr. T. Morris, Mr. D. Britch

Department of Computation, UMIST,
P. O. Box 88, Manchester, M60 1QD, United Kingdom
E-mail: t.morris@co.umist.ac.uk dbritch@co.umist.ac.uk

Abstract: *A software only wavelet based video codec has been designed and implemented. The codec takes a colour video sequence of arbitrary size as input and performs intra-frame compression on the sequence. Video sequences are transformed with a symmetrical wavelet transform, prior to quantisation and entropy encoding. A video codec quality analyser (CQA) is used to assess the subjective quality of the coded stream with respect to the uncoded one in order to provide a single quality measure that correlates with a subjective assessment of the data. Experiments using standard test sequences show that a higher compression ratio is achieved than obtainable using MJPEG, at the same quality level.*

Keywords: *wavelets, video coding*.

1. INTRODUCTION

Much research effort has been expended in the area of wavelet based image compression, with the results indicating that wavelet-based approaches outperform DCT-based techniques for still images [1, 2, 3, 4, 5]. However, when applied to video sequences, wavelet-based techniques have not always shown clear advantages compared to DCT-based techniques such as MPEG [6]. Furthermore, past research has concentrated on coding luminance information only; typically employing CIF and QCIF sized frames.

A key component of an efficient video coding algorithm is motion compensation. However at the present time it is too computationally intensive to be used in software video compression. A number of low end applications therefore use motion-JPEG, in essence frame by frame transmission of JPEG images [7], with no removal of interframe redundancy.

In considering the effectiveness of a video codec, two factors need to be considered. One is the amount of compression achieved and the other is the quality of the subsequently decoded data. The degree of compression is easily measured by comparing the volumes of the coded and raw data. The quality of the decoded data is much more difficult to quantify. It is well known that small errors randomly distributed over the whole frame reflect badly in objective measures, but they often correspond to invisible degradation of frame quality. On the other hand, relatively large errors concentrated in one region of a frame give a high PSNR while the subjective assessment is low because of a corruption of a particular region of the picture. This observation therefore calls into question the usefulness of objective measures.

In recent years, the disadvantages of both subjective and objective measures have motivated researchers to search for a better method of evaluating the quality of video sequences which have undergone lossy compression [8, 9]. A codec quality analyser (CQA) has been developed which aims to provide a consistent bias-free single quality measure, which correlates with a subjective assessment of video data [10].

This paper describes a baseline wavelet based video codec that we have developed. It is designed to take a colour video sequence as input, and compress it in the manner adopted by MJPEG. Test sequences were compressed with the codec, and the results obtained indicate that the codec outperforms MJPEG in terms of compression ratio, at the same level of quality.

The paper is organised as follows. The compression algorithm is described in Section 2. We then present and discuss our experimental results in Section 3. Section 4 evaluates the algorithm and outlines its future development.

2. ALGORITHM

An overview of the algorithm is shown in Figure 1. The algorithm comprises three main processes: wavelet decomposition of the input, scalar quantisation of the wavelet coefficients, and lossless entropy coding of the quantiser indices.

The encoder takes a colour video sequence with arbitrary sized dimensions as input, and transforms it from RGB representation to YCbCr representation, so that the data is represented in a form more suitable for compression.

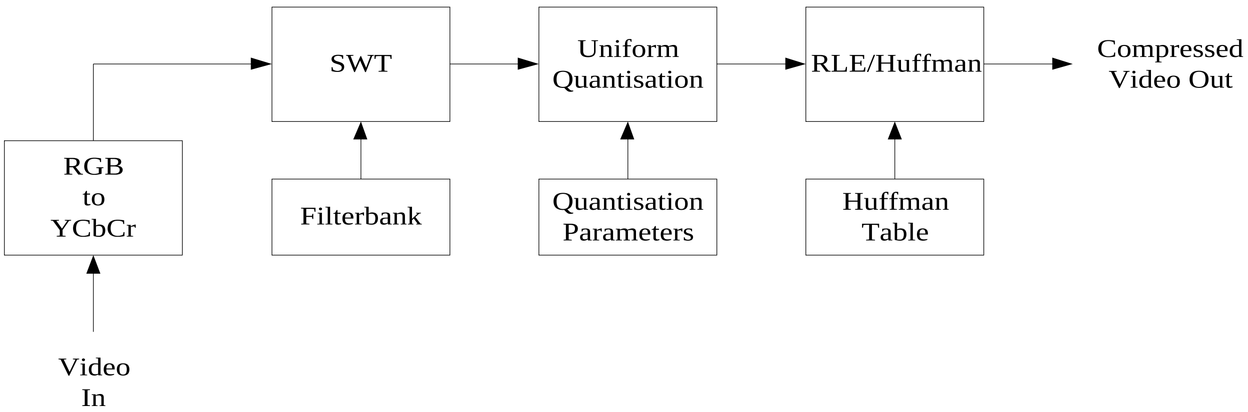

Figure 1 Simplified encoder diagram

Each frame is then successively decomposed by a symmetrical wavelet transform (SWT). The wavelet coefficients are then quantised by a uniform scalar quantiser. The quantiser outputs integer values that are subsequently entropy-encoded by run-length coding and Huffman coding. A single bitstream of data is produced which contains the compressed frame data, along with header information that allows the data to be decoded.

The decoder extracts the information needed in the decoding process from the compressed data. An entropy decoder reconstructs the compressed subbands, and the integer values are decoded to yield quantised wavelet coefficients, which are approximations to the original wavelet coefficients. The quantised coefficients are then inverse transformed before being converted back from YCbCr format to RGB, to produce the reconstructed frame.

2.1. The Wavelet Transform

The DWT is implemented using a two-channel perfect reconstruction linear phase filter bank [11]. Symmetric extension techniques are used to apply the filters near the frame boundaries; an approach that allows transforming images with arbitrary dimensions. A detailed treatment of symmetric wavelet transform methods is given in [12].

The SWT is applied to a frame by transforming first the rows and then the columns of the frame, yielding a two-dimensional, four-channel decomposition. The subband obtained by lowpass filtering in both the horizontal and vertical directions is then cascaded back through the analysis filter bank to produce a seven-channel decomposition. In total the cascade is repeated five times on the resulting lowpass subbands, until a sixteen-subband decomposition is achieved.

The filterbank used was the familiar biorthogonal 9/7 filter constructed by Cohen, Daubechies, and Feauveau [13]. This filterbank was selected based on its abilities to produce reconstructed images with superior quality to those generated by other filters [14].

2.2. Quantisation

Lossy compression is accomplished by uniform scalar quantisation of the SWT subbands. The quantiser is characterised by a set of decision points $\{d_i, i = 0, \ldots, N\}$ and a set of reconstruction levels $\{r_i, i = 0, \ldots, N - 1\}$. A floating point wavelet coefficient, x, is mapped to a reconstruction level r_i if x lies in the interval $\{d_i, d_i+1\}$. The quantiser outputs a set of indices that are entropy-encoded and transmitted in a compressed format. Since this operation is not invertible, the resulting compression is inherently lossy.

To achieve good results a separate quantiser should be designed for each scale, taking into account both the properties of the Human Visual System and the statistical properties of the scale's coefficients. Through experimentation it was found that excellent results could be obtained by quantising subbands 0 to 3 with one set of parameters, subbands 4 to 12 with another, and discarding subbands 13, 14, and 15 altogether. Figure 2 shows the location of the subbands within a frame.

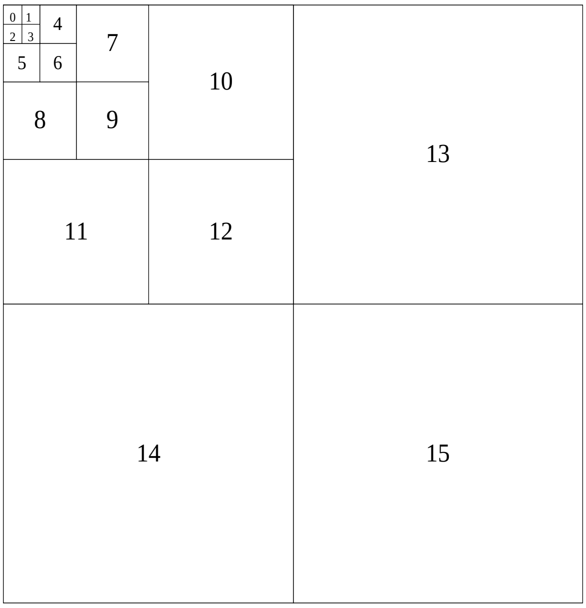

Figure 2 Frequency support of the subbands

Zafar and Zhang [15] have shown that luminance signals contain more than 60% of the total energy of the original signal, with the chrominance signals each having less than 20%. Therefore, the chrominance signals are quantised more coarsely than the luminance signal. Experimentation yielded the values shown in Table 1 as producing a good trade off between compression ratio and reconstruction quality.

Table 1 Reconstruction levels

	Subbands 0 to 3	Subbands 4 to 12
Y	512	128
Cb	128	32
Cr	128	32

2.3. Entropy Encoding

Following scalar quantisation of the subbands within a frame, the integer indices output by the quantiser are entropy-encoded by run-length coding and Huffman coding. The quantised wavelet coefficients are raster scanned for run length coding.

Along with the compressed frame information, the bitstream contains information that consists of video sequence dimensions, parameters for the scalar quantisers and tables for the Huffman coder. The Huffman coding tables are video sequence dependent and therefore must be contained in the coded data format.

The decoder parses the compressed data and extracts the tables needed in the decoding process to produce the reconstructed video sequence

2.4. Codec Quality Analyser

CQA is designed to give a subjective measure of the degradations introduced into a video stream by the coding process. This is achieved by measuring the subjectively important differences between the original and compressed video streams. Full details of these measurements can be found in [10]. The measurements are then combined using a neural network to give a single value that reflects the subjective quality of the video stream. A value of 1 indicates a sequence with an objectionable amount of degradation, while a value of 5 indicates a sequence with no perceptible degradation.

3. RESULTS AND DISCUSSION

Comparative tests were carried out between the proposed coding scheme and MJPEG on various test sequences. Video sequences of different resolutions were used, all at 30 fps and 24 bits per pixels.

In Table 2, coding results for test sequences Claire, Missa, Susie, Mobile and Football are presented in terms of the compression ratio that MJPEG achieved, together with the quality rating given by CQA.

Table 2 MJPEG results

Sequence	Resolution	Ratio	CQA
Claire	360 x 288	19:1	4
Missa	360 x 288	12:1	4
Susie	720 x 480	17:1	4
Mobile	720 x 576	9:1	4
Football	720 x 486	8:1	4

Figure 3 shows a frame from the Football sequence coded at three target compression ratios, using the proposed method. While there is no visible difference between the frame coded at 14:1 and 42:1, the frame coded at 70:1 is of obvious lower quality. The image background has lost much of the detail, and checkerboarding is present throughout the image.

Table 3 shows the coding results for the same sequences using the proposed method.

Table 3 Wavelet results

Sequence	Resolution	Ratio	CQA
Claire	360 x 288	20:1	4
Missa	360 x 288	15:1	4
Susie	720 x 480	22:1	4
Mobile	720 x 576	12:1	4
Football	720 x 486	12:1	4

The results show that the proposed method outperforms MJPEG in terms of compression ratio, at the same quality level. This indicates that the wavelet transform is a suitable alternative to the DCT for use in video coding.

4. CONCLUSIONS AND FUTURE WORK

This paper has presented a baseline wavelet colour video codec, which outperforms MJPEG in terms of compression ratio at the same subjective quality level.

Future research efforts will concentrate on three items. Firstly we aim to increase the compression ratio itself by improving the quantiser performance, and examining alternative scan orders for the wavelet coefficients, prior to run length coding. Secondly, we aim to investigate how different linear phase perfect reconstruction filter banks affect the compression ratio and quality of the decompressed data. Finally, work is in progress on a segmentation algorithm for locating objects within a scene. It is envisaged that this segmentation algorithm will be combined with the video codec in order to produce an increased compression ratio.

5. ACKNOWLEDGEMENTS

D. Britch is supported by the UK Engineering and Physical Sciences Research Council (EPSRC).

REFERENCES

[1] W. Li, Y. Q. Zhang, "Vector-based signal processing and quantization for image and video compression", *Proceedings of IEEE*, Vol. 83, Feb 1995, pp. 671-681.
[2] J. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients", *IEEE Trans. Signal Processing*, Vol. 41, Dec 1993, pp. 3443-3463.
[3] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, "Image coding using the wavelet transform", *IEEE Trans. Image Processing*, Vol. 1, No. 2, April 1992.
[4] M. G. Albanesi, I. de Lotto, and L. Carrioli, "Image compression by the wavelet decomposition", *Signal Processing*, Vol. 3, No. 3, June 1992.
[5] A. S. Lewis, and G. Knowles, "Image compression using the 2D wavelet transform", *IEEE Trans. Image Processing*, Vol. 1, No. 2, April 1992.
[6] Netravali, A., Haskell, B., *Digital Pictures – Representation and Compression*, Plenum Press, New York, 1989.
[7] G. K. Wallace, "The JPEG still picture compression standard", *Comm ACM*, Vol. 34, No. 4, April 1994, pp. 30-44.
[8] K. Hosaka, "A new picture quality evaluation method", *PCS'86*, Tokyo, Japan, April 1986.
[9] V. R. Algazi, Y. Kato, M. Miyahara, K. Kotani, "Comparison of image coding techniques with a picture quality scale", *Proc SPIE: Applications of Digital Image Processing*, XV: 1771, 1992.
[10] T. Morris, K. Angus, R. Butt, A. Chilton, P. Dettman, and S. McCoy, "CQA – Subjective video quality analyser", *British Machine Vision Conference*, Nottingham, UK, 1999.
[11] Strang, G., Nguyen, T, *Wavelets and Filter Banks*, Wellesley-Cambridge Press, Wellesley MA, USA, 1997.
[12] C. Brislawn, "Classification of nonexpansive symmetric extension transforms for multirate filter banks", *Applied and Computational Harmonic Analysis*, Vol. 3, March 1996, pp. 337-357.
[13] A. Cohen, I. Daubechies, and J. C. Feauveau, "Biorthogonal bases of compactly supported wavelets", *Comm. Pure Appl. Math.*, Vol. 45, 1992, pp. 485-560.
[14] J. D. Villasenor, B. Belzer, and J. Liao, "Wavelet filter evaluation for image compression", *IEEE Trans. Image Proc.*, Vol. 4, No. 8, August 1995, pp. 1053-1060.
[15] Zhang, T. Q., Zafar, S., *Video Data Compression for Multimedia Computing: Statistically Based and Biologically Inspired Techniques*, Chapter 1, ed: H. L. Hua, Kluwer Academic Publishers, 1997.

(a)

(b)

(c)

Figure 3 Comparison of the performance of the proposed method on the sequence "Football": (a) 14:1, (b) 42:1, (c) 70:1. The resolution is 720 x 486 pixels per frame.