Proc. IWSSIP, Manchester, September 2002.

OBJECT-BASED INTRA-FRAME VIDEO CODING USING THE LIFTING TRANSFORM

DR D. BRITCH
E-mail: D.Britch@mmu.ac.uk
Department of Computing and Mathematics
Manchester Metropolitan University,
Manchester, Ml 5GD, UK

DR. T. MORRIS
E-mail: T.Morris@co.umist.ac.uk
Department of Computation
UMIST
P. O. *Box* 88, *Manchester, M60 1QD, UK*

An object-based based video codec has been designed and implemented. The codec takes a colour video sequence of arbitrary size as input and performs intra-frame object-based compression on the sequence. Video sequences are segmented using an optic flow based algorithm, before being transformed with a fast lifting transform, prior to vector 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 a similar non object-based codec, at the same quality level.

1. Introduction

Object-based coding has been suggested as an alternative to conventional pixel-based coding [1], with segmentation performing the task of object identification. However, segmentation cannot be seen as a tool that is easily applied to generic scene content. Substantial gains can only be achieved if large parts of a scene contain regions that can be extracted from the remaining parts of the scene.

In recent years, the disadvantages of both subjective and objective measures for measuring decoded frame quality have motivated researchers to search for a better method of evaluating the quality of video sequences which have undergone lossy compression [2, 3]. Such a method has been produced and a codec quality analyser (CQA) has been developed that aims to provide a consistent bias-free single quality measure, which correlates with a subjective assessment of video data [4].

This paper describes an object-based video codec that we have developed. It is designed to take a colour video sequence as input, and perform intra-frame object-based compression on the sequence. Test sequences were compressed with the codec, and the results obtained indicate that the object-based codec outperforms the non object-based codec in terms of compression ratio, at the same quality level.

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

The algorithm comprises five main processes: optic flow based segmentation, lifting-based decomposition of the detected objects, vector quantisation of the lifting coefficients, zerotree coding of the quantised coefficients followed by lossless entropy coding of the zero tree data.

The encoder takes a colour video sequence with arbitrary sized dimensions as input, and codes the first frame using a non object-based approach. Horn and Schunk's optic flow algorithm [5] is then used to segment the remainder of the input sequence in order to separate the sequence into foreground and background objects. A minimum bounding box is used to obtain the area that the foreground object occupies, and then the bounding box is extended by five pixels on every side in order to store a small amount of surrounding background with the object. The extra pixels are stored for use in updating the background when the object moves. Along with the object, its location in the frame is also stored. The foreground objects in each frame are then successively decomposed by a fast lifting transform [6].

An effective method to reduce the amount of data in many compression algorithms is to use an RGB to YCbCr decorrelation transform followed by subsampling of the chrominance components [7]. However, it is possible to use the multi-resolutional nature of the lifting transform to achieve the same effect. After decorrelating the data using an RGB to YCbCr transform and then applying the lifting

transform, the HL, LH, and HH sub bands of the chrominance component's lifting decomposition are discarded and all other subbands retained. This has the effect of achieving a 2:1 subsampling in both the horizontal and vertical dimensions of the chrominance components [8].

The lifting coefficients are then quantised by a trained Lloyd-Max vector quantiser [9, 10, 11] that outputs quantised coefficients. These coefficients are subsequently entropy encoded by zerotree coding [12] and Huffman coding, producing a single bitstream of data that contains the compressed frame data, along with header information that allows the data to be decoded.

The non object-based codec is essentially identical to the object-based codec, with the exception that the non object-based codec has had the optic flow based segmentation removed.

3. Results and Discussion

Comparative tests were carried out between the proposed coding scheme and the non object-based coding scheme. Video sequences of different spatial resolutions were used, all at 30 fps and 24 bits per pixels. The quantiser was trained using a maximum of 20 iterations and a threshold of 0.01, meaning that when the distortion between iterations drops below 1 %, the training algorithm terminates.

Table 1 shows the coding results achieved for five video test sequences with both the non object-based codec and the object-based codec. The results are presented in terms of the compression ratio achieved, together with the quality rating given by CQA.

Table 1

Sequence	Non obj-based		Object-based	
	Ratio	CQA	Ratio	CQA
Claire	52:1	4	87:1	4
Missa	29:1	4	35:1	4
Susie	32:1	4	36:1	4
Mobile	25:1	4	154:1	4
Football	23:1	4	25:1	4

The results show a significant improvement over the non object-based approach. In all five video sequences the compression ratio has increased, with the quality of the decoded stream remaining the same. Figure 1 shows a frame of the Mobile sequence coded using the non object-based approach, while Figure 2 shows the same frame coded with the proposed object-based approach. The frame coded with the non object-based method is of very low quality. However, the frame coded using the proposed object-based method has no perceptible degradation.

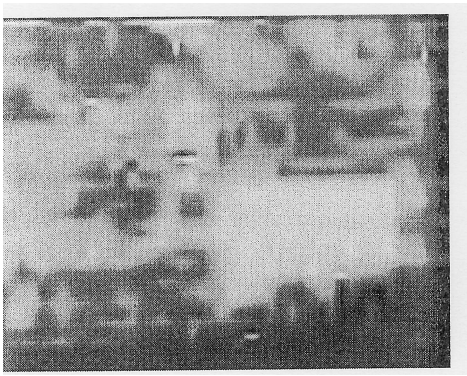

Figure 1. Mobile coded at 154: 1 (non object-based)

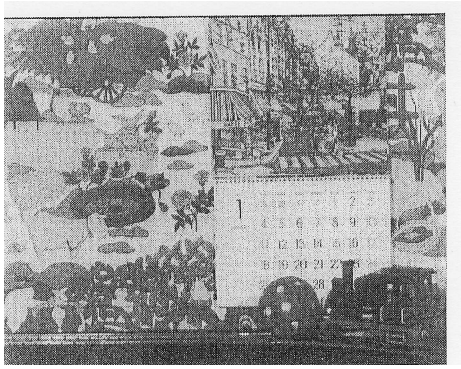

Figure 2. Mobile coded at 154: 1 (object-based)

In the Susie and Football sequences, the improvement in compression ratio is negligible, but in the remaining three sequences the results show the substantial benefits that object-based coding can produce.

4. Conclusions and Future Work

This paper has presented an object-based colour video codec using vector quantisation, which out-performs a similar non object-based codec in terms of compression ratio, at the same subjective quality level. The results show the benefits that object-based coding can achieve when the scene contains suitably sized regions that can be extracted from background of the frame.

Future research efforts will concentrate on a number of items. Firstly we aim to further investigate the quantisation strategy in order to increase the compression ratio.

Secondly, to allow comparison of this video codec with others in existence it will be necessary to modify the codec so that video sequences can be coded at user specified bit rates.

Finally, although it was possible to accurately detect the boundaries of objects located in the video test sequences, it is unlikely that this would be possible in video sequences more representative of the real world. Therefore, a substantial amount of research needs to be performed into the topic of motion segmentation.

References

1. T. Sikora, "MPEG digital video-coding standards", *IEEE Signal Processing Magazine* (September 1997).
2. K. Hosaka, "A new picture quality evaluation method", *PCS* '86, Tokyo, Japan (1986).
3. V.R. Algazi, Y. Kato, M. Miyahara, and K. Kotani, "Comparison of image coding techniques with a picture quality scale", *Proceedings of the SPIE: App. of Digital Image Proc.*, XV: 1771 (1992).
4. 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).
5. B. K. P. Horn, and B. G. Schunk, "Determining optical flow", *Artificial Intelligence,* Vol. 17, pp. 185-203 (1981).
6. W. Sweldens, "The lifting scheme: A custom-design construction of biorthogonal wavelets", *Appl. Comp. Harm. Anal.*, Vol. 3, No.2, pp. 186-200 (1996).
7. T. Q. Zhang, and S. Zafar, *Video Data Compression for Multimedia Computing: Statistically Based and Biologically Inspired Techniques,* Chapter1, ed: H. L. Hua, Kluwer Academic Publishers (1997).
8. T. Fukuhara, K. Katok, K. Hosaka, and A. Leung, "Motion-JPEG 2000 standardisation and target market", *Proc. IEEE Int. Conf Image Processing (ICIP 2000),* Vancouver, Canada, Vol. II, pp. 57-66 (2000).
9. S. P. Lloyd, "Least squares quantisation in PCM", *IEEE Trans. on Inf Theory,* IT- 28,pp.129-136 (1982).
10. J. Max, "Quantising for minimum distortion", *IEEE Trans.* on *Inf Theory,* IT-6, pp. 7-12 (1960).
11. Y. Linde, A. Buzo, and R. Gray, "An algorithm for vector quantiser design", *IEEE Trans.* on *Communications,* 26(1), pp. 84-95 (1980).
12. J. M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients", *IEEE Trans.* on *Signal Processing,* Vol. 41, pp. 3445-3462 (1993).