# ADAPTIVE SUBBAND VQ FOR VERY LOW BIT RATE VIDEO CODING

*Stathis P. Voukelatos and John J. Soraghan*
Signal Processing Division, Dept. of Electronic and Electrical Eng., University of Strathclyde,
Glasgow G1 1XW, Scotland, U.K., E-Mail: stathis@spd.eee.strath.ac.uk

## ABSTRACT

A novel adaptive VQ based subband coding scheme for very low bit rate coding of video sequences is presented. Overlapped block motion estimation/compensation is employed to exploit interframe redundancy. A 2D wavelet transform (WT) is applied to the resulting displaced frame difference (DFD) signal. The WT coefficients are encoded using an adaptive vector quantization scheme in combination with a dynamic bit allocation strategy based on marginal analysis. Simulation results on videophone-type test sequences are given to evaluate the performance of the codec at very low bit rates. A comparative performance with the H.261 and H.263 video coding standards is also shown.

## 1 INTRODUCTION

The potential applications of very low bit rate video coding in a number of forthcoming visual services has led to an increased research and standardization effort in the area. The well established ITU-T (former CCITT) H.261 [1] recommendation is aimed primarily at videophone and videoconferencing services at transmission rates of $p \times 64$ kbps. In very low bit rate environments, below 30 kbps, such as the PSTN, the performance of the H.261 coder is inadequate. The reconstructed video sequences contain annoying blocking artifacts, due to the block based DCT transform coding employed in the H.261 codec. Recently the ITU-T H.263 [2] recommendation was defined aiming at videotelephony applications over the PSTN.

The wavelet transform has been shown to be an efficient coding method for still images and video. Unlike block transform based coding it does not suffer from blocking artifacts and hence it is able to produce better subjective quality especially at low bit rates. In typical subband based video coding, a 2D wavelet transform is applied in the spatial domain and some form of motion compensated prediction is performed to exploit redundancy between adjacent frames. In [3] the multiresolution motion estimation was introduced, where motion vectors are estimated in the WT domain using a hierarchical approach. This method is also used in [4] in combination with an adaptive bit plane run length coding scheme for the quantization of the subband error signals. Alternatively, full resolution motion estimation can be performed on the original video frames, followed by a wavelet transform of the resulting DFD signal. This approach is followed in [5], where successive approximation lattice VQ is used to encode the WT coefficients and in [6], where entropy constrained scalar quantizers are employed. By applying the WT in the temporal domain as well 3D WT coding schemes can be designed [7].

In this paper we present a novel wavelet transform based very low bit rate video coding system [8]. The overlapped block motion compensation (OBMC) [9] technique is employed to exploit interframe redundancies. The OBMC increases the coding gain when a WT based coder is applied on the DFD signal. A 2D WT is applied on the resulting DFD signal and an adaptive subband vector quantization (ASBVQ) scheme has been developed for the encoding of the WT coefficients. The paper is organized as follows: In Section 2 an overview of the complete codec is given. Following that, the overlapped block motion estimation technique is illustrated and its advantages over conventional block matching are highlighted. In Section 4, the adaptive subband VQ algorithm is presented and in Section 5, simulation results and a comparative performance with the H.261 and H.263 video coding standards are given. We conclude the paper with Section 6.

## 2 OVERVIEW OF THE ASBVQ CODEC

The block diagram of the complete ASBVQ codec is given in Fig. 1. A motion compensated predictor, based on the overlapped block motion compensation technique, generates a prediction of the current frame. A 2D wavelet transform is applied on the DFD signal. Two decomposition levels producing seven subbands are used. The WT coefficients of each subband are grouped into vectors and the most significant vectors according to their energy are preserved. The output of the vector selection process is a *significance map* for each subband. The significance map is a binary image where a "1" corresponds to a vector that has been preserved and will be

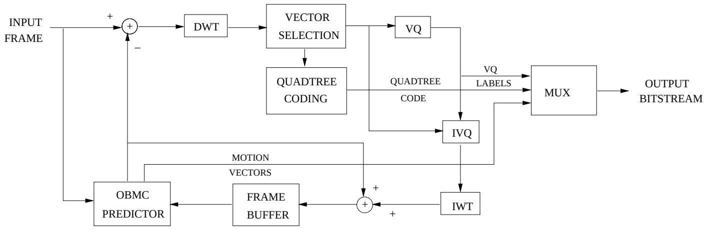


Figure 1: Adaptive subband VQ video codec

used in the VQ stage, while a "0" corresponds to a vector that has been discarded. Each of these significance maps has to be coded so that the decoder can determine the positions of the selected vectors. The encoding of each map is done using a quadtree structure. The selected coefficient vectors are coded with an adaptive VQ scheme. Fixed rate coding can be achieved by assigning a constant number of bits to each frame of the sequence.

# 3 OVERLAPPED BLOCK MOTION COMPENSATION

Block matching motion estimation/compensation is the most widely used method in interframe coding systems, as it can successfully reduce the energy of the DFD signals, which is the main objective from a coding point of view. The DFD signal produced by the conventional block matching approach will typically contain discontinuities at the boundaries of neighbouring blocks where different motion vectors have been assigned. When a block based transform, such as the DCT, is applied on the DFD signal these blocking artifacts do not compromise seriously the coding efficiency, since the edges of the motion compensation blocks usually coincide with the edges of the transform blocks. However, when a global transform, such as the DWT is employed, these discontinuities produce high frequency components. Hence, the coding gain of the transform is reduced since significant amounts of signal energy are spread in the high frequency subbands.

The overlapped block motion compensation (OBMC) [9] method reduces the blocking artifacts generated by the conventional block matching, by using overlapping blocks in the motion estimation and compensation. The operation of the OBMC is illustrated in Fig. 2. During motion estimation the current frame of the sequence is divided in blocks of size $2M \times 2M$, which overlap by $M$ pixels and the best matching block for each $2M \times 2M$ block is found from the previous decoded frame. Motion compensation is performed by shifting the best matching block of each $2M \times 2M$ overlapping block and weighting it with a smooth window function. Then all these windowed blocks are summed together to form the prediction of the current frame.

In our implementation we chose $M = 16$, resulting in blocks of size $32 \times 32$ overlapping by 16 pixels. Half-pixel accuracy is employed in the estimation of the motion vectors, since it achieves significant reduction in the energy of the DFD signal compared with one-pixel accuracy. The motion vectors are coded using DPCM to exploit the existing correlation in the motion between neighbouring blocks. The DPCM error values are encoded with variable length coding.

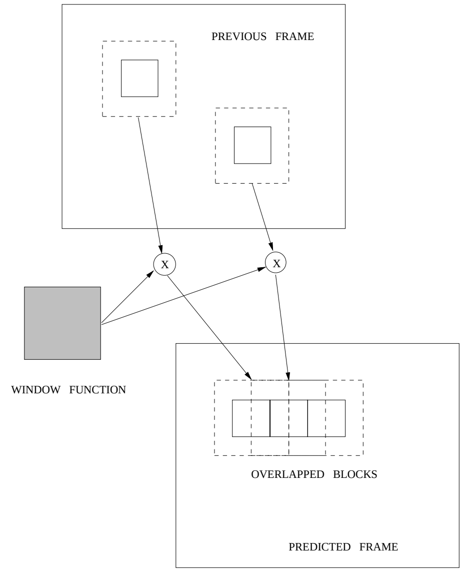


Figure 2: Overlapped block motion compensation

# 4 ADAPTIVE SUBBAND VECTOR QUANTIZATION

The wavelet coefficient vectors that have been selected from each subband are coded using an adaptive VQ scheme. A codebook is designed for each subband using the LBG [10] algorithm. Thus, no fixed codebooks based on training sets are employed. Since the subband codebooks are designed adaptively for each new frame, they are incorporated in the compressed bitstream and are transmitted to the decoder as side information. The bit allocation scheme used to determine the codebook sizes for each subband is based on the marginal analysis [11] optimization method. The available bit budget is distributed on-line, during compression time, by combining the bit allocation scheme with the splitting method for LBG codebook design. This modification of the LBG algorithm by incorporating the bit allocation operations results in an algorithm that produces a set of codebooks of various sizes. We introduce some notation to help us present the adaptive vector quantizer design procedure:

$N$ = Number of subbands.
$C_{old}(i), C_{new}(i)$ = Codebooks for subband $i$
$S_i$ = Codebook size of subband $i$.
$D_i(s)$ = Distortion of subband $i$ when quantized with codebook of size $s$.
$R_i(s)$ = Number of bits for coding subband $i$ with code-

book of size $s$.
$T$ = Target bit rate.
$B(S_1, \ldots, S_N)$ = Current bit rate.

We now present the main steps of the combined codebook design and bit allocation procedure. The iterative codebook design algorithm terminates when the target bit rate $T$ is reached.

1. For $i = 1, \ldots, N$, set $S_i = 1$ and $C_{old}(i) = \{$centroid of vectors of subband $i\}$.

2. Compute $D_i(1)$ for each $i$.

3. For $i = 1, \ldots, N$, split $C_{old}(i)$ into two vectors and optimize the codebook of size 2 using the LBG algorithm. Let $C_{new}(i)$ be the new set of codebooks.

4. Compute $D_i(2)$ for each $i$.

5. Set $S_j = 2 \times S_j$ and $C_{old}(j) = C_{new}(j)$ where:

$$j = \text{argmax}\{\frac{D_i(S_i) - D_i(2 \times S_i)}{R_i(2 \times S_i) - R_i(S_i)} \; ; \; i = 1, \ldots, N\}$$

   Split the codewords of $C_{old}(j)$ and optimize the codebook of double size using the LBG algorithm. Let $C_{new}(j)$ be the new codebook. Compute $D_j(2 \times S_j)$.

6. Calculate the current bit rate $B(S_1, \ldots, S_N)$.

7. If $B > T$ stop, else goto 5.

After the termination of the algorithm the set of codebooks $C_{old}(i)$, for $i = 1, \ldots, N$, are used to quantize the WT coefficients of each subband.

## 5 SIMULATION RESULTS

In this section we present comparative simulation results of our ASBVQ coding scheme with an H.261 and an H.263 codec. The first 150 frames of the color test sequence "Mother and Daughter" in QCIF format were used for the experiments. The frame rate of the sequence is 25 frames/sec. A study of the performance of the three coders was made for bit rates of 9.6, 14.4 and 28.8 kbps, which are the common PSTN modem rates. The average PSNR results are given in Table 1. The ASBVQ simulations start with a perfect first frame. The H.263 coder was operated using the Unrestricted Motion Vectors (Annex D) and Advanced Prediction Mode (Annex F) operating options defined in the recommendation. The desired frame rates for the ASBVQ and H.261 were obtained by temporally subsampling the original sequence. For the H.263 the frame rates given in Table 1 are an average, since the buffer regulation scheme used by the coder performs frame skipping, when the complexity of the sequence increases, resulting in variable frame rate.

| | Bit Rate (kbps) | Frame Rate (fps) | Y PSNR (db) | U PSNR (db) | V PSNR (db) |
|---|---|---|---|---|---|
| ASBVQ | 9.71 | 6.25 | 30.49 | 38.03 | 36.54 |
| H.263 | 9.65 | 5.52 | 31.23 | 37.00 | 36.88 |
| H.261 | 9.60 | 4.16 | 28.80 | 35.16 | 34.82 |
| ASBVQ | 14.37 | 8.33 | 31.11 | 38.26 | 37.21 |
| H.263 | 14.47 | 8.05 | 31.89 | 37.33 | 37.31 |
| H.261 | 14.40 | 6.25 | 29.30 | 35.33 | 34.97 |
| ASBVQ | 29.37 | 12.5 | 32.33 | 38.49 | 38.07 |
| H.263 | 28.85 | 12.76 | 33.16 | 38.17 | 38.15 |
| H.261 | 28.80 | 12.5 | 30.07 | 35.72 | 35.30 |

Table 1: PSNR performance on "Mother and Daughter"

Figure 3: Original frame

In Fig. 3 frame 96 of the original sequence is given. The reconstructed frames using the three coders at 14.4 kbps are given in Fig. 4. It can be observed that the quality of the H.261 is significantly inferior to the other two, as it suffers from very severe blocking artifacts. The visual quality of the H.263 and the ASBVQ is generally similar, with the H.263 achieving some greater fidelity in certain areas of the sequence that contain complex motion. This can be attributed to the option provided by Annex F of the recommendation of using four motion vectors instead of one for each macroblock, resulting in better modeling of complex not translational motion. Such a feature has not been incorporated yet in the ASBVQ codec.

## 6 CONCLUSIONS

In this paper a wavelet based very low bit rate video codec was presented. Overlapped motion estimation and compensation is employed since it generates a DFD signal which can be efficiently coded in the WT domain. For the quantization of the wavelet coefficients an adaptive VQ scheme combined with a bit allocation strategy based on marginal analysis has been developed. Simulation results at very low bit rates are given for standard colour sequences. A comparative performance with the H.263 and H.261 video coding standards demonstrated that our codec achieves comparable visual quality with the H.263 and significantly higher than the H.261.


### Acknowledgments

This work has been supported by the GEC-Marconi Hirst Research Centre and a University of Strathclyde studentship. The H.261 results were obtained with the PVRG-P64 Codec 1.1 developed by Andy Hung at the Portable Video Research Group, Stanford University, USA. The H.263 simulations used version 1.4a of the H.263 simulation software developed by Telenor research, Norway.


### References


[1] ITU-T Recommendation H.261, "Video Codec for Audiovisual Services at $p \times 64$ kbits/s".

[2] Draft ITU-T Recommendation H.263, "Video Coding for Low Bit Rate Communication".

[3] S. Zafar, Y. Q. Zhang and B. Jabbari, "Multiscale Video Representation Using Multiresolution Motion Compensation and Wavelet Decomposition", *IEEE Journal on Selected Areas in Communications*, vol. SAC-11, no. 1, pp. 24-35, Jan. 1993.

[4] K. H. Goh, J. J. Soraghan and T. S. Durrani, "Very Low Bit-Rate Colour Video Coding Scheme Using Subband Motion Estimation/Compensation", *Proc. EUSIPCO-94*, pp. 1214-1217, Edinburgh, Scotland, Sept. 1994.

[5] D. G. Sampson, E. A. B. da Silva and M. Ghanbari, "Low Bit-Rate Video Coding Using Wavelet Vector Quantization", *IEE Proc. Vision, Image and Signal Processing*, vol. 142, no. 3, pp. 141-148, June 1995.

[6] Y. H. Kim and J. W. Modestino, "Adaptive Entropy Coded Subband Coding of Image Sequences", *IEEE Trans. on Communications*, vol. COM-41, no. 6, pp. 975-987, June 1993.

[7] K. N. Ngan and W. L. Chooi, "Very Low Bit Rate Video Coding Using 3D Subband Approach", *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 4, no. 3, pp. 309-316, June 1994.

[8] S. P. Voukelatos and J. J. Soraghan, "Very Low Bit Rate Color Video Coding Using Adaptive Subband Vector Quantization with Dynamic Bit Allocation", to appear in *IEEE Trans. on Circuits and Systems for Video Technology.*

[9] J. Katto, J. Ohki, S. Nogaki and M. Ohta, "A Wavelet Codec with Overlapped Motion Compensation for Very Low Bit-Rate Environment", *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 4, no. 3, pp. 328-338, June 1994.


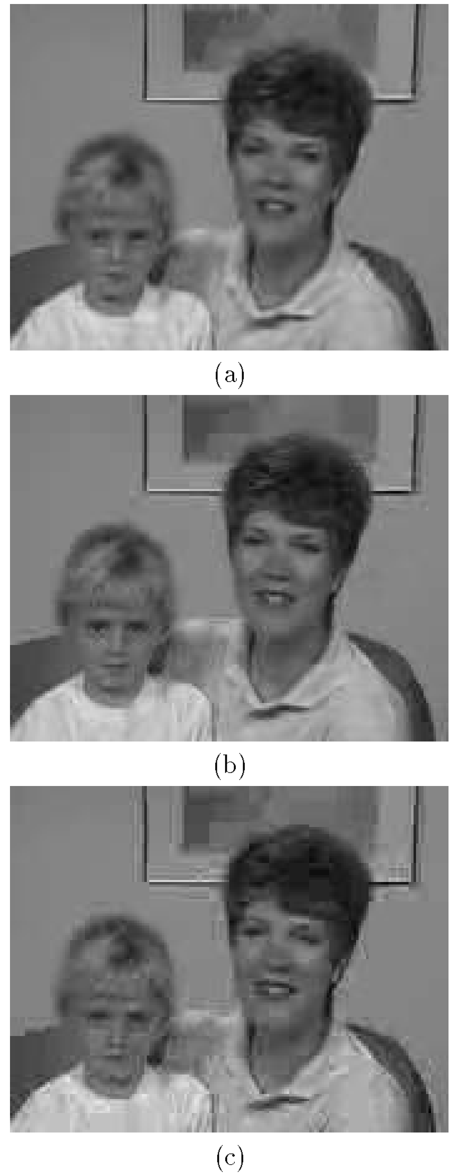


Figure 4: Reconstructed frames at 14.4 kbps (a) ASBVQ, (b) H.263, (c) H.261


[10] Y. Linde, A. Buzo and R. M. Gray, "An Algorithm for Vector Quantizer Design", *IEEE Trans. on Communications*, vol. COM-28, no. 1, pp. 84-95, Jan. 1980.

[11] B. Fox, "Discrete Optimization via Marginal Analysis", *Management Science*, vol. 13, no. 3, Nov. 1965.