3-D SUBBAND CODING OF VIDEO USING RECURSIVE FILTER BANKS

Marek Domański and Roger Świerczyński

Politechnika Poznańska, Instytut Elektroniki i Telekomunikacji, ul. Piotrowo 3a, 60-965 Poznań, Poland
Phone: +48 61 782 762, Fax: +48 61 782 572, E-mail: domanski@et.put.poznan.pl , roger@et.put.poznan.pl

Abstract

A video coding technique based on a three-dimensional subband analysis with recursive spatial filter banks is proposed. Moreover a simple technique to compress digital data in the subbands is described. In order to avoid annoying artifacts at edges and thin lines the filter banks are switched adaptively. Flat areas are processed with recursive filters exhibiting long impulse responses and good selectivity, while object edges and other detailed regions are processed with recursive filters with highly attenuated impulse responses and poorer selectivity. For very simple encoding scheme good visual quality has been obtained for real test video sequences in the CIF format encode at the bitrates about 150 kbps. Obviously further bit rate reduction could be obtained using a more sophisticated coder. The very important advantage of the technique proposed is its simplicity.

1. INTRODUCTION

Recently, rapidly emerging applications in many areas such like videophone and videoconferencing, multimedia data bases, distant teaching, personal video storage and many others have strongly stimulated research on digital video compression. Great interest has been shown for video coding for low and very low bit rate channels, i.e., for channels of about 100 kbps to 1 Mbps and under 64 kbps, respectively. Digital video transmission throughout such channels is related to a need for very high compression, even higher than achievable with hitherto commonly used techniques [1,2].

In order to avoid the problems faced by application of the hitherto popular block-based DCT techniques to the very low bit rate coding, the techniques which are free of the blocking effects are examined extensively. Among them the object-based methods [3] as well as the methods based on extensions of the well-known image subband coding (SBC) (e.g., [4,5]) are of particular interest for video coding.

The more straightforward approach to subband coding of video consists in replacing of the DCT block-based coding inside the prediction loop that implements interframe prediction with motion compensation [4,6-8]. The prediction error is encoded using subband decomposition. Disadvantage of this approach is related to the fact that the regions of significant prediction error tend to be unnecessarily extended due to the downsampling [9].

Another approach consists in application of the three-dimensional (3-D) subband coding where the three-dimensional spatio-temporal frequency band of the input video sequence is splitted up into the low- and high-frequency bands in the temporal frequency domain and then again each of these channels is splitted up into some subbands in the domain of the spatial frequencies [10-14]. The proposals combine 3-D SBC with classic motion compensated prediction [13,14] or geometrical vector quantization [11,12]. The paper [10], where promising results for 9.6 kbps have been reported, describes encoding of the channels obtained by a 3-D subband analysis using a hybrid technique where the base temporal subband is coded with a DCT block-based coder while the vector quantization is applied in the other subbands.

Promising results have been already obtained for still image data compression using subband coding and recursive filter banks [7,15,18,19]. Such filters can be implemented very efficiently and the problems resulting from the nonlinear phase characteristics can be avoided by application of reversive systems where images are processed in opposite directions in the coder and in the decoder, thus compensating for the shifts caused by the filters. Low-order recursive filters need only 0.5-1.0 multiplication per pixel to implement a filter pair [15].

Nevertheless subband image coding applications require filter banks with specific features. Selectivity is a conflicting requirement to short filter responses which are important for processing of the portions of images where edges and lines occur. There are two ways of avoiding this difficulty. The first is to use asymmetrical filter banks where analysis lowpass filter is long, while the analysis highpass is very short [4,20]. Another approach used also in this paper consist in adaptive switching of two filter banks. Flat areas are processed with recursive filters exhibiting long impulse responses and good selectivity, while object edges and other detailed regions are processed with recursive filters with highly attenuated impulse responses and poorer selectivity [20,21].

The authors have also successfully used a very simple technique for subband encoding where data in the high-frequency subbands are transmitted only for those portions of images where some changes in the low-frequency channel have been detected [16].

In this paper we are going to prove that application of this coding technique called as hierarchical quantization together with recursive switchable filter banks leads to a simple and efficient systems for digital video compression. We propose a 3-D subband coding approach based on application of spatial recursive filter banks in reversive arrangements and an original technique to reduce the bit rate in the low temporal frequency channel.

2. 3-D SUBBAND CODER

The input video sequence is analyzed by temporal, horizontal and vertical filter banks. For spatial analysis, both horizontal and vertical, in contrary to the papers [10-14], separable recursive half-band filters in polyphase implementation are used because of their simplicity [4,15-17].

Nevertheless , in the temporal analysis, we follow the suggestions of [11,13] using very simple linear-phase two-tap filters

$$H(z) = 0.5(1 \pm z^{-1}), \quad (1)$$

where "+" and "-" correspond to the low- and high-pass filters, respectively. This filter bank has a very simple implementation and exhibits small group delay (half of sampling period) resulting in small system response times which are very critical for videophone and videoconferencing applications. It enables perfect reconstruction in the very practical DFT arrangement.

Our approach to encode channel information is very simple (see Fig.1). We exploit lower spatial resolution of the human visual system for the high temporal frequency channel. Therefore the high temporal frequency subbands are decimated, i.e., only the spatial subbands 0 - 6 are encoded and transmitted. However, four spatial subbands are encoded and transmitted in the low temporal frequency subband. Encoding of the subband 0 (low spatial frequencies) from the high temporal frequency channel as well as the subbands 1-9 from the low temporal frequency channel is controlled by the signal DL.

The signal DL is a difference of the two consecutive frames from the subband 0 in the low temporal frequency channel, i.e., in order to avoid the "dirty window" effect, a frame of the DL signal is created from the information originating from the four consecutive frames (see Fig. 1).

A pixel is active if the respective value of DL is positive. Only active pixels from the other channels are transmitted. The other pixels from these channels are reconstructed in the decoder from the previous frames. In contrary to these four channels, the subband 0 from the low temporal frequency subband is reconstructed by differential decoding, that is by adding the samples of the signal DL to the samples of the previous frame. At the output of the decoder, Hufmann coding is applied to all the signals.

The compression gain can be increased by increasing the number of subbands in the spatial domain. Figure 1 presents a system where the 2-D spatial frequency band is decomposed into 10 subbands. There are three levels of decomposition. The signal is decomposed into 4 subimages at each level using a separable filter bank consisting of horizontal and vertical two-band filter banks.

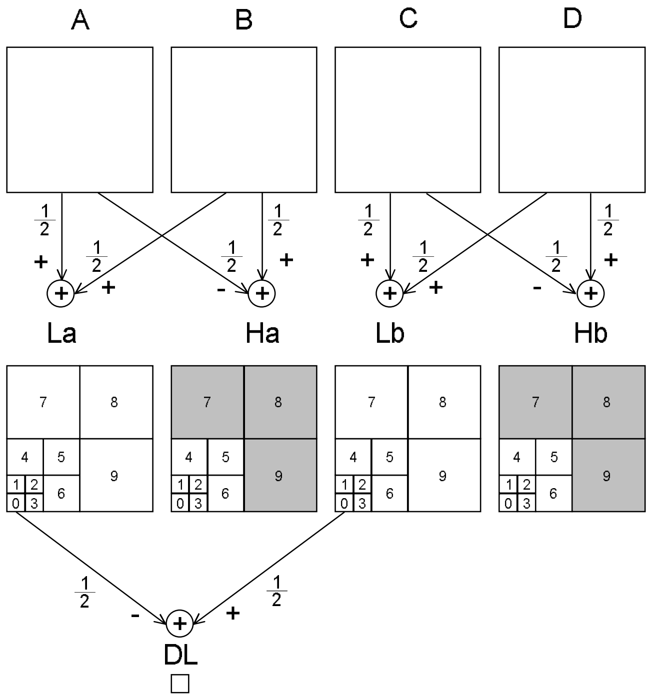

Fig. 1. Channel encoding principle

3. FILTER BANK SWITCHING

In our experiments, we use switchable filter banks as proposed in [21]. We use simple criteria for selecting an appropriate filter. First, the frame L (low temporal frequency) is divided into non-overlapping blocks of size 16x16. All border blocks are set to be filtered with the short filter (i.e. the filter with a quickly decaying impulse response). For each other block, the normalized variance is calculated. If variance is larger then a predefined threshold - the block is also set to be filtered with the short filter, otherwise we choose the long filter (the same in both directions). Analysis is performed in the following way (for lines):

$$O_i[0] = a_{i0}[0] \cdot in_i[1] + in_i[0];$$
$$for \;\; n \in <1, MaxX-2>$$
$$O_i[n] = a_{i0}[n] \cdot (in_i[n+1] - O_i[n-1]) + in_i[n];$$
$$O_i[MaxX-1] = in_i[MaxX-1];$$

where:

- $in_i[n]$ - input to the i-th allpas section;
- $O_i[n]$ - output of the i-th allpas section;
- $a_{i0}[n]$ - switchable filter coefficient;
- $MaxX$ - line size;

This procedure is repeated two times for the obtained low-frequency subband. Then we code obtained subbands (0-3) the method as just described. Synthesis is performed in the receiver as follows (for lines):

$$O_i[0] = r_i[0];$$
$$set \;\; r_i[MaxX] = 0;$$
$$for \;\; n \in <1, MaxX-1>;$$
$$O_i[n] = a_{i0}[MaxX-1-n] \cdot (r_i[n+1] - O_i[n-1]) + r_i[n];$$

where:

- $r_i[n]$ - time-reversed input to the i-th allpas section;
- $O_i[n]$ - time-reversed output of the i-th allpas section;
- $a_{i0}[n]$ - switchable filter coefficient;
- $MaxX$ - line size.

An example of application of switchable filters is presented below (Figures 2-5 at the end of this paper). This is a frame from standard sequence Salesman. Assume that there was no motion so far (it means that no high-frequency information has been sent yet). One can see that short filter produces sharper image, but suffers from blocking effects. Otherwise, long filter exhibits strong ringing noise around sharp edges. Switching filters, i.e. application of the long filter for flat regions and non-sharp edges and the short filter for sharp edges, yields better visual quality. Moreover, the high-frequency information (spatial) is transmitted only in the regions with high motion.

Of course, the motion detector used is mostly sensitive to moving sharp edges. This implies that the spatial high-frequency information is predominantly transmitted for moving edges. For better performance, it is wise to use the filter with short impulse response in such a region.

4. EXPERIMENTAL RESULTS

The technique has been examined using standard videophone test sequences like "Salesman" and "Miss of America". The input luminance sequences are in the CIF format, i.e., 288 lines and 356 columns with 10 pps.

For output sequences with good subjective opinion we obtained bitrates of 145 kbps for the sequence "Miss of America" and 167 kbps for "Salesman".

5. CONCLUSIONS

A very fast and efficient technique has been proposed. High processing speed has been obtained by using recursive filters for spatial analysis as well as a simple subband encoding method.

REFERENCES

[1] IEEE Transactions on Circuits and Systems on Video Technology, Special issue on very low bit rate video coding, June 1994.

[2] Proceedings of International Workshop on Coding Techniques for Very Low Bit-rate Video, Colchester UK, 1994.

[3] H.G. Musmann, Object-based analysis-synthesis coding, Proc. Int. Symp.Circuits Syst., IEEE 1994.

[4] J.W. Woods (ed.), Subband image coding, Kluwer 1991.

[5] M.Domański, R.Świerczyński, Subband coding in the YUV space. XVIIth Nat. Conf. Circuit Theory and Electronic Networks, Wrocław - Polanica Zdrój, 1994, pp.2211-226.

[6] P.H.Westerink, Subband coding of images, PhD thesis, Delft Univ. of Technology, 1989.

[7] J.H. Husoy, Subband coding of still images and video, PhD thesis, Univ.Trondheim, 1991.

[8] D. Qian, A motion compensated subband coder foe very low bit rates, to be published in Image Communications, 1995.

[9] H.G. Musmann, private communication, 1995.

[10] K.Ngan, W.Chooi, Very low bit rate video coding using 3D subband approach, IEEE Trans. Circ. Syst. Video Techn., vol. 4, 1994, pp.309-316.

[11] C.Podilchuk et al., Three-dimensional subband coding of video, IEEE Trans. Image Proc., vol.4, 1995, pp.125-139.

[12] C.Podilchuk, Low bit rate subband video coding. Proc. IEEE Int. Conf. Image Proc., Austin TX, 1994, pp. III 280-284.

[13] J.Ohm, Three-dimensional subband coding with mo-

tion compensation, IEEE Trans. Image Proc., vol.3, 1994, pp.559-571.
[14] D.Taubman, A.Zakhor, Multirate 3-D subband coding of video, IEEE Trans. Image Proc., vol.3, 1994, pp.572-588.
[15] M.Bleja, M.Domański, Image data compression using subband coding, Annales des Telecommunications, vol. 45, 1990, pp.477-486.
[16] M.Domański, R.Świerczyński, Subband coding of images using hierarchical quantization, Signal Processing VII: Theories and Applications, EUSIPCO 1994, pp. 1218-1221.
[17] M.Domański, R.Świerczyński, Design of nonlinear-phase filter banks for subband coding of images. Proc. IEEE Int. Conf. Image Proc., Austin TX, 1994, pp.893-897.
[18] T.Ramstad, IIR filterbank for subband coding of images, Proc. IEEE 1982 Int. Symp. on Circ. and Syst., Helsinki 1988, pp 827-830.
[19] M. Domański, Cyfrowe filtry rekursywne dla subzakresowego kodowania obrazów, Krajowe Sympozjum Telekomunikacji, Bydgoszcz 1988, pp. 254-262, in Polish.
[20] O.Egger, W.Li, Subband coding of images using asymmetrical filter banks, IEEE Trans. Image Proc., 4/1995, pp. 478-485.
[21] M.Smith, W.Chung, Recursive time-varying filter banks for subband image coding, IEEE Trans. Image Proc., vol 4, 1995, pp. 885-895.

ACKNOWLEDGMENTS
This work has been supported by the Polish government grant KBN 8 S504 002 06.
Roger Świerczyński is a fellow of the Foundation for Polish Science .

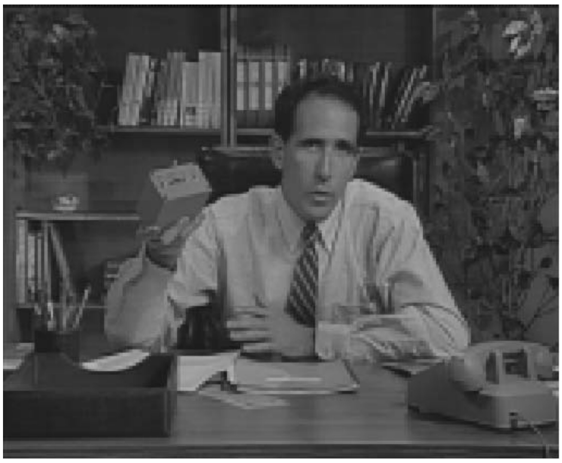

Fig. 2. Frame coded/decoded using short filter.

Fig. 4. Filter selecting map - black blocks are coded with short filter - others with long one.

Fig. 3. Frame coded/decoded using long filter.

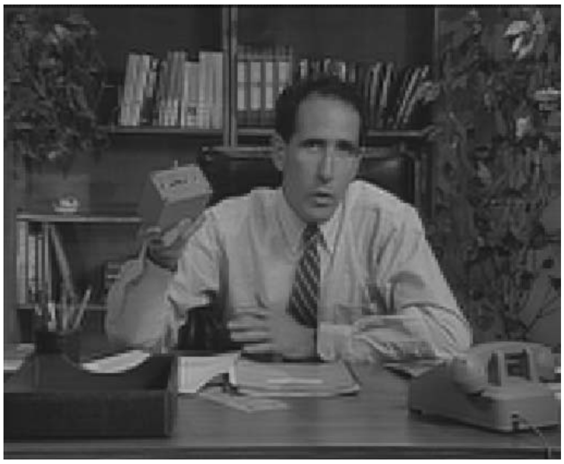

Fig. 5. Frame coded/decoded using switchable filter.