SIDE INFORMATION INTERPOLATION WITH SUB-PEL MOTION COMPENSATION FOR WYNER-ZIV DECODER

Sven Klomp, Yuri Vatis, Jörn Ostermann
Institut für Informationsverarbeitung
Universität Hannover, Appelstr. 9A, 30167 Hannover, Germany

Keywords: Distributed video coding, Wyner-Ziv, Slepian-Wolf, motion-compensated temporal interpolation, sub-pel motion estimation, side information.

Abstract: Using Distributed Video Coding (DVC), the complex task of exploiting the source statistics can be moved from the encoder to the decoder. Such a DVC decoder needs side information to exploit the statistics. In common DVC codecs, the side information is obtained by interpolating the current frame from already decoded frames. This paper proposes an interpolation technique for the side information that uses motion compensation with sub-pel accuracy, and compares different interpolation filters for calculating the sub-pel values. Using a six tab Wiener filter, we observe a gain of up to 1.8 dB for the DVC coded frames.

1 INTRODUCTION

Current video coding solutions, such as MPEG or ITU-T H.26x standards, perform well for broadcasting, streaming and other applications, wherein a video is encoded once and decoded several times. The encoder of such a solution exploits the source statistics, whereby the decoder can be kept very simple. For opposite scenarios with many encoders, Distributed Video Coding (DVC) might be more suitable than conventional video coding since the decoder performs the complex task of exploiting the source statistics.

DVC is based on the Slepian-Wolf (Slepian and Wolf, 1973) and Wyner-Ziv (Wyner and Ziv, 1976) theorems. These theorems state that it is possible to compress two statistically dependent signals in a distributed way (separate encoding, jointly decoding) using a rate equal to that used in a system where the signals are encoded and decoded together. Current approaches mostly implement the unsymmetrical case, where the two signals are coded with different bitrates.

A general block diagram of an unsymmetrical Wyner-Ziv (WZ) codec is shown in Figure 1. At the encoder, the sequence is divided into key frames and Wyner-Ziv frames controlled by the group-of-picture (GOP) size (e.g. at GOP size 4 every fourth frame is coded as key frame). The key frames are coded with a conventional intra frame coder (e.g. H.264)

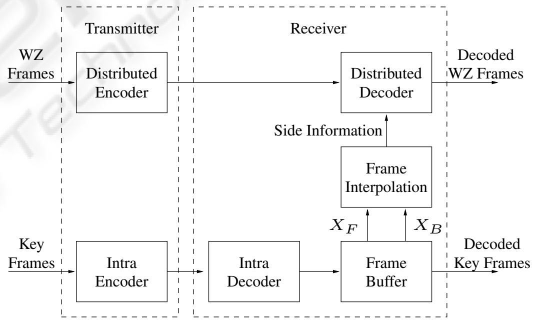

Figure 1: DVC Architecture.

and the Wyner-Ziv frames with a distributed coder. This paper deals with the frame interpolation which is independent of the distributed decoder. Therefore, the exact approach on how the WZ bitstream is generated is not of importance here. Detailed information about different distributed coder can be found in (Girod et al., 2005) and (Puri and Ramchandran, 2002). Our results are based on the transform domain Wyner-Ziv coder proposed in (Brites et al., 2006). The frame interpolation block uses the previous (X_F) and next (X_B) key frame to estimate the current WZ frame. This estimation result is called side information and is required by the distributed decoder as a base for the decoding process. Since the side information is only an estimation of the WZ frame, the distributed decoder uses the WZ bits to correct errors.

Klomp S., Vatis Y. and Ostermann J. (2006).
SIDE INFORMATION INTERPOLATION WITH SUB-PEL MOTION COMPENSATION FOR WYNER-ZIV DECODER.
In *Proceedings of the International Conference on Signal Processing and Multimedia Applications*, pages 178-182
DOI: 10.5220/0001569201780182

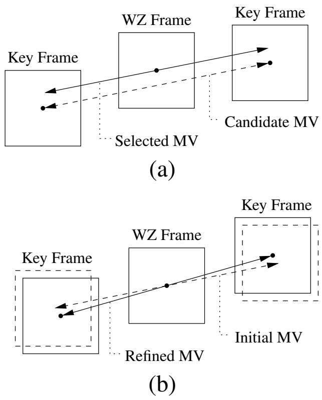

Figure 2: Motion Vector: a) Selection; b) Refinement.

Therefore, the rate of the WZ bitstream is directly affected by the quality of the side information.

Current approaches use motion-compensated temporal interpolation (MCTI) (Ascenso et al., 2005) to calculate the side information. In this paper, an improved MCTI based on sub-pel motion estimation is proposed. The new techniques for the temporal interpolation are introduced in Section 2. In Section 3, the results obtained with these techniques are presented. This paper finishes with conclusions in Section 4.

2 MOTION-COMPENSATED TEMPORAL INTERPOLATION WITH SUB-PEL ACCURACY

The frame interpolation block in Figure 1 uses MCTI for calculating the side information. A full search block matching algorithm estimates the motion vectors between the previous (X_F) and the next (X_B) key frame with full-pel accuracy. Since this vector field will result in overlapped and uncovered areas after the frame interpolation, the motion estimation scheme proposed by (Ascenso et al., 2005) is used:

For each 16x16 block of the interpolation frame, a vector is selected from the previously estimated candidates that intercepts the interpolation frame closest to the centre (Figure 2(a)). This motion vector is used as initial value for the bidirectional motion estimation where the motion is refined for a smaller search range, but with sub-pel accuracy. Since linear motion is assumed between the key frames, the forward and backward motion vector are symmetrical (Figure 2(b)). In the last step, the motion vector field is smoothed by using weighted vector median filters (Alparone et al., 1996). The WVM filter compares the motion vectors of neighbouring blocks to detect outliers and to reduce the number of false motion vectors.

In order to estimate and compensate sub-pel motion vectors, the key frames have to be interpolated at these interim values.

Key Frame Interpolation

For the interpolation of the pixel values at half-pel positions, a six tap Wiener filter as defined in H.264 (Richardson, 2003) is used. The filter coefficients are defined as

$$(1, -5, 20, 20, -5, 2)\,/32. \qquad (1)$$

All half-pel values that are horizontally or vertically adjacent to integer positions are interpolated with this filter. Then, the remaining values can be interpolated with the already calculated samples, using the same Wiener filter.

If higher precision motion vectors are required, more sub-pel positions have to be calculated. In H.264, quarter-pel samples are obtained by using a bilinear filter applied at the already calculated half-pel positions and existing full-pel positions. The remaining quarter-pel positions without neighbouring half-pel or full-pel positions are obtained like the half-pel positions by using already calculated samples.

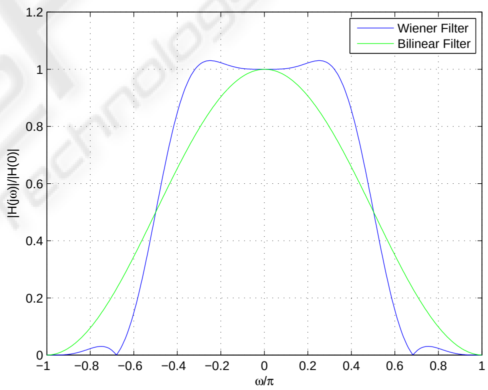

Figure 3: Frequency response of a Wiener filter and a bilinear filter.

Evaluations of the bilinear interpolated quarter-pel samples have shown that the resulting motion vectors are not suitable for motion compensated interpolation. The frequency response of the bilinear filter (Figure 3) indicates, that the signal is distorted at low frequencies. In addition, aliasing in the original frame is not accurately suppressed by the filter. In contrast, the Wiener filter, with its precipitous sides in the frequency response, is more suitable for interpolation. Therefore, the interpolation can be improved by using the Wiener filter for the quarter-pel values, too. Knowing the impulse response of the common filter at particular positions, the impulse response at other positions can be computed by shifting the impulse response (Vatis et al., 2005). This process is depicted in Figure 4: The six tap Wiener filter is interpolated with a spline function. This function is shifted by $1/4$ pel and thereafter again scanned at full-pel positions. After quantisation of the impulse response, the new filter coefficients for quarter-pel positions are

$$(5, -18, 114, 37, -11, 1)/128. \quad (2)$$

This filter is no longer symmetrical and is designed for quarter-pel positions with a neighbouring full-pel position on the left side. If the left neighbour is a half-pel sample, the filter has to be mirrored. Likewise, the remaining quarter-pel positions are calculated using already calculated quarter-pel samples and the shifted Wiener filter.

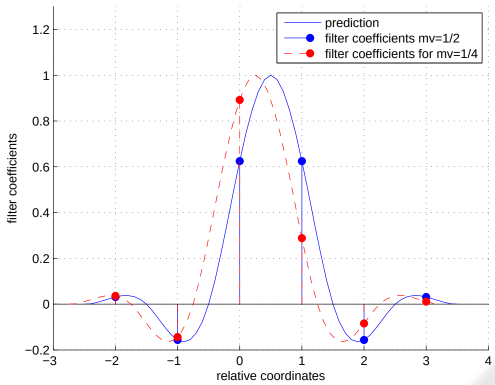

Figure 4: Prediction of the impulse response of a 6-tap 1D Wiener filter at quarter-pel positions from the impulse response at half-pel positions.

3 EXPERIMENTAL RESULTS

For the evaluation of the rate-distortion performance, four side information interpolation methods are considered: i) full-pel MCTI; ii) half-pel MCTI; iii) quarter-pel MCTI with bilinear filter and iv) quarter-pel MCTI with Wiener filter. The performance of the H.264 intra frame coder is also considered for comparison. A transform domain Wyner-Ziv coder, as proposed in (Brites et al., 2006), is used as distributed coder.

All sequences have a frame rate of 15fps, except for the CIF sequence Concrete (Figure 8), where the frame rate is 30fps. To get similar PSNR for key and WZ frames, the quantisation parameter of the H.264 coder is adjusted for each rate-distortion point.

Performance gains of up to 0.55 dB are achieved with half-pel motion compensation for the flower sequence (Figure 5). Quarter-pel MC with Wiener filter actually yields up to 0.75 dB. In case of MCTI with bilinear interpolation, quarter-pel motion compensation produces results worse than half-pel motion compensation due to the distorting characteristic of the filter. The gain of sub-pel motion compensation decreases slightly for lower bitrates, since the key frames are more distorted and thus lack details required for accurate motion estimation.

For City (Figure 6), the performance is increased by 0.6 dB and 0.75 dB for half-pel MC and quarter-pel MC with Wiener filter, respectively. Quarter-pel interpolation with bilinear filter does not improve over half-pel MC.

The results for the Foreman sequence (Figure 7), with up to 0.35 dB gain for half-pel MC, are well below the other sequences. The same applies for quarter-pel MC with Wiener filter with a gain of up to 0.45 dB. Pure intra frame coding using H.264 outperforms the WZ approach already at low bitrates, since the intra prediction modes of the H.264 codec work very well for the this sequence.

Figure 8 shows that for CIF sequences, the performance can also be increased with sub-pel motion compensation. Half-pel MC and quarter-pel MC with Wiener filter yield of up to 1.4 dB and 1.8 dB, respectively.

As mentioned in Section 1, this approach affects only the side information and therefore useful for the most WZ codecs. Examinations of a pixel based WZ codec (Girod et al., 2005) have shown that the performance gain for the sub-pel MC is almost the same. These results are not further investigated, since the overall performance of the pixel domain codec is lower than the performance of a transform domain codec.

In contrast to the results in (Li and Delp, 2005), where only the side information is examined and not the overall performance after WZ decoding, our results point out, that motion-compensated temporal interpolation of side information is significantly improved by using motion estimation with half-pel accuracy. By using a Wiener filter instead of a bilinear filter for interpolating quarter-pel samples, the performance is increased further.

4 CONCLUSIONS

In this paper, the advantage of sub-pel motion-compensated temporal interpolation is investigated and compared with common full-pel MCTI in the context of Distributed Video Coding. For interpolating the sub-pel positions, the H.264 technique is used and improved by a six tab Wiener filter for quarter-pel samples. Compared to the full-pel MCTI, these modifications achieve coding gains of up to 0.75 dB or up to 20% WZ bitrate reduction for QCIF and up to 1.8 dB or up to 50% WZ bitrate reduction for CIF.

ACKNOWLEDGEMENTS

The work presented was developed within DISCOVER (Distributed Coding for Video Services), an European Commission Future and Emerging Technologies (FET) project, funded under the European Commission IST FP6 programme. Further information can be found at www.discoverdvc.org. The DISCOVER software started from the IST WZ software developed at the Image Group from Instituto Superior Técnico (IST) of Lisbon by Catarina Brites, João Ascenso and Fernando Pereira.

REFERENCES

Alparone, L., Barni, M., Bartolini, F., and Cappellini, V. (1996). Adaptive weighted vector-median filters for motion fields smoothing. In *IEEE ICASSP*, Georgia, USA.

Ascenso, J., Brites, C., and Pereira, F. (2005). Improving frame interpolation with spatial motion smoothing for pixel domain distributed video coding. In *5th EURASIP*, Slovak Republic.

Brites, C., Ascenso, J., and Pereira, F. (2006). Improving transform domain wyner-ziv video coding performance. In *IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, Toulouse, France.

Girod, B., Aaron, A., Rane, S., and Rebollo-Monedero, D. (2005). Distributed video coding. *Proc. of the IEEE*, 93(1):pp. 71–83.

Li, Z. and Delp, E. J. (2005). Wyner-ziv video side estimator: Conventional motion search methods revisited. In *IEEE Int. Conf. on Image Processing*, pages pp. 825–828, Genova, Italy.

Puri, R. and Ramchandran, K. (2002). Prism: A new robust video coding architecture based on distributed compression principles. In *Proc. of 40th Allerton Conf. on Comm., Control and Computing*, Allerton, IL.

Richardson, I. (2003). *H.264 and MPEG-4 Video Compression*, chapter 6.4.5.2. John Wiley & Sons Ltd., West Sussex, England.

Slepian, J. and Wolf, J. (1973). Noiseless coding of correlated information sources. *IEEE Trans. on Inform. Theory*, 19(4):pp. 471–480.

Vatis, Y., Edler, B., Wassermann, I., Nguyen, D., and Ostermann, J. (2005). Coding of coefficients of two-dimensional non-separable adaptive wiener interpolation filter. In *VCIP 2005*, Beijing, China.

Wyner, A. and Ziv, J. (1976). The rate-distortion function for source coding with side information at the decoder. *IEEE Trans. on Inform. Theory*, 22(1):pp. 1–10.

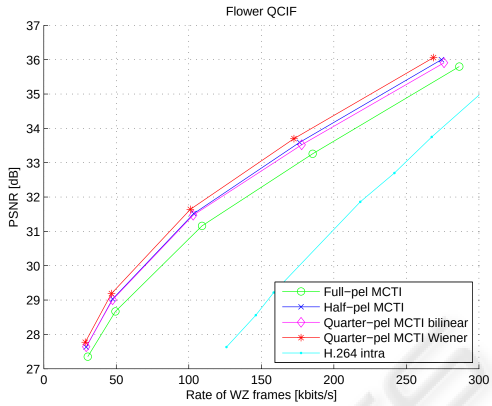

Figure 5: RD performance for the Flower QCIF Sequence coded with GOP size 2 (125 frames).

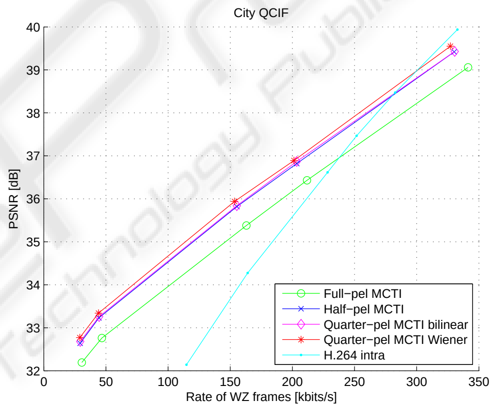

Figure 6: RD performance for the City QCIF Sequence coded with GOP size 2 (150 frames).

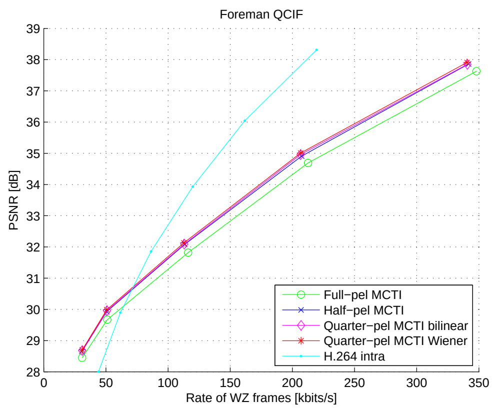

Figure 7: RD performance for the Foreman QCIF Sequence coded with GOP size 2 (150 frames).

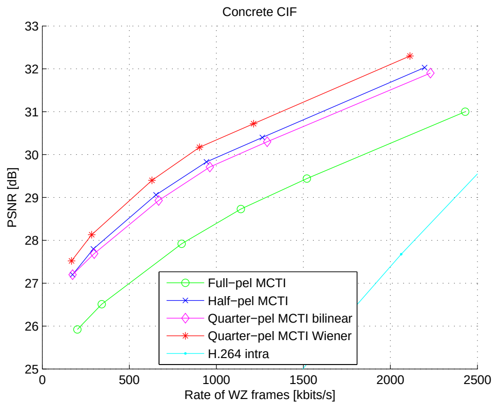

Figure 8: RD performance for the Concrete CIF Sequence coded with GOP size 2 (250 frames).