Wavelet-Based Video Coding Using Image Warping and Overlapped Block Motion Compensation

Detlev Marpe, Guido Heising, Hans L. Cycon and Alexander P. Petukhov

Abstract— **This paper describes a video coding algorithm that combines new ideas in motion estimation, wavelet filter design, and wavelet-based coding techniques. A motion compensation technique using image warping and overlapped block motion compensation (OBMC) is employed to reduce temporal redundancies in a given image sequence. This combined motion model has the advantage of representing more complex motion than simple block matching schemes. To further improve the quality of the temporal prediction, an adaptive high-quality texture interpolation is used. Spatial decorrelation of the motion compensated residual images is performed using an one-parametric family of biorthogonal IIR wavelet filters coupled with a highly efficient pre-coding scheme. Experimental results demonstrate substantial improvements in objective quality of 1.0–2.2 dB PSNR compared to the H.263+ standard. Especially at very low bit-rates, where the reconstruction quality of block-based coders suffers from visually annoying blocking artifacts the proposed coding scheme produces a superior subjective quality.**

Keywords— **Very low bit-rate video coding, motion compensation, image and video compression, wavelet transform**

I. Introduction

CONVENTIONAL methods of hybrid video coding like those implemented in the low bit-rate video coding standard H.263 [1] consist of a motion compensation/estimation and a transform coding part which both rely on block-based techniques. These techniques have the well-known drawback of producing blocking artifacts especially at very low bit-rates. In this work, we propose a video coding scheme which combines very effective non block-based or overlapping block-based techniques in order to improve coding efficiency and visual quality.

The temporal predictor of our proposed hybrid coding scheme utilizes two different motion models: a *bilinear image warping motion model* using a regular grid of squares [6], [16] and an *overlapped block motion compensation* (OBMC) with cosine weighted window functions [14]. Although the image warping model induces a subdivision into square blocks, the obtained motion vector field is smooth especially across the block boundaries. However, in the presence of objects performing different movements, motion discontinuities along object boundaries are not well represented in the image warping model. In this case, OBMC offers an instrument for a better prediction by means of a superposition of overlapping displaced blocks from the reference frame, each weighted by a smooth cosine window. Thus, we propose an adaptive block-wise switching between both motion models by minimizing a given costfunction.

Prediction residual frames (interframes) as well as intraframes are decorrelated using a discrete wavelet transform (DWT) which is realized by an appropriately designed pair of infinite impulse response (IIR) filters. In contrast to biorthogonal wavelets with compact support [4] most frequently used in the image coding community, biorthogonal bases generated by recursive filters [3] offer significantly more flexibility for the adaptation to specific coding problems. To further enhance the transform coding scheme, an additional pre-coding part is introduced where the quantized wavelet coefficients are pre-processed prior to entropy coding. This pre-coder is based on the concepts of partitioning, aggregation and conditional coding (PACC), a framework, which was first introduced in [10] and which has proven to be very efficient in still image compression and very low bit-rate video coding [11], [12].

The organization of the paper is as follows. Section II describes the motion models, the estimation of the model parameters and the motion compensation algorithm. In Section III, we discuss the employed wavelet transform together with the PACC pre-coding framework and some related pre-coding methods. We present experimental results of our proposed video coding scheme in comparison with the H.263+ coder in Section IV. The conclusion can be found in Section V.

II. Temporal prediction scheme

A. Motion Model

Two different motion models are employed in order to increase the quality of the predicted frames. A bilinear geometric transform [16] is used for warping prediction and cosine weighted windows (Fig. 3) are utilized for OBMC [14]. According to Fig. 1, the current frame k is subdivided into squares of 16 by 16 pels, thus obtaining a regular grid. In contrast to H.263 the positions of the motion vectors are on the vertices of the squares which leads to an additional row and column of motion vectors. The estimated vectors are used for warping predicition, OBMC or both (in case of neighboring blocks utilizing different motion models).

A dense motion vector field is achieved by using a bilinear geometric transform (warp) which smoothly varies over the image. In addition to translational motion the model can describe rotation, shear and change in scale. The 8 parameters of the bilinear transform can be described by the motion vectors of 4 points. Therefore the motion vectors assigned to the four vertices of a square are used. The

D. Marpe and G. Heising are with the Image Processing Department, Heinrich-Hertz-Institute (HHI) for Communication Technology, Berlin, Germany. H. L. Cycon is with the Fachhochschule für Technik und Wirtschaft, Berlin, Germany. A. P. Petukhov is with the Dept. of Mathematics, St. Petersburg State Technical University, St. Petersburg, Russia.

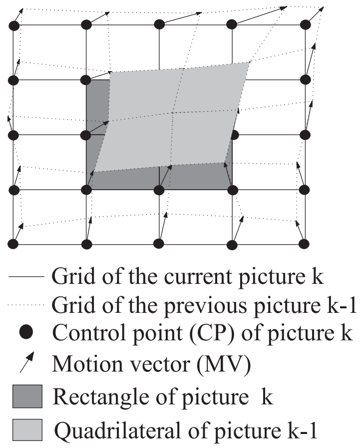

Fig. 1. Principle of image warping prediction.

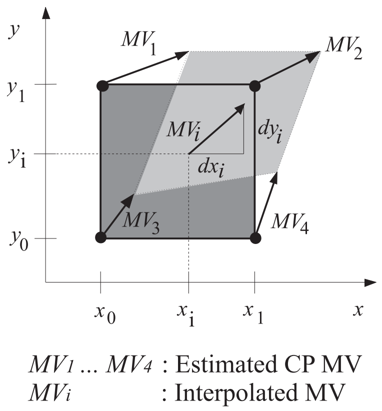

Fig. 2. Spatial interpolation of motion vectors.

motion vector MV_i inside a square block is interpolated from the four surrounding control point motion vectors $MV_1 \dots MV_4$ (*cf.* Fig. 2) using the following equation:

$$
\begin{aligned}
MV_i &= \begin{pmatrix} dx_1 \\ dy_1 \end{pmatrix} \cdot (\tilde{y}_i - \tilde{x}_i \cdot \tilde{y}_i) + \begin{pmatrix} dx_2 \\ dy_2 \end{pmatrix} \cdot (\tilde{x}_i \cdot \tilde{y}_i) \\
&+ \begin{pmatrix} dx_3 \\ dy_3 \end{pmatrix} \cdot (1 - \tilde{x}_i - \tilde{y}_i + \tilde{x}_i \cdot \tilde{y}_i) \qquad (1) \\
&+ \begin{pmatrix} dx_4 \\ dy_4 \end{pmatrix} \cdot (\tilde{x}_i - \tilde{x}_i \cdot \tilde{y}_i)
\end{aligned}
$$

where

$$\tilde{x}_i = \frac{x_i - x_0}{x_1 - x_0}, \quad \tilde{y}_i = \frac{y_i - y_0}{y_1 - y_0}.$$

This kind of warping prediction, also known as control grid interpolation [13], leads to a motion field without motion discontinuities. The smoothness across border lines is achieved by employing the same two motion vectors on both sides of the border, *e.g.* for $x_0 - x_i \approx \pm 0$ according to Eq. (1) only MV_1 and MV_3 contribute to MV_i. Hereby blocking artifact free predictions are obtained, which is essential for the performance of the proposed loop filter described in Section II-C and the wavelet transform of Section III.

However, in the case of sequences with differently moving objects, the warping model is not able to deal with motion discontinuities at the object borders. In these blocks, overlapped block motion compensation is employed by superimposing 4 predicted values $p_1, \dots, p_4$ using non linear weighting functions $w_1, \dots, w_4$:

$$
\begin{aligned}
p(x_i, y_i, k) &= \sum_{j=1}^{4} p_j(x_i, y_i, k) \cdot w_j(x_i, y_i) \\
&\stackrel{\text{def}}{=} p_1 \cdot (\hat{y}_i - \hat{x}_i \cdot \hat{y}_i) + p_2 \cdot (\hat{x}_i \cdot \hat{y}_i) \qquad (2) \\
&+ p_3 \cdot (1 - \hat{x}_i - \hat{y}_i + \hat{x}_i \cdot \hat{y}_i) \\
&+ p_4 \cdot (\hat{x}_i - \hat{x}_i \cdot \hat{y}_i),
\end{aligned}
$$

where

$$\hat{x}_i = \frac{1}{2}\left(1 - \cos\left(\pi \cdot \frac{x_i - x_0}{x_1 - x_0}\right)\right), \qquad (3)$$

$$\hat{y}_i = \frac{1}{2}\left(1 - \cos\left(\pi \cdot \frac{y_i - y_0}{y_1 - y_0}\right)\right) \qquad (4)$$

and

$$p_j(x_i, y_i, k) = p_j(x_i + dx_j, y_i + dy_j, k-1), \quad j = 1, \dots, 4.$$

Thus, the four predicted values $p_1, \dots, p_4$ are computed by employing the translational motion model with one of the motion vectors $MV_1, \dots, MV_4$ of the four surrounding vertices for each prediction. For each block, one bit is transmitted indicating the used prediction type.

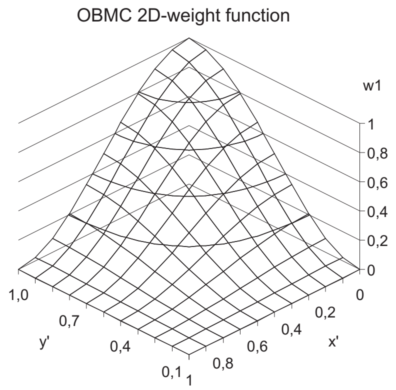

Fig. 3. 2D weighting function for OBMC-based prediction.

By combining these two compensation schemes, namely warping prediction and OBMC, no blocking is introduced *a)* between adjacent blocks using the same prediction scheme and *b)* around vertices of neighboring blocks using different schemes. The latter is a result of using, on the one hand, the same vertex motion vectors for both schemes. On the other hand, both schemes converge to the same predicition in the vicinity of a vertex. For example, let us assume that inside the square of Fig. 2 OBMC is used and outside warping prediction is employed. In that case, close to the upper left vertex warping prediction relies only on the motion vector MV_1 while for OBMC only the predicted value p_1 of the respective vertex is used, which leads to equal predictions. Only little blocking artifacts may be introduced around the center of the border line of two adjacent blocks using two different prediction schemes.

B. Motion Estimation

For motion estimation the current picture is subdivided into squares, thus creating a control grid. For every control point (vertex of a square), a motion vector is estimated by *octagonal matching*, *i.e.* finding the minimum mean-squared error (MSE) between the predicted and the current four surrounding squares of a control point (Fig. 1).

In doing so the motion vectors of the eight neighboring control points are fixed and only the motion vector of the center control point is changed. Because of their interdependency, the motion vectors are iteratively refined. Therefore in each iteration the control points are scanned from top left to bottom right of the image. Two iterations or, in case of larger motion, three iterations are performed using a 5-step search algorithm in each iteration with step-sizes varying in the set $\{8, 4, 2, 1, 0.5\}$. For low motion regions, the larger steps can be omitted.

In order to improve the prediction in presence of large motion a *hierarchical motion estimation scheme* changing the grid size from coarse to fine is employed, *e.g.* an initial grid size of 32×32 followed by 16×16 pels can be used. The vectors of the finer grid are initialized by bilinear interpolation and subsequent rounding to half pel accurate values of the prior estimated vectors of the coarse grid. This also leads to a more homogeneous motion vector field. To further force the estimation of smooth vector fields a Lagrangian multiplier λ is used to choose the best control point motion vector considering the prediction error (in the MSE sense) and the local motion vector variance:

$$\begin{aligned} MV_i^{\text{opt}} \;=\; & \arg\min_{MV_j} \Big\{ \text{MSE}\left(MV_j\right) \\ & + \lambda \cdot \sum_{k=1}^{8} \left(MV_k - MV_j\right)^2 \Big\}. \end{aligned} \quad (5)$$

The decision whether to use warping prediction or OBMC for a block is also based on Eq. (5). The prediction type yielding lower cost is chosen. To avoid a large increase of the computational workload, the MSE for both schemes, *i.e.* warping prediction and OBMC, is computed for the motion vector in question only under consideration of the bottom right subblock of the octagon (*cf.* Fig.1). The three other subblocks are using the prediction schemes which were previously assigned during motion estimation of vectors at causal positions.

C. Loop Filter

A warping predictor utilizes a higher amount of fractional motion which leads to a strong lowpass effect when a bilinear pixel interpolation is employed. A high quality texture interpolation reduces the lowpass effect of the first order filtering of the bilinear texture interpolation. Hereby the peak signal-to-noise ratio (PSNR) can be increased up to 1.0–1.5 dB, or equivalently, bit savings up to 30% can be achieved compared to bilinear pixel interpolation [7].

Best results were obtained by using the predicted pixel at the half or quarter pel position next to the unknown value for prediction. Tensor products of one-dimensional 10-tap hamming weighted sinc interpolation filters with pre-stored coefficients are used to compute the predictions at quarter pel positions.

As this interpolation scheme emphasizes the high frequency image components the predicted image gets much sharper and less high frequency coefficients must be coded. Unfortunately, also the included coding artifacts are more visible. Therefore, a fully non-block based coding scheme such as the proposed warping predictor and the wavelet-based residual coder should be used to avoid any blocking artifacts. In order to also reduce the ringing effects a block-wise adaptive interpolation scheme is employed. In this scheme, after the motion vectors are estimated it is adaptively decided for each square block whether to use the high quality or the bilinear texture interpolation. The MSE is used as a criterion for the decision, which leads to a slightly increased PSNR and an improved visual quality. This scheme can also be viewed as an adaptive loop filter. As additional side information, one bit per block is transmitted indicating the type of loop filter.

III. Wavelet-Based Coding

Encoding of the motion compensated P-frames as well as of initial intraframes (I-frames) is performed by means of a conventionally structured transform coding scheme. First, a wavelet transform is applied to an entire frame. Uniform scalar quantization with a central dead-zone around zero similar to that designed for H.263 is then used to map the dynamic range of wavelet coefficients to a reduced alphabet of decision levels. The actual statistical coding of the quantized wavelet coefficients is complemented by a 3-stage pre-coding process based on the PACC coding principle. In this section, we give a brief description of this coding framework along with a discussion of some issues concerning construction and choice of an appropriate family of wavelet filters.

A. Biorthogonal Wavelet Bases Associated with Recursive Filters

Orthogonal wavelet bases, associated with recurcive filters, were investigated in detail by Herley and Vetterly [8]. In video and image processing biorthogonal pairs of bases consisting of functions having odd or even symmetry are usually more effective than orthogonal non-symmetric wavelet bases. For example, in wavelet-based image compression the so-called 9/7-wavelet with compact support, which was constructed in [4], is most frequently used. Concerning the bases from [8], which combine orthogonality and symmetry, their adaptation to a concrete problem is rather problematic due to inherently severe constraints.

The standard approach of constructing a dual pair of biorthogonal bases consists in reducing this problem to

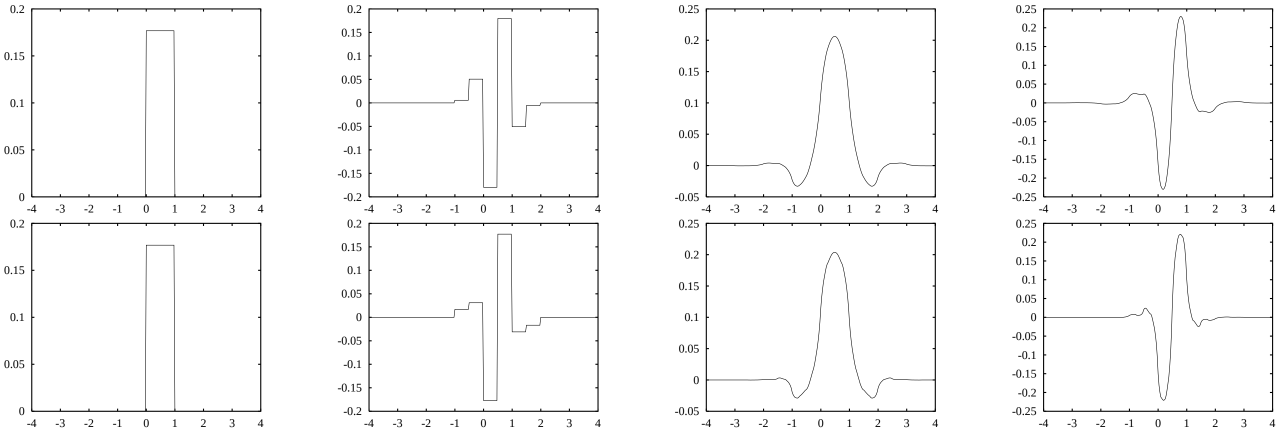

Fig. 4. From left to right: scaling function of analysis, analyzing wavelet, scaling function of synthesis, and synthesizing wavelet used for I-frame coding (top row) and P-frame coding (bottom row).

finding solutions of the matrix equation

$$M(z)\tilde{M}^T(z^{-1}) = 2I, \tag{6}$$

where I is the identity matrix, and

$$M(z) = \begin{pmatrix} h(z) & h(-z) \\ g(z) & g(-z) \end{pmatrix}, \quad \tilde{M}(z) = \begin{pmatrix} \tilde{h}(z) & \tilde{h}(-z) \\ \tilde{g}(z) & \tilde{g}(-z) \end{pmatrix}$$

are so-called *modulation matrices*, whose components in our case are rational functions, satisfying the normalizing conditions $h(1) = \tilde{h}(1) = \sqrt{2}$. Here, and in the following, filters h and g denote low-pass and high-pass filters of the decomposition algorithm, respectively, while $\tilde{h}$ and $\tilde{g}$ denote the corresponding filters for reconstruction.

In the polynomial case, the relations

$$g(z) = z^{-1}\tilde{h}(-z^{-1}), \; \tilde{g}(z) = z^{-1}h(-z^{-1})$$

necessarily hold. However, for rational solutions of Eq. (6) the last conditions can be violated. We consider an one-parameter family of filters

$$\begin{aligned} h(z) &= \frac{1}{\sqrt{2}}(1+z), & (7)\\ \tilde{h}(z) &= \frac{(2+a)(z^{-1}+3+3z+z^2)(z^{-1}+b+z)}{4\sqrt{2}(2+b)(z^{-2}+a+z^2)}, & (8)\\ g(z) &= \frac{(2+a)(z^{-1}-3+3z-z^2)(-z^{-1}+b-z)}{4\sqrt{2}(2+b)}, & (9)\\ \tilde{g}(z) &= \frac{1}{\sqrt{2}}\frac{1-z^{-1}}{z^{-2}+a+z^2}, & (10) \end{aligned}$$

where $b = \frac{4a-8}{6-a}$, $\quad |a| > 2, \quad a \neq 6$. When $a = 6$ we get a more simple pair of filters

$$h(z) = \frac{1}{\sqrt{2}}(1+z), \quad \tilde{h}(z) = \sqrt{2}\frac{z^{-1}+3+3z+z^2}{z^{-2}+6+z^2}.$$

The case $|a| \leq 2$ leads to unstable recursive filters because in this case the absolute value of roots of the denominator equals 1.

We note that the choice of the filters g and $\tilde{g}$ in Eqs. (9,10) is unusual since denominator $z^{-2}+a+z^2$ has moved from g to $\tilde{g}$. From the point of view of wavelet terminology, this modification does not lead neither to a change of the dual pair of *multiresolution analyses* nor to a change of wavelet spaces. However, it modifies the wavelet bases.

The nature and statistics of the images to be transformed is different for intra- and interframes. In intraframe mode we have to compress ordinary still images whereas interframe mode provides, in fact, partial derivatives of video sequences in time. So, we optimized bases for these modes separately. Our numerical simulations gave the optimal values $a = 8$ for intraframe and $a = 25$ for interframe mode in relations (7)–(10). Graphs of the optimal bases are presented in Fig. 4.

B. The PACC Pre-Coding Framework

For encoding the quantized wavelet coefficients, we follow the conceptual ideas presented in[10]. Next, we give a brief review of the involved techniques. For more details, the readers are referred to [10], [11].

B.1 Partitioning

Fig. 5 shows a block diagram of the proposed pre-coder, following the PACC concept: the initial "partitioning" process divides each frame into three sub-sources: the significance map, indicating the position of significant coefficients, the magnitude map, holding the absolute values of significant coefficients and the sign map with the phase information of the wavelet coefficients. Note, that all three sub-sources inherit the subband structure from the quantized wavelet decomposition, so that there is another partition of each subsource according to this band structure.

B.2 Aggregation

In a second stage, the pre-coder performs an "aggregation" of insignificant coefficients across the scales using a quadtree related data structure [9], [10]. These so-called *zerotrees* connect insignificant wavelet coefficients which

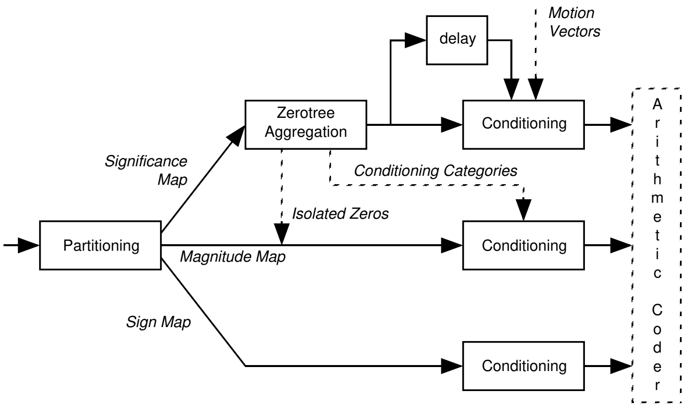

Fig. 5. Schematic representation of the PACC pre-coder.

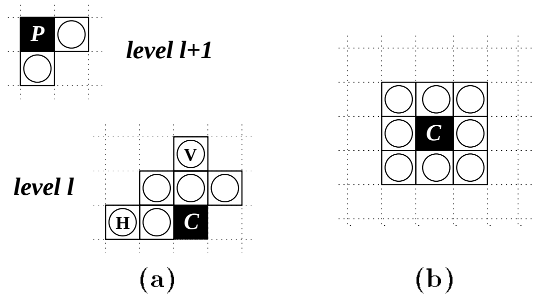

Fig. 6. (a) Two-scale template (white circles) with an orientation dependent design for conditional coding of significance maps; V, H: additional element used for vertical and horizontal oriented bands, respectively. (b) 8-neighborhood used for conditioning of magnitudes.

share the same spatial location across different scales of the octave-band decomposition. However, in contrast to other zerotree-based coding methods, we only consider zerotree roots localized in the lowest frequency bands which guarantee a sufficient coding benefit.

B.3 Conditional Coding

The final "conditioning" part of the pre-coder supplies the elements of each source with a "context", *i.e.* an appropriate model for the actual coding process in the arithmetic coder [15]. Fig. 6 (a) shows the prototype template used for conditioning of elements of the significance map. It consists of two parts. The first part is a causal neighborhood of the actual coding event C which depends on scale and orientation of a given band. Except for the lowest frequency bands, the second part of the template uses additional information of the next upper level, *i.e.* significance information on the next coarser scale, which is given by the neighbors of the parent P of C, thus allowing a "prediction" of the non-causal neighborhood of C. The processing of subbands is performed in the order from lowest to highest frequency bands, and the partitioned data of each band is processed such that the significance information is coded (and decoded) first.

This allows the construction of special conditioning states for the coding of magnitudes using the local significance information. Thus, the actual conditioning of magnitudes is performed by classifying significant coefficients according to the local variance estimated by the significance of their 8-neighborhood, as depicted in Fig. 6 (b). For conditional coding of sign maps a second-order Markov model is used whose states are built of two preceding signs of a given sign event with respect to the orientation of the related band [10], [11].

IV. Experimental Results

Simulations were carried out using the MPEG-4 test sequences *Container Ship* and *Foreman* (30 Hz, 300 frames, 176×144 pels). As a reference system we used an implemenation of the H.263+ coder (H.263, Version 2) [2], [5] provided by the University of British Columbia (UBC)[1] (software version 3.1.1). There are several modes of operation permitted by a H.263 compliant encoder as defined by a number of negotiable options. For the runs of the H.263+ coder (TMN9), the unrestricted motion vector mode (Annex D), the advanced prediction mode (Annex F), the improved intra coding mode (Annex I), and the deblocking filter mode (Annex J) have been enabled. For each run of both coders, the quantization step-size of intra- and interframe coding was fixed. To obtain the desired frame rate of 10 Hz, the original sequences were temporally subsampled.

Two sets of experiments were carried out. First, we evaluated the performance of the intraframe coding of our wavelet-based coding scheme in comparison to the advanced intra coding mode of H.263+. Fig. 7 shows the results for encoding the first frame of the *Foreman* sequence in intra-mode. The proposed new scheme outperformed the H.263+ coder by 0.7–2.7 dB PSNR for the luminance component with higher coding gains towards higher bit-rates (*cf.* Fig. 7). Subjectively, comparing the reconstructed I-frames of our proposed coder to those of TMN9 at the same bitrate, we observed an improved quality in favor of our proposed scheme especially at lower rates where the reconstructions of our coder were much sharper and free of any blocking artifacts.

For the second part of our experiments, we used a set of different fixed quantizer step-sizes for prediction residual coding. The first I-frame was coded with identical step-size but with a smaller dead-zone (relative to that of the P-frames). Figures 8 and 9 show the rate-distortion (RD) graphs of the luminance (Y) component as a result of these coding experiments. We measured Y-PSNR versus bit-rate of the complete bitstream. The RD-curves were plotted using averaged Y-PSNR over bit-rate, where the latter was calculated as the arithmetic mean of total bits per frame (including the intraframe) multiplied by the desired frame rate of 10 Hz.

Compared to the reference coder, the proposed coding scheme achieved a gain in averaged PSNR of 1.0–1.5 dB for the very active *Foreman* sequence (*cf.* Fig. 8). For the *Container Ship* sequence (*cf.* Fig. 9), the gain in PSNR

[1] http://www.ece.ubc.ca/image

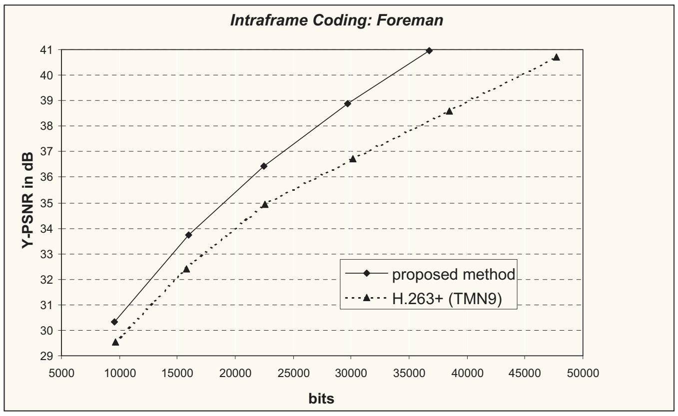

Fig. 7. Rate-Distortion comparison of intraframe coding for *Foreman*.

amounts to 1.8–2.2 dB. In addition, the visual quality of the reconstructed frames appeared to be much improved. However, our reconstructed frames were suffering from some ringing noise at very low bit-rates. We are currently investigating how to reduce these artifacts by improving our temporal prediction scheme and by studying the influence of different choices of wavelet filters.

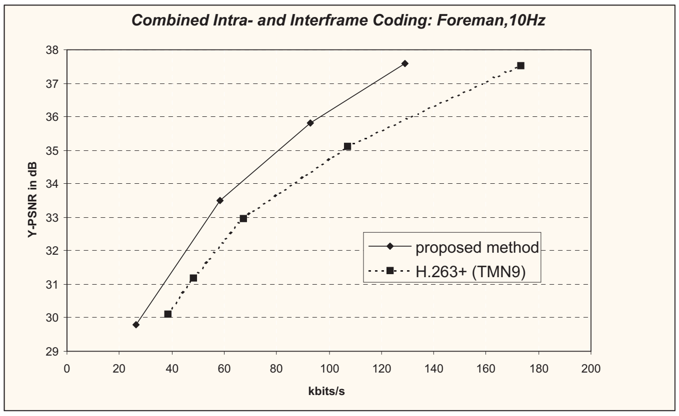

Fig. 8. Averaged Y-PSNR [dB] over bit-rate [kbits/s] for *Foreman* at 10 Hz.

V. Conclusions

A new blocking artifact free video coding scheme combining warping based prediction, overlapped motion compensation and wavelet-based residual coding has been presented. The proposed coding method has proven to be very efficient both with respect to objective and subjective evaluation.

Acknowledgments

Parts of this work were supported by Deutsche Telekom Berkom GmbH, Berlin, Germany, and by Russian Foundation for Basic Research under grant 97-01-00443.

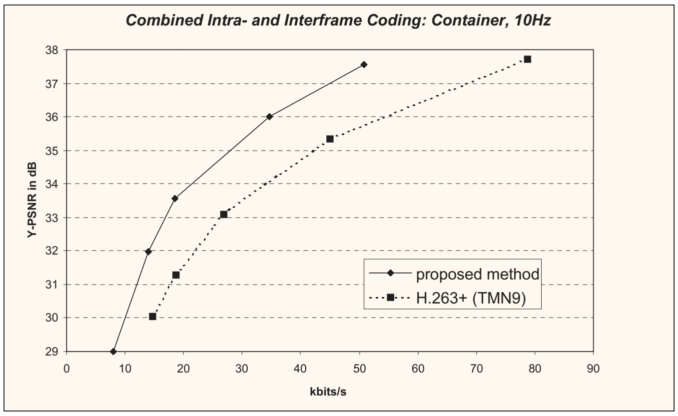

Fig. 9. Averaged Y-PSNR [dB] over bit-rate [kbits/s] for *Container Ship* at 10 Hz.

References

[1] ITU-T *Recommendation H.263* "Video Coding for Low Bit-Rate Communication", Mar. 1996.
[2] ITU-T *Draft Recommendation H.263 Version 2* "Video Coding for Low Bit-Rate Communication", Sept. 1997.
[3] A. P. Petukhov, "Biorthogonal Wavelet Bases with Rational Masks and Their Application", Preprint 1998.
[4] A. Cohen, I. Daubechies and J.-C. Feauveau, "Biorthogonal Bases of Compactly Supported Wavelets", *Comm. on Pure and Appl. Math.*, **45**, pp. 485–560, 1992.
[5] G. Côté, B. Erol, M. Gallant, and F. Kossentini, "H.263+: Video Coding at Low Bit-Rates", *IEEE Trans. on Circuits and Systems for Video Techn.*, **8**, No. 7, pp. 849–866, Nov. 1998.
[6] G. Heising and G. Ruhl, "Video Coding Using Spatial Extrapolation Based Motion Field Segmentation", *IEEE Int. Conf. Image Proc.*, **2**, pp. 482–485, Sep. 1996.
[7] G. Heising, "Blocking Artifact Free Video Coding by Combining Warping Based Prediction with Wavelet Error Coding", *Proc. Picture Coding Symposium (PCS) '97*, pp. 309–314, 1997.
[8] C. Herley and M. Vetterli, "Wavelet and Recursive Filter Banks", *IEEE Trans. on Signal Proc.* **41**, No. 8, pp. 2536–2556, 1993.
[9] A. Lewis and G. Knowles, "Image Compression Using the 2D Wavelet Transform", *IEEE Trans. on Image Proc.*, **1**, No. 2, pp. 244–250, April 1992.
[10] D. Marpe and H. L. Cycon, "Efficient Pre-Coding Techniques for Wavelet-Based Image Compression", *Proc. PCS '97*, pp. 45–50, 1997.
[11] D. Marpe and H. L. Cycon, "Very Low Bit-Rate Video Coding Using Wavelet-Based Techniques", *IEEE Trans. on Circ. and Sys. for Video Techn.*, **9**, No. 1, pp. 85–94, Feb. 1999.
[12] D. Marpe, "A Layered Approach to the Integration of Lossy and Lossless Image Compression", *Proceedings Picture Coding Symposium 1999*, Portland, USA, April 1999.
[13] G. J. Sullivan and R. L. Baker, "Motion Compensation for Video Compression Using Control Grid Interpolation", *Proc. Int. Conf. ASSP*, Toronto, Canada, pp. 2713–2716, May 1991.
[14] H. Watanabe and S. Singhal, "Windowed Motion Compensation", *Proc. SPIE Conf. Visual Comm. Image Proc.*, Vo. 1605, pp. 582–589, Nov. 1991.
[15] I. Witten, R. Neal and J. Cleary, "Arithmetic Coding for Data Compression", *Comm. ACM*, **30**, pp. 520–540, June 1987.
[16] G. Wolberg, *Digital Image Warping*, IEEE Computer Society Press, Los Alamitos, CA, 1990.