ADAPTIVE INTERPOLATION FILTER FOR MOTION COMPENSATED HYBRID VIDEO CODING

Thomas Wedi

Institut fuer Theoretische Nachrichtentechnik und Informationsverarbeitung
Universitaet Hannover, Appelstr. 9a, 30167 Hannover, Germany
E-Mail: wedi@tnt.uni-hannover.de

ABSTRACT

Common hybrid video coding standards such as H.263, MPEG-1,2,4 are based on motion compensated prediction. For the motion compensated prediction purpose, displacement vectors with fractional-pel resolutions (e.g. 1/2-pel or 1/4-pel) are used. In order to estimate and compensate fractional-pel displacements, the image signal has to be interpolated. The most popular and very simple interpolation method is the bilinear interpolation. Recent approaches like the MPEG-4 ACE-profile and H.26L use Wiener interpolation filters. Up to now these filters are based on filter-coefficients that are invariant. The same filter is used for all sequences and for all images of a sequence. In this paper an adaptive interpolation filter is introduced in order to improve the coding efficiency. The adaptive interpolation filter is based on filter-coefficients that are estimated for each image. Thus, the influences of the displacement estimation error and the aliasing components, which are deteriorating the motion compensated prediction are reduced. Analytical calculations and experimental results show a dependence between the variance of the displacement estimation error and the filter characteristic. Compared to the H.26L codec the adaptive interpolation filter improves the PSNR up to 1.0 dB for a sequence. For single images with high displacement estimation errors, gains up to 1.5 dB are obtained.

1. INTRODUCTION

Standardised hybrid video codec like H.263, MPEG-1,2,4 are based on motion compensated prediction. Figure 1 shows the generalised block diagram of a hybrid video encoder. The current image $s(k)$ at time instance k is predicted by a motion compensated prediction from an already transmitted image $s'(k\text{-}1)$. The result of the motion compensated prediction is image $\hat{s}(k)$. Only the prediction error $e(k)$ and the motion information $\vec{d}(k)$ is coded and transmitted.

For the motion compensated prediction, the current image is partitioned into blocks. A displacement vector $\vec{d}(k)$ is assigned to each block, which refers to the corresponding position of its image signal in an already transmitted reference image $s'(k\text{-}1)$. The displacement vectors have a fractional-pel resolution. The most frequently used fractional-pel resolution of the displacement vectors $\vec{d}(k)$ is 1/2-pel (H.263, MPEG-1,2,4). Recent approaches like MPEG-4 ACE [4] and H.26L [3] are based on 1/4-pel displacement vector resolution.

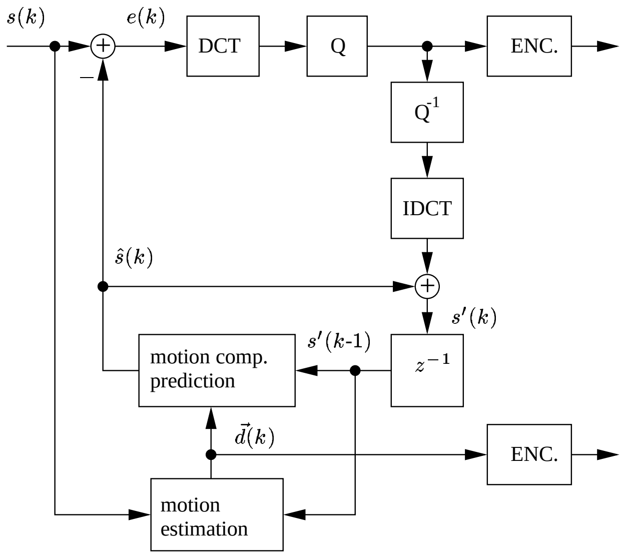

Fig. 1. Block diagram of a hybrid video encoder based on motion compensated prediction.

Displacement vectors with a fractional pel resolution may refer to positions in the reference image, which are located between the sampled positions of its image signal. In the following these positions are called subpel positions. In order to estimate and compensate fractional-pel displacements, the image signal on subpel positions has to be generated by interpolation. Up to now this interpolation is done by interpolation filters that are invariant. The same filter is used for all sequences and for all images of a sequence.

In this paper a motion compensated prediction scheme with adaptive interpolation filters is proposed in order to re-

Proc. Picture Coding Symposium (PCS 2001), Seoul, Korea, April 2001

duce the prediction error and improve the coding efficiency. The adaptive interpolation filter is based on filter-coefficients which are adapted during the motion compensated prediction process from image to image.

In Section 2 the motion compensated prediction process with the adaptive interpolation filter is introduced in detail. Section 3 presents an analysis of the influences on the adaptive interpolation filter. Experimental results are given in Section 4. The paper closes with a summary.

2. MOTION COMPENSATED PREDICTION USING AN ADAPTIVE INTERPOLATION FILTER

In this Section the motion compensated prediction module shown in Figure 1 is described in detail. This module uses the already transmitted signal $s'(k\text{-}1)$ and the estimated displacement vectors $\vec{d}(k)$ in order to create the prediction image $\hat{s}(k)$. Figure 2 shows the block diagram of the motion compensated prediction module for a displacement vector resolution of $1/M$. The motion compensated prediction is

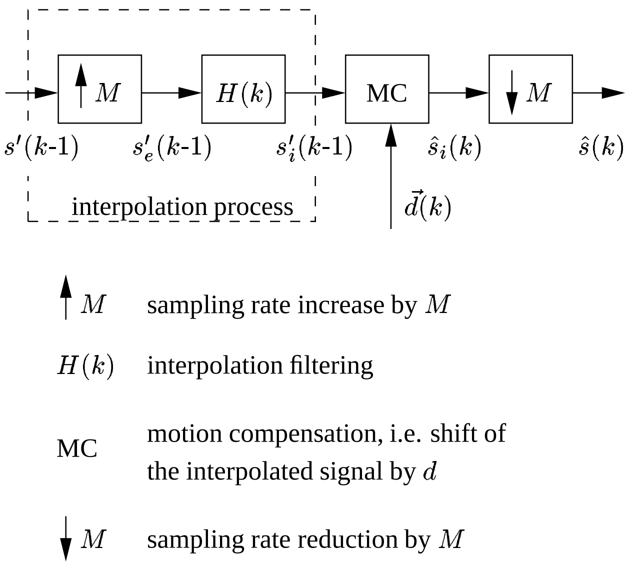

Fig. 2. Block diagram of the motion compensated prediction based on a displacement vector resolution of $1/M$.

performed in two steps. The first step is the interpolation process, where the spatial sampling rate of the already transmitted image $s'(k\text{-}1)$ is increased by a factor of M and filtered with an interpolation filter H. In case of an adaptive interpolation, the filter $H = H(k)$ is a function of the time instance k. The result of the first step is the interpolated image $s'_i(k\text{-}1)$. In the second step the interpolated signal is shifted according to the estimated displacement vector $\vec{d}(k)$ and the sampling rate is reduced by the factor of M. The result is the motion compensated image $\hat{s}(k)$.

The most popular and very simple interpolation method is the bilinear interpolation (H.263, MPEG-2 and MPEG-4). Recent approaches like MPEG-4 ACE and H.26L use Wiener interpolation filters. These Wiener interpolation filters were designed to interpolate the image signal while reducing aliasing components that are deteriorating the motion compensated prediction [1, 2, 5]. The Wiener filters were designed to interpolate $s'(k\text{-}1)$ of Fig. 2, regardless of the motion compensation module MC. Furthermore the filter-coefficients are invariant. The same filter $H(k)$=H is used for all sequences and for all images of a sequence. Different cameras, i.e. different image acquisition processes with different LP-filters produce different aliasing components in the image signal. These varying aliasing components can not be considered by invariant filters.

In this paper an adaptive interpolation filter is introduced in order to reduce the prediction error and improve the coding efficiency. For each image at each time instance k new filter-coefficients for filter $H(k)$ are estimated and applied. The motion compensated prediction with the adaptive interpolation filter consists of the following four steps:

1. Estimation of displacement vectors $\vec{d}(k)$. For this purpose an invariant Wiener filter is applied.

2. Estimation of filter $H(k)$ that minimises the energy of the prediction error $e(k)$ when performing the MC with $d(k)$ from step 1.

3. Coding of filter coefficients of $H(k)$.

4. Motion compensated prediction by using $\vec{d}(k)$ from step 1 and $H(k)$ from step 2. (see Fig. 2)

Further iterations where step 1 is performed with the new estimated filter $H(k)$ does not lead to significant improvements. With the adaptive filter the following two advantages are achieved compared to the Wiener filter:

1. Adaptation:

Weiner Filter:	no adaptation $H(k)$=H=const
Adapt. Filter:	adaptation to the statistics of each image

2. Optimisation criterion for filter design:

Weiner Filter:	interpolate $s'_i(k\text{-}1)$ while reducing aliasing components
Adapt. Filter:	minimise the energy of the prediction error $e(k)$

In the following section an analysis is given that explains why an adaptive interpolation filter reduces the remaining prediction error and improves the coding efficiency.

3. INFLUENCES ON THE ADAPTIVE INTERPOLATION FILTER

There are two reasons why an adaptive interpolation filter reduces the remaining prediction error and improves the coding efficiency. The first reason is the consideration of varying aliasing signals and the second reason is the consideration of varying displacement estimation errors. Both influences on the adaptive interpolation filter are analysed in the following two subsections.

3.1. Influence of Aliasing

Due to non-ideal low-pass filters in the image acquisition process, the Nyquist Sampling Theorem is not fulfilled and aliasing disturbs the motion compensated prediction. This leads to an additional prediction error, which has to be coded. Since the non-ideal low-pass filters differ for various image acquisition processes, the aliasing of various sequences also differs. An invariant interpolation filter as it is used up to now cannot consider these varying aliasing signals.

3.2. Influence of Displacement Estimation Errors

Due to the limited resolution of the displacement vectors (e.g. 1/2- or 1/4-pel) and the limited validity of the translational motion model, displacement estimation errors occur. A displacement estimation error leads to an inaccurate displacement vector. This means that the displacement vector refers to an inaccurate spatial position in the reference image. Due to this inaccurate spatial position a prediction error remains, which has to be coded. If this inaccurate spatial position is a subpel position, an optimal interpolation filter can consider the inaccuracy of the displacement vector in order to reduce the prediction error.

In the following the influence of the displacement estimation error on the motion compensated prediction image is analysed analytical. In this equation (1) the one dimensional representation with one spatial coordinate x is chosen.

$$\hat{s}(x,t) = s'(x - \hat{d_x}, k-1) \tag{1}$$

The prediction result $\hat{s}(x,t)$ at time instance k is equal to the shifted version of the already transmitted image signal. Here $\hat{d_x}$ is the estimated displacement in x-direction. It is assumed, that $\hat{d_x}$ is given by

$$\hat{d_x} = d_x + e_x \tag{2}$$

where d_x is the exact displacement and e_x is the displacement estimation error. With equation (1) $\hat{s}(x,t)$ is given by

$$\hat{s}(x,t) = s'(x - (d_x + e_x), k-1) \tag{3}$$

If it is assumed, that e_x can be modelled by a random process with a probability density function $p(e_x)$, the calculation of $\hat{s}(x,t)$ can be expressed by an expectation value. Therefore $\hat{s}(x,t)$ is given by the following equation.

$$\begin{aligned} \hat{s}(x,t) &= E\left[s'\left(x - (d_x + e_x), k-1\right)\right] \qquad (4) \\ &= \int_{-\infty}^{\infty} p(e_x) \cdot s'\left(x - (d_x + e_x), k-1\right)\, de_x \end{aligned}$$

The integration over $p(e_x) \cdot s'(\dots)$ in equation (4) can be interpreted as a filtering of $s'(\dots)$ with a filter that is give by the probability density $p(e_x)$. If the displacement estimation error variance σ_e^2 of $p(e_x)$ is increased, the distribution of $p(e_x)$ gets wider. This means that the signal $s'(\dots)$ becomes a wider influence on $\hat{s}(x,t)$. Thus the filter characteristic depends on σ_e^2 and $p(e_x)$ respectively. Since the image content differs from scene to scene, σ_e^2 also differs. An invariant interpolation filter as it is used up to now cannot consider these varying displacement estimation errors.

4. EXPERIMENTAL RESULTS

For the experimental results, the proposed motion compensated prediction scheme with the adaptive interpolation filter is integrated into the H.26L (TML-4) codec. The H.26L (TML-4) codec is based on 1/4-pel displacement vector resolution and a Wiener interpolation filter with six symmetric filter-coefficients (6-tap filter). For each image of a sequence the interpolation filter is adapted and its filter- coefficients are DPCM coded and transmitted as side information. The applied sequences are *Mobile & Calendar* and *Foreman* in CIF format.

There is no image in each of the sequences, where the use of the adaptive filter leads to a performance loss. This is due to the fact that the adaptive filter is at least as good as the invariant filter and that the additional side information is very small. For a symmetric 6-tap adaptive interpolation filter only three coefficients have to be coded and transmitted as side information for each image. Figure 3 show the operational rate distortion curves for these sequences. The coding gain has been measured by the PSNR of the reconstructed image signal. Due to the adaptive interpolation filter the PSNR can be improved up to 1.0 dB for a sequence. For single images with high displacement estimation errors, gains up to 1.5 dB are obtained.

In Fig. 4 two different images and the corresponding filters for two different time instances of the sequence *Foreman* are shown exemplarily. In part a) of this figure an image of the sequence is shown where high displacement estimation errors occur. In part b) an image of the sequence

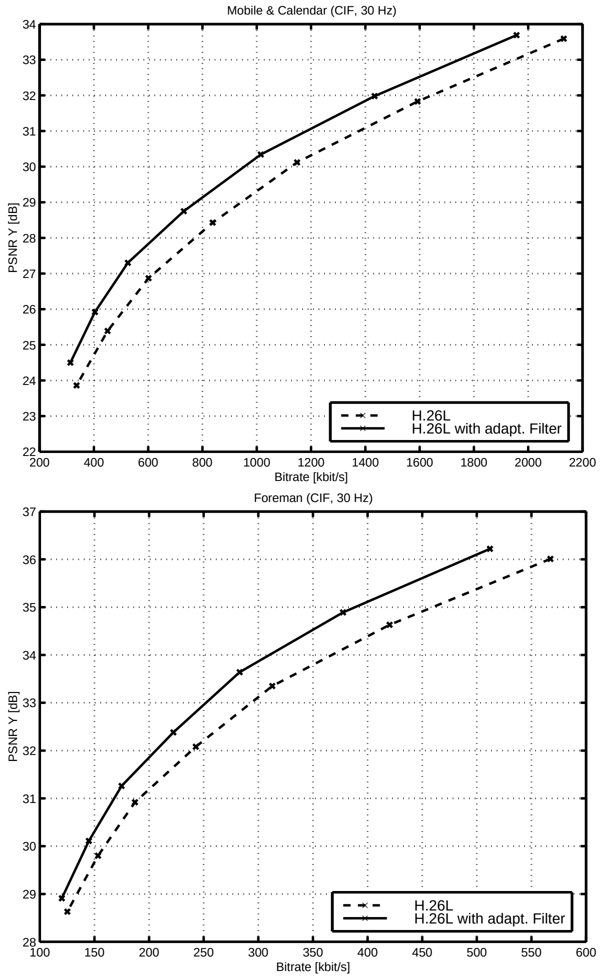

Fig. 3. Operational rate distortion curves for test sequences *Mobile & Calendar* and *Foreman*.

is shown where the displacement error variance is much smaller than in part a). It can be seen that the filter characteristic becomes sharper when the displacement estimation error decreases.

5. SUMMARY

A motion compensated hybrid video coding scheme using an adaptive interpolation filter is presented. The adaptive interpolation filter is based on filter-coefficients which are estimated during the motion compensated prediction process for each image. The cost function for the estimation of the filter-coefficients is the energy of the motion compensated prediction error. Due to the adaptive interpolation filter the PSNR can be improved up to 1.0 dB in average for a sequence. For single images with high displacement estimation errors, gains up to 1.5 dB are obtained.

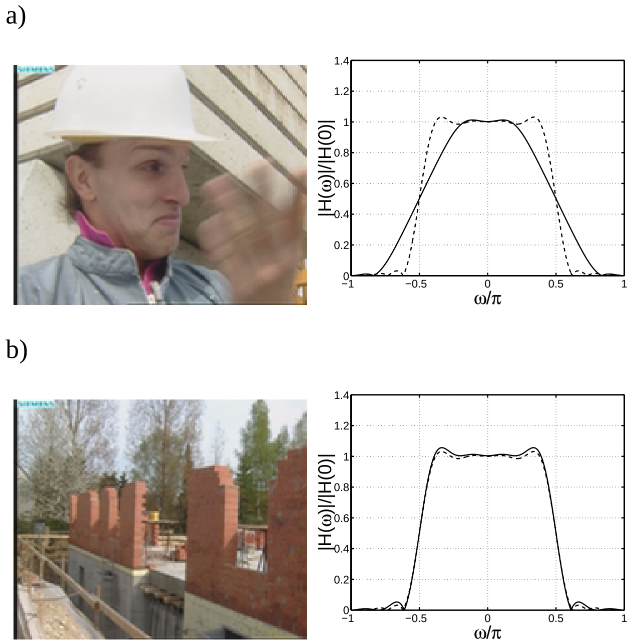

Fig. 4. Original images (left) and magnitude responses of corresponding adaptive filters for two time instances a) and b) of sequence *Foreman*. Dashed line: Wiener filter, solid line: adaptive filter

6. REFERENCES

[1] O. Werner, ”Drift analysis and drift reduction for multiresolution hybrid video coding”, Signal Processing: Image Communication, Vol. 8, No. 5, July 1996, p.387-409, ISSN: 0923-5965

[2] U. Benzler, O. Werner, ”Improving multiresolution motion compensating hybrid coding by drift reduction”, Picture Coding Symposium 1996, March 1996, Melbourne.

[3] ITU-T SG16/Q15 (G. Bjontegaard), ”H.26L Test Model Long Term Number 5 (TML-5)”, ITU-T SG16/Q15 doc. Q15-K-59, (downloadable via ftp://standard.pictel.com/video-site), Aug. 2000

[4] MPEG-4: ISO/IEC, ”Final Proposed Draft Amendment 1 for ISO/IEC 14496-2”,
Doc. ISO/IEC/JTC1/SC29/WG11 N2802, Vancouver, July 1999.

[5] MPEG-4: ISO/IEC (U. Benzler), ”Results of core experiment P8 (Motion and Aliasing Compensated Predicition”, Doc. ISO/IEC/JTC1/SC29/WG11 N2625, October 1997.