A TIME-RECURSIVE INTERPOLATION FILTER FOR MOTION COMPENSATED PREDICTION CONSIDERING ALIASING

Thomas Wedi

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

ABSTRACT

In this paper a new interpolation filter for motion compensated prediction in a hybrid video codec is presented. It allows to interpolate an aliasing affected image signal more accurate than conventional interpolation filters like bilinear or Wiener filter. In contrast to conventional filters, the new time-recursive filter uses not only the image signal of the current frame, but also the image signal of previous frames in order to interpolate the current frame. The filter has been tested with synthetic and real sequences. The experimental results for the synthetic sequence show, that with conventional filters it is not possible to predict an aliasing affected image perfectly, even if the estimated displacements perfectly fit to the sequence. Furthermore it is shown, that with the new filter it is possible to predict an aliasing affected image perfectly. For the used real sequence a coding gain of 0.8 dB is achieved compared to an MPEG-4 codec.

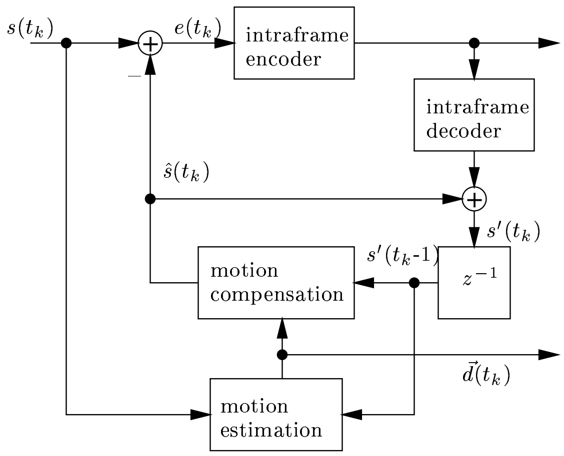

Figure 1: Block diagram of a hybrid video encoder based on motion compensated prediction.

1. INTRODUCTION

Standardized hybrid video codecs like H.263, MPEG-2, MPEG-4 etc. are based on motion compensated prediction [3, 4, 8]. Figure 1 shows the block diagram of such a hybrid video encoder. In the motion estimation module the displacement vector field $\vec{d}(t_k)$ between consecutive images is estimated. The displacement vectors and the previous reconstructed image $s'(t_k\text{-}1)$ are used in the motion compensation in order to create the prediction image $\hat{s}(t_k)$. The remaining prediction-error $e(t_k)$ and the displacement vector field are encoded and transmitted.

The standardized coding techniques use displacement vectors with fractional-pel resolution [4, 5, 8]. In order to estimate and compensate fractional-pel displacements, the image signal between the sampled positions of image $s'(t_k\text{-}1)$ has to be reconstructed. This reconstruction is equivalent to an increase of the sampling rate of $s'(t_k\text{-}1)$ [1, 2]. In the following the increase of the sampling rate is denoted as upsampling. In case of $\frac{1}{L}$-pel resolution of the displacement vectors, $s'(t_k\text{-}1)$ has to be sampled up by the factor of L. The result of such an upsampling is denoted by $s'_u(t_k\text{-}1)$. In the following the samples of $s'(t_k\text{-}1)$ are called fullpel samples and the samples to be reconstructed, between the fullpel samples are called subpel samples.

Due to non ideal LP-filters in the image acquisition process, there is aliasing in the sampled image signal $s(t_k)$. Thus, the aliasing in the image to be coded $s(t_k)$ should also be predicted by a motion compensated prediction in order to reduce the prediction error.

Conventional motion compensated prediction uses bilinear interpolation to generate the upsampled image $s'_u(t_k\text{-}1)$ and thus do not take into account aliasing [8]. Recent approaches use Wiener Filters in order to reduce the aliasing effect [6, 7]. But, since the aliasing signal is unknown, Wiener filters can not compensate the aliasing perfectly [1, 2, 6]. Therefore the accuracy of the prediction and the coding efficiency is still limited, even if there are no displacement estimation errors.

In the following section, a new interpolation filter is presented, which allows to interpolate an aliasing affected image perfectly. The filter uses not only the image signal of the current frame, but also the image signal of previous frames in order to interpolate the current frame.

2. THE TIME-RECURSIVE INTERPOLATION FILTER

The idea of the new interpolation filter presented in this paper is the following: If an unmoved, analog image signal is sampled at the same positions at different times, the two sampled signals are identical even if the sampling rate doesn't fit to the Nyquist frequency. Therefore the sampled images also have the same aliasing signal. If the image signal is displaced by one sample, the corresponding, displaced samples are identical and have the same aliasing. This shows, that at full-pel displacements the aliasing does not affect the prediction. Furthermore, it induces a new interpolation approach for increasing the sampling rate. If it is known that a signal is shifted by a fractional-pel displacement, we may insert the former samples from the fullpel positions at the corresponding subpel positions in the current upsampled image, knowing that they contain the correct sampled image signal. Even the aliasing is predicted correctly.

Based on this idea a new filter is presented in order to generate the upsampled image $s'_u(t_k\text{-}1)$, which takes fullpel samples of previous frames to reconstruct the subpel samples of the frame to be interpolated. The filter is a time-recursive interpolation filter and is therefore called TRI-Filter. In Figure 2 the block diagram of the TRI-Filter is shown for an $\frac{1}{L}$-pel resolution of the displacement vectors.

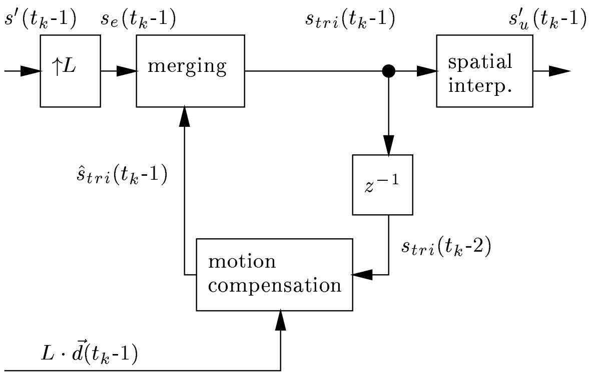

Figure 2: Block diagram of the time-recursive interpolation filter (TRI-Filter) with motion vectors of $\frac{1}{L}$-pel resolution. Image $s'(t_k\text{-}1)$ is sampled up by the factor of L to image $s'_u(t_k\text{-}1)$.

The filter consists of three major components. The first component is the expander $\uparrow L$, the second is the recursive structure with the *merging* and the *motion compensation* module and the third component is the *spatial interpolation*. These three components are described in the following.

In the first component, the module $\uparrow L$ expands the sampling rate of image $s'(t_k\text{-}1)$ by the factor of L, resulting in image $s_e(t_k\text{-}1)$.

$$s_e(t_k\text{-}1, x, y) = \begin{cases} s'(t_k\text{-}1, \frac{x}{L}, \frac{y}{L}), & x, y = 0, \pm L, \pm 2L, \ldots \\ m, & \text{otherwise} \end{cases} \tag{1}$$

As equation (1) shows, this is done by inserting $s'(t_k\text{-}1)$ on fullpel positions and filling a marker value m on the subpel positions. The marker value indicates samples which are not yet reconstructed. In (1) x and y denote image coordinates.

In the second component, the samples of image $s_{tri}(t_k\text{-}2)$ are used to replace the marked samples of $s_e(t_k\text{-}1)$. This is done by a *motion compensation*, where the samples of $s_{tri}(t_k\text{-}2)$ are shifted in the direction of the motion. The result of the motion compensation is image $\hat{s}_{tri}(t_k\text{-}1)$. In the module *merging* the marked samples of $s_e(t_k\text{-}1)$ are replaced by corresponding samples of $\hat{s}_{tri}(t_k\text{-}1)$. The result is image $s_{tri}(t_k\text{-}1)$, which is given in the following equation (1).

$$s_{tri}(t_k\text{-}1, x, y) = \begin{cases} \hat{s}_{tri}(t_k\text{-}1, x, y), & s_e(t_k\text{-}1, x, y) = m \\ s_e(t_k\text{-}1, x, y), & \text{otherwise} \end{cases} \tag{2}$$

In the third component, only the remaining marked samples in $s_{tri}(t_k\text{-}1)$ are reconstructed by a *spatial interpolation* (e.g. bilinear or Wiener filter). The result is the upsampled image $s'_u(t_k\text{-}1)$.

At the first time instance t_k=1, where no previous image $s_{tri}(t_k\text{-}2)$ is available, $s_{tri}(t_k\text{-}1)$ is initialized with $s_e(t_k\text{-}1)$.

Because of the recursive structure of the filter, image $s_{tri}(t_k\text{-}1)$ is filled step by step, reducing the number of marked samples. Therefore the samples of $s_{tri}(t_k\text{-}1)$ can remain for many time instances in the image memory. To prevent that inaccurate samples remain to long in the image memory, they are removed after a certain time.

3. EXPERIMENTAL RESULTS

For the verification of the new TRI-Filter a synthetic and a real video test sequence are used. The synthetic test sequence is called *SynWaves* and was created by sampling an analog signal which is composed of two sinusoidal waves. In figure 3a the two analog sinusoidal waves and their samples are shown in dependence of the x-coordinate. The frequency of the

a)

b)

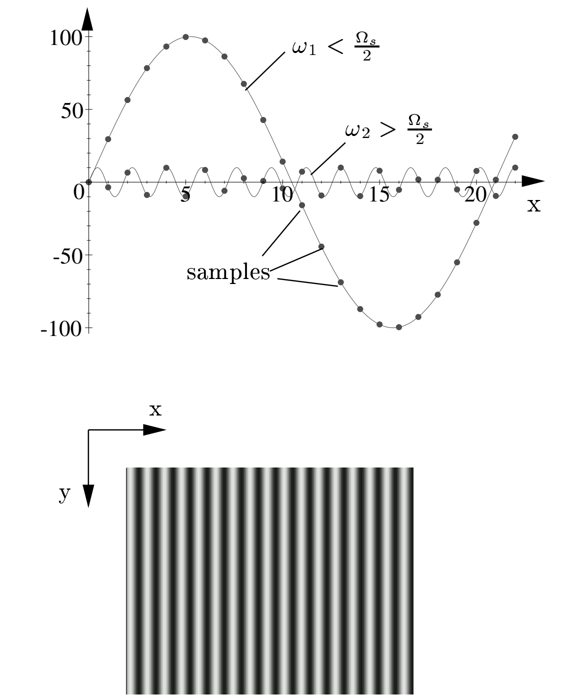

Figure 3: Synthetic Sequence SynWaves: a) Sampling of two sinusoidal waves for the creation of the sequence. b) Exemplary image of the sequence.

first wave ω_1 is below and the frequency of the second wave ω_2 is above the Nyquist frequency $\Omega_s/2$. Therefore the second wave generates aliasing in the sampled signal. Furthermore the analog signal is translational moved with an halfpel displacement between successive frames in order to synthesize a displacement that can be estimated correctly by the used encoders. Figure 3b shows one image of this synthetic test sequence, which is called SynWaves.

The developed TRI-Filter is integrated in two different, available MPEG-4 codecs. One with a displacement resolution of halfpel (L=2) which uses a bilinear interpolation [8] and one with quarter-pel displacements (L=4) [7]. The quarter-pel version uses a Wiener Filter, which was developed to compensate the aliasing signal in the prediction process. Table 1 shows the

Name of Codec	MV res.	Interp. Filter
MPEG4-hp	1/2 pel	bilinear
MPEG4-qp	1/4 pel	Wiener
TRI-hp	1/2 pel	TRI-Filter
TRI-qp	1/4 pel	TRI-Filter

Table 1: Used original and modified MPEG-4 codecs for the experimental verifications.

four different codecs with their motion vector resolution (MV res.) and their used interpolation filter.

In figure 4 the results for the synthetic test sequence SynWaves are shown. Although the image signal is

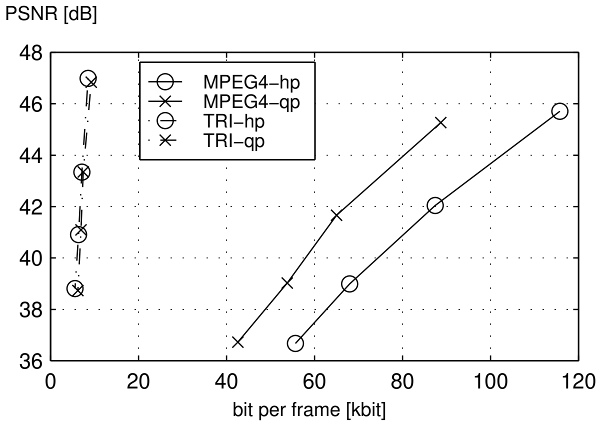

Figure 4: Operational rate-distortion curves obtained from the synthetic sequence SynWaves (SIF format).

displaced by half a pel between successive images, the conventional MPEG4-Codecs can not predict the signal perfectly. Even the quarter-pel version with the Wiener Filter can not compensate the aliasing sufficiently. The rates obtained with the developed TRI-Filter confirm, that the TRI-Filter can predict the aliasing affected input image perfectly. The remaining errors result exclusively from the quantization errors of the intraframe encoder.

In figure 5 the results for the real test sequence *mobile & calendar* (SIF format) are shown. Although the gain is not as large as with the synthetic test sequence, there is a significant gain when using the TRI-Filter. The reason for the decreased gain is, that the motion compensated samples in $\hat{s}_{tri}(t_k-1)$ are affected by displacement estimation errors and errors caused by an

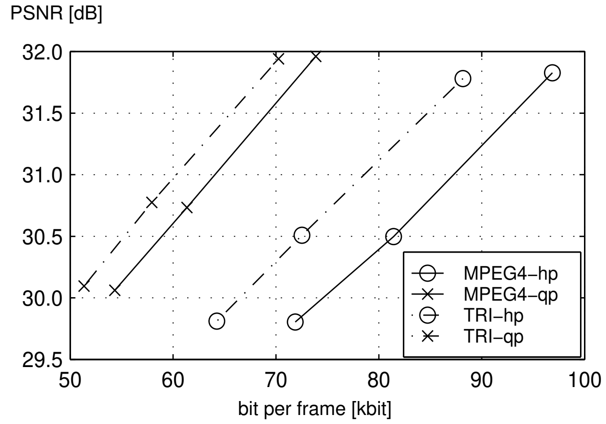

Figure 5: Operational rate-distortion curves obtained from the test sequence mobile & calendar (SIF format).

inaccurate translational motion model. Nevertheless a gain of 0.8 dB can be obtained in case of half-pel motion vectors and bilinear interpolation in the MPEG4 Codec. In case of quarter-pel motion vectors and with a Wiener Filter in the MPEG4 Codec a gain of 0.3 dB is attained. The gain could be enhanced by using better motion models in order to create a more accurate image $\hat{s}_{tri}(t_k - 1)$.

4. CONCLUSIONS

A new time-recursive interpolation filter called TRI-Filter for motion compensated prediction in an hybrid video codec is presented. The filter has been developed to interpolate an aliasing affected image and uses not only the image signal of the current frame, but also the image signal of previous frames in order to interpolate the current frame. Therefore the motion vector field is used in order to use fullpel samples of previous frames to reconstruct subpel samples of the current frame.

The TRI-Filter was tested using a synthetic and real sequence.The results for the synthetic sequence show, that conventional spatial interpolation filters as they are used in actual codecs can not interpolate an aliasing affected image perfectly, even when the motion model and the displacements perfectly fit to the sequence. In contrast to this, the results for the synthetic sequence confirm, that the new TRI-Filter interpolates an aliasing affected image perfectly, when the used motion model and the displacements perfectly fits to the sequence.The results for the real sequence show, that the TRI-Filter improves the interpolation result and the coding efficiency by about 0.8 dB PSNR compared to an MPEG-4 codec.

5. REFERENCES

[1] A.V. Oppenheim, R.W. Schaefer, "Discrete-Time Signal Processing", Prentice Hall, ISBN 0-13-216292-X, 1989.

[2] A.M. Tekalp, "Digital Video Processing", Prentice Hall, ISBN 0-12-190075-7, 1995.

[3] T. Ebrahimi, M. Kunt, "Visual Data Compression for Multimedia Applications", Proceedings of IEEE, Vol. 86, No. 6, June 1998.

[4] H.G. Musmann, P. Pirsch, H.J. Grallert, "Advances in picture coding", Proceedings of IEEE, Vol. 73, No. 4, S.523-548, April 1985

[5] B. Girod, "Motion-Compensating Prediction with Fractional-Pel Accuracy", IEEE Transactions on Communications, Vol. 41, No. 4, April 1993.

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

[7] MPEG-4: ISO/IEC, "Description of Core Experiments on coding efficiency in MPEG-4 video", Doc. ISO/IEC/JTC1/SC29/WG11 N1810, July 1997.

[8] MPEG-4: ISO/IEC, "Final Committee Draft for ISO/IEC 14496-2",
Doc. ISO/IEC/JTC1/SC29/WG11 N2202, March 1998.