SPATIO-TEMPORAL FADING SCHEME FOR ERROR CONCEALMENT IN BLOCK-BASED VIDEO DECODING SYSTEMS

Markus Friebe and André Kaup

Chair of Multimedia Communications and Signal Processing,
University Erlangen-Nuremberg, Cauerstr. 7, 91058 Erlangen
{friebe,kaup}@LNT.de

ABSTRACT

In this contribution we propose a new spatio-temporal fading scheme for block loss recovery in block-based video decoding systems. Based on a boundary error criterion obtained from temporal error concealment, either spatial, temporal, or fading of both methods is used for recovering lost image samples in one macroblock. The weights for fading are interpolated from the boundary error. A weighted absolute difference between well received macroblock boundary samples from the current frame and motion compensated macroblock boundary samples from the previous frame represents the boundary error. It is shown, that in case of transmission errors this method can successfully be used for block loss recovery.

Index Terms— Error concealment, video restoration, video coding and transmission

1. INTRODUCTION

When transmitting block-based coded video over errorprone networks, packets may be lost in case of transmission errors. Due to entropy coding, image information is lost until a synchronization marker is reached. Here, visually annoying block artifacts are introduced in decoded video frames. Generally, restoration techniques are either spatial to conceal Intra-coded frames, temporal or spatio-temporal to conceal Inter-coded frames.

The Boundary Matching Algorithm (BMA) [1] recovers a motion vector of the lost macroblock by using the information of surrounding errorfree received motion vectors. This is a temporal technique. The Decoder Motion-Vector Estimation Algorithm (DMVE) [2] is also a temporal technique and uses image samples, which encircle the lost macroblock. Using these encircled image samples, the DMVE is looking for the best match in the previous frame. H.264 Intra [3] is using surrounding errorfree or concealed image samples for spatial linear interpolation. A content-based adaptive spatio-temporal method (CABLR) [4] is using temporal image information for macroblock loss recovery, if the temporal information fits well. Otherwise correctly received or concealed spatial neighboring macroblocks are used to recover a lost macroblock. Finally a range constraint is applied on the spatially recovered macroblock.

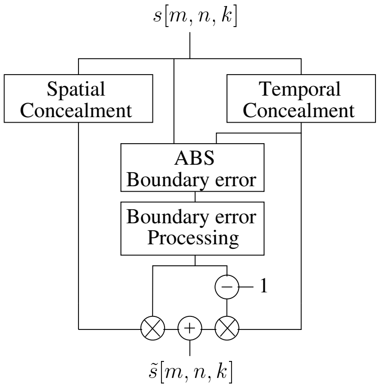

Fig. 1. Block diagram of the spatio-temporal fading scheme.

H.264 Intra introduces blurred image areas, because of spatial linear interpolation. BMA and DMVE are both temporal methods and introduce block artifacts in case of object occlusions and uncovering. CABLR conceals either temporal or spatial and through this switching scheme, either blurred image areas or block artifacts are visible. In this contribution we propose a new spatio-temporal fading scheme, which can be used for block loss recovery in Intra- and Inter-coded frames. This fading scheme combines the result obtained from spatial with that obtained from temporal error concealment. Using this fading scheme, block artifacts can be further reduced while keeping the reconstructed image area relatively sharp. The proposed spatio-temporal fading scheme will be further described in the following section.

2. SPATIO-TEMPORAL FADING SCHEME

A block diagram of the spatio-temporal fading scheme is shown in Fig. 1. First temporal and spatial concealment estimates the lost image samples in one macroblock. Descriptions about these methods are given in section 2.1 and 2.2. Based on the result of temporal concealment, a boundary error is computed. A post-processing unit computes a weighting matrix for the missing macroblock. Using these weighting matrix,

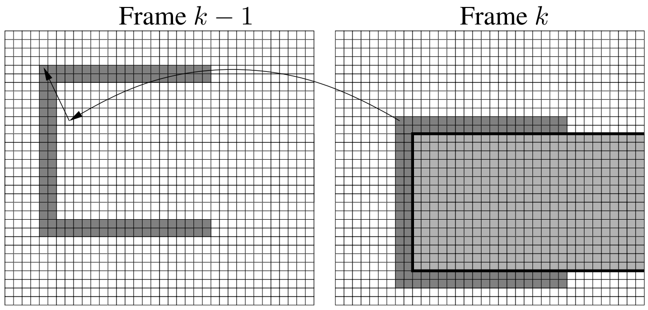

Fig. 2. Left: Frame $k-1$. Right: Frame k. The dark-gray marked area represents the $SAD_w[mv_m, mv_n]$ evaluation area with $w[m,n] = 1$. The light-gray area represents the lost image samples with $w[m,n] = 0$. Both for the left most consecutive macroblock loss.

pelwise weighted mean values between spatial and temporal error concealment results are computed. This post-processing unit is further described in section 2.3.

2.1. Temporal Error Concealment

The motion vector of the lost macroblock $\tilde{mv}_m$ and $\tilde{mv}_n$ is estimated similar to DMVE [2] by minimizing a weighted sum of absolute differences $SAD_w[mv_m, mv_n]$. Eq. (1) and Eq. (2) show, how $SAD_w[mv_m, mv_n]$ is calculated and how the motion vector is estimated from the minimum $SAD_w[mv_m, mv_n]$. Image samples s in the current frame k are compared to motion compensated image samples in the previous frame $k-1$. A weighting function represents correct ($w[m,n] = 1$) and incorrect ($w[m,n] = 0$) received image samples in the current frame k. In Fig. 2, the dark-gray marked area represents correct and the light-gray area the incorrect received image samples for the left most consecutive macroblock loss. For the right most consecutive macroblock loss, the dark-gray marked area is flipped in horizontal direction. The small arrow in Fig. 2 shows one of the evaluated motion vectors. The size of a lost macroblock is given by M and N in row and column direction respectively. m_0 and n_0 represents the left top image sample in the lost macroblock.

$$\begin{aligned} SAD_w[mv_m, mv_n] = \\ \sum_{m=m_0-2}^{m_0+M+1} \sum_{n=n_0-2}^{n_0+N+1} w[m,n] \cdot |s[m,n,k] \\ -s[m+mv_m, n+mv_n, k-1]| \end{aligned} \tag{1}$$

and

$$[\tilde{mv}_m, \tilde{mv}_n] = \arg \min_{mv_m, mv_n} (SAD_w[mv_m, mv_n]) \tag{2}$$

The difference between DMVE and this method is the weighting function. DMVE treats concealed areas as correctly received image samples by setting $w[m,n] = 1$. In this approach, concealed areas are lower weighted by setting $w[m,n] = 0.3$ because concealed areas are estimated and not perfectly reconstructed. After estimating the lost motion vector, the temporal error concealment result represents the image samples

$$s_{temp}[m,n] = s[m_0 + m + \tilde{mv}_m, n_0 + n + \tilde{mv}_n, k-1] \tag{3}$$

with $m = 0, ..., M-1 \wedge n = 0, ..., N-1$ in the previous frame $k-1$.

2.2. Spatial Error Concealment

For spatial error concealment, a frequency selective signal extrapolation method [5] is used for block loss recovery. This method is extrapolating image samples from surrounding well decoded image samples to the lost macroblock. The known surrounding image samples within a specific area $\mathcal{A}$ are approximated by a set of 2D-DFT basis functions. Such basis functions $\varphi_{u,v}[m,n]$ are selected and scaled by $c_{u,v}$ to minimize an error criterion. The error criterion is computed between the sum of the selected basis functions $g[m,n]$ and the surrounding well received image samples within the specific area. u and v are indices for basis functions in row and column direction.

$$g[m,n] = \sum_{(u,v \in \mathcal{A})} c_{u,v} \varphi_{u,v}[m,n] \tag{4}$$

This is an iterative method and in each iteration step, one 2D-DFT basis function is selected and scaled. The iteration stops, if the error criterion is smaller than a given threshold or a maximum number of iterations is reached. The area $\mathcal{L}$ within $g[m,n]$, which represents the lost macroblock is used for spatial error concealment.

$$s_{spatial}[m,m] = g[i,j] \tag{5}$$

with $m = 0, ..., M-1 \wedge n = 0, ..., N-1 \wedge i, j \in \mathcal{L}$.

2.3. Boundary error processing and Interpolation

In Fig. 2, the boundary error $e[m,n]$ is represented by the weighted absolute difference between the dark-gray marked area in frame k and the motion compensated dark-gray marked area in frame $k-1$. The weights $w[m,n]$ for computing the boundary error are the same as in section 2.1. The boundary error measures pelwise, if temporal error concealment fits well to the lost macroblock area.

$$\begin{aligned} e[m,n] &= w[m,n] \cdot |s[m,n,k] \\ &\quad -s[m + \tilde{mv}_m, n + \tilde{mv}_n, k-1]| \end{aligned} \tag{6}$$

with $m = m_0 - 2, ..., m_0 + M + 1 \wedge n = n_0 - 2, ..., m_0 + N + 1$. In Fig. 3 an example for the boundary error $e[m,n]$ is shown. At the bottom macroblock border temporal concealment fits relatively well, but not at the top right macroblock corner. After computing the boundary error, this error is low-

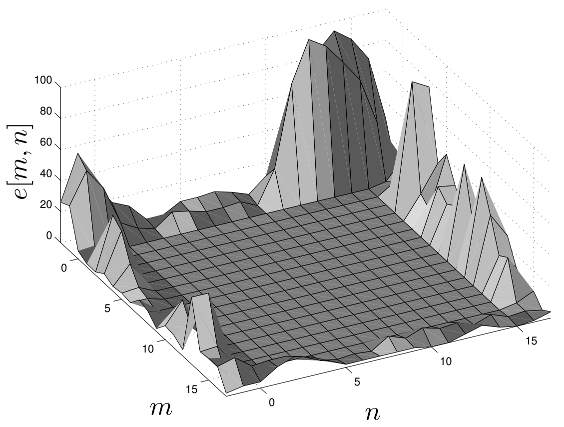

Fig. 3. Boundary error $e[m,n]$ for a 16×16 block loss.

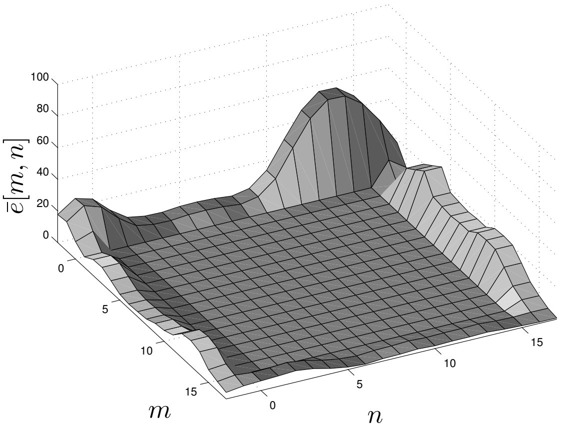

Fig. 4. Mean boundary error $\bar{e}[m,n]$.

pass filtered to avoid high frequencies in the resulting weighting matrix for fading between temporal and spatial error concealment. A lowpass filter computes a mean boundary error $\bar{e}[m,n]$ within a 3×3 window along the dark-gray marked area in Fig. 2.

$$\bar{e}[m,n] = \frac{1}{\sum_{i=-1}^{1}\sum_{j=-1}^{1} w[m-i,n-j]} \cdot \sum_{i=-1}^{1}\sum_{j=-1}^{1} e[m-i,n-j] \quad (7)$$

For filtering symmetrically boundary extensions are used. The mean boundary error $\bar{e}[m,n]$ can be seen in Fig. 4. The area $\mathcal{L}$ represents the area ($m = 0, ..., M-1 \wedge n = 0, ..., N-1$) encircled by the lowpass filtered boundary error. $\mathcal{L}$ is then bilinear interpolated from the nearest $\bar{e}[m,n]$ neighbors, outside of $\mathcal{L}$, in vertical and horizontal direction. The distances to the nearest neighbors represent the weights for bilinear interpolation. In the following, the resulting matrix $\tilde{w}[m,n]$ of size $M \times N$ from bilinear interpolation is then compared to a

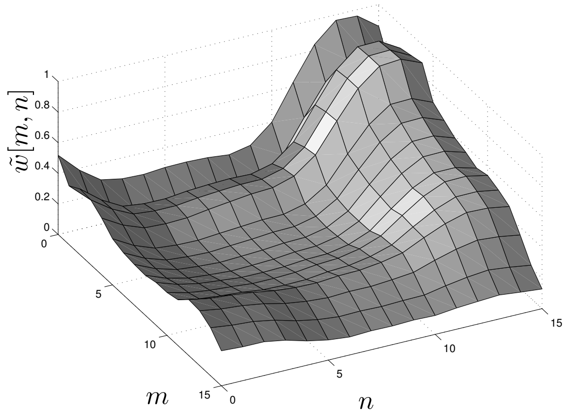

Fig. 5. Weighting matrix $\tilde{w}[m,n]$.

threshold T_w and scaled by T_w.

$$\tilde{w} = \begin{cases} 1 & , \quad \tilde{w} > T_w \\ \frac{\tilde{w}}{T_w} & , \quad \text{else} \end{cases} \quad (8)$$

In Fig. 5 the weighting matrix $\tilde{w}[m,n]$ for the boundary error in Fig. 3 is shown. At the bottom macroblock border, the weights are relatively low and temporal concealment is higher weighted than spatial. Here, temporal concealment fits relatively well. Whereas, at the top right macroblock corner, temporal concealment does not fit well, the weights are relatively high and spatial concealment is stronger weighted. The lost macroblock $\tilde{s}[m,n]$ is estimated from spatial $s_{spatial}[m,n]$ and temporal $s_{temp}[m,n]$ error concealment by

$$\begin{aligned} \tilde{s}[m,n] &= (1-\tilde{w}[m,n]) \cdot s_{temp}[m,n] \\ &\quad + \tilde{w}[m,n] \cdot s_{spatial}[m,n]. \end{aligned} \quad (9)$$

$\tilde{s}[m,n]$ is estimated using a pelwise weighted mean value between spatial $s_{spatial}[m,n]$ and temporal $s_{temp}[m,n]$ error concealment. Finally, the weighting function $w[m,n]$ is updated within the concealed area.

3. SIMULATION RESULTS

For simulations we consider the case that consecutive macroblocks are lost in Intra-coded frames. Let N_{MB} represent the number of macroblocks in one row and assume, that the previous frame $k-1$ was decoded without errors. Synchronization markers are inserted at macroblock column positions $N_{MB}-2$ and transmission errors appears in each third row at macroblock column position 3. Best objective and subjective results has been achieved using $T_w = 35$ as threshold and $w[m,n] = 0.3$ for weighting the concealed areas. For BMA the necessary surrounding motion vectors has to be precomputed and DMVE uses a two line encirclement. The colorspaces Y, U and V are filtered separately for all methods. For DMVE, CABLR and the proposed method, the motion vector is only estimated for the luminance and further used

Table 1. Mean Y PSNR results in dB for frames 1-100

Video	Fore-man	Basket-ball	Sales-man	Flower garden	Coast-guard
H.264 Intra [3]	21.99	15.93	22.26	15.44	17.74
Zero MV	27.02	15.60	32.91	16.48	22.59
BMA [1]	30.76	20.27	33.41	23.31	26.32
DMVE [2]	32.75	21.16	32.60	26.51	**28.46**
CABLR [4]	33.15	20.58	32.63	25.86	27.68
Proposed	**33.51**	**22.37**	**33.87**	**27.14**	28.28

Fig. 6. Left: Original image. Right: Consecutive 16×16 block loss.

for the chrominances. If the luminance in CABLR is recovered spatially, then also the chrominances are recovered spatially. In Table 1 mean luminance PSNR results for different videos and error concealment methods are shown. The PSNR values are only evaluated at the lost areas. Spatial methods like H.264 Intra [3] yields the lowest PSNR values. CABLR performs well in video sequences with less motion and homogeneous image areas like foreman. DMVE yields the highest results for the coastguard sequence. The proposed method on average achieves the highest PSNR results over all video sequences. In Fig. 6, a part of a basketball sequence frame is shown without errors at the left hand side and with consecutive macroblock loss at the right hand side. The results for H.264 Intra are shown at the left hand side and DMVE at the right hand side in Fig. 7. H.264 Intra is introducing blurred image areas because of the linear interpolation. DMVE introduces some block artifacts in case of object occlusions and uncovering as seen at the arm or the head in the upper image area. Also CABLR is introducing some block artifacts as seen in Fig. 8 at the left hand side. At the referee's head, block artifacts are visible. The result for the proposed method is shown in Fig. 8 at the right hand side. In case temporal concealment does not perform well like at the referee's head, spatial concealment is more weighted than temporal. Here block artifacts are reduced.

4. CONCLUSION

We proposed a spatio-temporal fading scheme for error concealment in block-based coded video decoding systems. The simulation results have shown that the proposed method outperforms the reference methods both in subjective and objective video quality. The proposed method is therefore a good candidate for block loss recovery in Intra- and Inter-coded video frames.

Fig. 7. Left: H.264 Intra [3]. Right: DMVE [2].

Fig. 8. Left: CABLR [4]. Right: Proposed.

5. REFERENCES

[1] W.-M. Lam, A-R. Reibman, and B. Liu, "Recovery of Lost or Erroneously Received Motion Vectors," in *Proc. Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, Minneapolis, April 1993, pp. V417–V420.

[2] J. Zhang, J. F. Arnold, and M. R. Frater, "A Cell-Loss Concealment Technique for MPEG-2 Coded Video," in *IEEE Trans. Circuits Syst. Video Technol.*, June 2000, vol. 10, pp. 659–665.

[3] Y.-K. Wang, M. M. Hannuksela, and V. Varsa, "The Error Concealment Feature in the H.26L Test Model," in *Proc. Int. Conf. on Image Processing (ICIP)*, Rochester, 2002, pp. 729–732.

[4] J. Park, R. J. Marks D.-C. Park, and M. A. El-Sharkawi, "Content-Based Adaptive Spatio-Temporal Methods for MPEG-Repair," in *IEEE Trans. Image Process.*, August 2004, vol. 13, pp. 1066–1077.

[5] A. Kaup, K. Meisinger, and T. Aach, "Frequency selective signal extrapolation with applications to error concealment in image communication," *Int. J. Electron. Commun. (AEÜ)*, vol. 59, pp. 147–156, May 2005.