INTER-SEQUENCE ERROR CONCEALMENT OF HIGH-RESOLUTION VIDEO SEQUENCES IN A MULTI-BROADCAST-RECEPTION SCENARIO

Tobias Tröger, Jens-Uwe Garbas*, Henning Heiber†, Andreas Schmitt† and André Kaup**

*Multimedia Communications and Signal Processing
University of Erlangen-Nuremberg
Cauerstr. 7, 91058 Erlangen, Germany
{troeger, garbas, kaup}@lnt.de

†Development Infotainment
Audi AG
85045 Ingolstadt, Germany
{henning.heiber, andreas.schmitt}@audi.de

ABSTRACT

In this paper, a new approach is proposed for the concealment of lost samples in a high-resolution, high-quality video utilizing a low-resolution, highly compressed video with equal content. It is shown, that this inter-sequence error concealment is a very robust and flexible technique which outperforms conventional methods even for low bit rates. By adopting an affine motion model, the proposed technique also performs well in case of different image sizes, cropped image content and arbitrarily shaped loss area. An optimization method limits the computational complexity and maximizes the image restoration quality. If two or more low-resolution reference video sequences are available, the algorithm can be easily expanded. A typical application for inter-sequence error concealment is the restoration of a DVB-T video sequence in a terrestrial multi-broadcast-scenario with DVB-T, DVB-H and T-DMB.

1. INTRODUCTION

Currently, various broadcasting techniques are deployed worldwide and there are still more to come. "Digital Video Broadcasting - Terrestrial" (DVB-T) is a digital broadcasting system for terrestrial transmission of SDTV or HDTV video, audio streams and data [1]. An IP-based transmission of low-resolution digital multimedia is provided by "Digital Video Broadcasting - Handheld" (DVB-H) which can be seen as an adapted version of the DVB-T technique especially optimized for mobile reception conditions [2]. An extension of the well-known "Digital Audio Broadcasting Standard" (DAB) for the additional transmission of digital, low-resolution TV signals is called "Terrestrial Digital Multimedia Broadcasting" (T-DMB) [3].

In a typical broadcasting scenario, compressed and packetized video signals are transmitted over error-prone channels. As a result, packet errors occur at receiver-side which can be expressed by the corresponding packet error rate (PER). Utilizing a two-stage combined channel coding and interleaving scheme as for example the DVB-T standard does, the PER can be reduced significantly after transmission. However, as soon as the limit of the employed channel code is reached, no further error correction can be achieved. Therefore, the decoded video signal is degraded with macroblock or slice losses due to the block-based coding principle of hybrid video coders.

State-of-the-art error concealment techniques predict the lost image information from temporal or spatial or both temporal and spatial neighboring pixels. We call these methods in general intra-sequence error concealment techniques (IASEC).

Lost motion vectors can be recovered with the "Boundary Matching Algorithm" (BMA) by using the information of surrounding error-free received motion-vectors [4]. If both motion vector and corresponding prediction error of a macroblock are lost, an extended version of BMA, which is called EBMA, additionally assumes the prediction error from neighboring blocks if available [4]. The "Decoder Motion-Vector Estimation Algorithm" (DMVE) minimizes the difference between surrounding image samples of a lost macroblock and those of the candidate block in the preceding frame also by utilizing a matching principle [5]. All three mentioned algorithms are temporal techniques. "H.264 Intra" is a spatial error concealment technique which uses surrounding error-free or concealed image samples of the lost image area for weighted pixel averaging [6]. In our simulations we consider BMA, DMVE and H.264 Intra as reference intra-sequence concealment techniques.

Considering a future multi-broadcast-reception scenario, two or more video signals with equal image content may be available at receiver-side. A typical application scenario would be the reception of both DVB-T and DVB-H or both DVB-T and T-DMB. In this case, the specific transmission properties of each broadcasting technique lead to differences in spatial image resolution, image quality and degree of distortion. Typically, broadcasting techniques for mobile reception as for example DVB-H or T-DMB apply low spatial image resolutions whereas DVB-T sequences usually have a higher resolution. We distinguish between high-resolution sequences (HRS) and corresponding low-resolution reference sequences (LRRS) in the following. The utilized video coding standards and the compression factors define the particular image quality of both video sequences. Taking into account the image quality, the LRRS is typically coded with a low bit rate and therefore has moderate image quality when it is transmitted in a DVB-H or T-DMB network. The HRS, however, has a high image quality as it is displayed at large screens. Finally, the amount and distribution of distorted image samples in the compared video sequences depends on the deployed error protection schemes and the underlying channel characteristics. The LRRS is considered as error-free in our case as in the given scenario it is better protected against transmission errors as the HRS.

Based on the characterized scenario, we show in our work how lost macroblocks or slices of a high-resolution video sequence can be concealed utilizing a perfectly synchronized reference video sequence with error-free image content but low spatial resolution. In contrast to the intra-sequence case, we call this technique inter-sequence error concealment (ISEC).

2. INTER-SEQUENCE ERROR CONCEALMENT

2.1 Image Matching Procedure

First, we introduce the general procedure for the inter-sequence error concealment of a high-resolution video sequence using a low-resolution reference sequence. In order to process corresponding frames, synchronization is required as a precondition. Here, we assume both video sequences being synchronized. The proposed algorithm contains three main steps and can be applied to video sequences which have equal image content but differ in spatial resolution and in image quality due to the given transmission parameters. Within a certain range, even the image content of corresponding frames can vary. Hence, the crucial point of this approach is its generality.

Let us consider a high-resolution frame $A(m,n)$ and a corresponding low-resolution reference frame $B(r,s)$ with equal content, where $m \in \{1,...,M\}$, $n \in \{1,...,N\}$, $r \in \{1,...,R\}$ and $s \in \{1,...,S\}$ depict the pixel positions ($M > R$, $N > S$). As the image content of both frames is similar, frame $B(r,s)$ can be understood as a projection of frame $A(m,n)$. In the first step of the proposed algorithm,

this projection shall be inverted by mapping image $B(r,s)$ onto image $A(m,n)$. As the exact reprojection properties are unknown in general, optimal mapping is aggravated. Therefore, we adopt an affine motion model in our work for parameterization of the image transformation process.

In [7], the affine motion model is characterized which is well-known in video processing for global motion estimation (GME) [8]. By adopting the motion model, an affine image transformation can include translation, rotation, zoom, scaling and shear. The positions m and n of a transformed image are based on the original coordinates r and s, where $a_1, ..., a_6$ are the transformation parameters.

$$m = a_1 \cdot r + a_2 \cdot s + a_3 \tag{1}$$

$$n = a_4 \cdot r + a_5 \cdot s + a_6 \tag{2}$$

As the exact properties of the original image transformation process of frames $A(m,n)$ and $B(r,s)$ are unknown, we have to consider the relevant affine model parameters. Frame $B(r,s)$ is either supposed to be a non-uniformly scaled version, a non-uniformly scaled as well as truncated version or uniformly scaled as well as truncated version of frame $A(m,n)$. Only the latter case guarantees that no perspective distortions occur. Then, the aspect ratio of the image content is kept constant. Let us discuss the edge cases: In case of exclusive non-uniform scaling (case 1), the image transformation can be defined only by parameters a_1 and a_5. All other parameters are zero ($a_2 = 0$, $a_3 = 0$, $a_4 = 0$, $a_6 = 0$). In case of a truncated projection (case 2), however, we have translation and uniform or non-uniform scaling. Therefore, we need a four parameter model consisting of a_1, a_3, a_5, and a_6 ($a_2 = 0$, $a_4 = 0$). As a consequence of a truncated projection, frame $B(r,s)$ does not contain the full image content of frame $A(m,n)$. Here, lost marginal samples in frame $A(m,n)$ which lie outside the reprojection of $B(r,s)$ are concealed according to weighted pixel averaging [9]. As we want our algorithm to work independently of the projection properties and the spatial resolution of $B(r,s)$, the four parameter model is always used. Then, $\mathbf{a} = [a_1, a_3, a_5, a_6]^T$ denotes the transformation parameters which are unknown and shall be best-fitted.

The formal definition of a reprojection of $B(r,s)$ depending on the model parameters $\mathbf{a}$ of a transformation process is given by

$$\tilde{B}_a(m,n) = \mathscr{R}\{B(r,s)|\mathbf{a}\}, \tag{3}$$

where $\tilde{B}_a(m,n)$ is the reprojection candidate. The best-fitting candidate $\tilde{B}_a^{best}(m,n)$ has to be determined which maximizes the restoration quality in the concealed image parts of frame $A(m,n)$. As the image information of distorted samples is lost, our criterion for optimal reprojection is the mean of squared errors (MSE) between the correctly received image samples of the reference image $A(m,n)$ and the corresponding samples of the reprojection candidate $\tilde{B}_\mathbf{a}(m,n)$.

$$MSE(\mathbf{a}) = \frac{1}{\sum_{m=1}^{M}\sum_{n=1}^{N} W(m,n)} \sum_{m=1}^{M} \sum_{n=1}^{N} W(m,n) \left(A(m,n) - \tilde{B}_a(m,n)\right)^2 \tag{4}$$

Error mask $W(m,n)$ denotes a binary matrix which defines if a particular image sample is used for the optimization or not. A sample is valid if it is not distorted in frame $A(m,n)$ and if it is element of the projection candidate $\tilde{B}_a(m,n)$.

$$W(m,n) = \begin{cases} 1 & \text{if sample is valid} \\ 0 & \text{else} \end{cases} \tag{5}$$

The final minimization problem for the determination of the best-fitting affine model parameters $\mathbf{a}_{best}$ out of a set $\mathbb{D}$ can be formulated as

$$\mathbf{a}_{best} = \underset{\mathbf{a} \in \mathbb{D}}{\operatorname{argmin}}(MSE(\mathbf{a})). \tag{6}$$
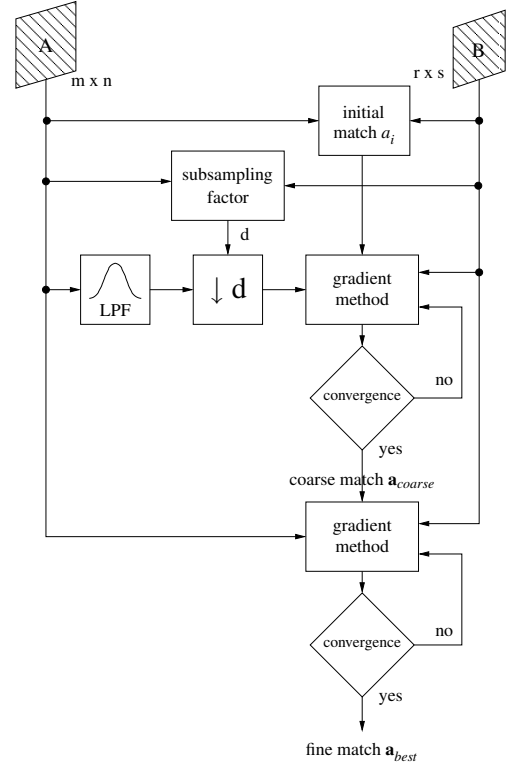

Figure 1: Parameterization of image transformation

As the proposed algorithm for inter-sequence concealment is supposed to work independently of the image resolution used in the LRRS, set $\mathbb{D}$ can not be constrained without impairing the image restoration quality. Due to the high computational complexity of a full search in $\mathbb{D}$, a gradient method is utilized for optimization of $\mathbf{a}$. We decided in favor of the Levenberg-Marquardt approach which is quite robust even in case of inaccurate start values [10].

Fig. 1 shows the framework for the affine model parameterization of the reprojection. Starting with an inital match $\mathbf{a}_i$, the gradient method is performed in two stages. We define the initial match $\mathbf{a}_i = [a_1^i, a_3^i, a_5^i, a_6^i]^T$ as the image size ratios for scaling combined with zero translation ($\mathbf{a}_i = [\frac{M}{R}, 0, \frac{N}{S}, 0]^T$). The first stage is applied to a subsampled version of frame $A_d(m,n)$ where d is the subsampling factor. It can be set for example as the minimum of the aspect ratios which are rounded down.

$$d = \min(\lfloor \frac{M}{R} \rfloor, \lfloor \frac{N}{S} \rfloor) \tag{7}$$

For two-dimensional low-pass filtering of the decimation step, a simple kernel can be used. Of course, the binary error mask $W(m,n)$ has to be decimated in the same way as frame $A(m,n)$. If convergence is reached in the first optimization step, we get a coarse match a_{coarse}. By refining a_{coarse} instead of the initial match $\mathbf{a}_i$ in the second stage, convergence is improved and the overall computational complexity is reduced. The second stage is performed on the full-resolution image $A(m,n)$ and leads to the best-fitting transformation parameters $\mathbf{a}_{best}$.

In the second step of the proposed algorithm for ISEC, image $B(r,s)$ is finally transformed to $\tilde{B}_a^{best}(m,n)$ with the determined model parameters $\mathbf{a}_{best}$ (6). In other words, image $B(r,s)$ is upsampled, low-pass filtered and if neccessary shifted. To achieve optimal reprojection, the same low-pass filter as in the optimization process has to be applied for the final image transformation. We use linear interpolation, here.

Finally, the distorted image samples of frame $A(m_e,n_e)$ are replaced in the third step of the algorithm with corresponding image samples of the transformed image $\tilde{B}_a^{best}(m_e,n_e)$, where m_e and n_e

denote the position of erroneous samples.

The proposed algorithm is independant of the shape of lost image areas and can be applied for each frame of a HRS separately. To minimize the computational complexity, the transformation parameters can be recalculated only with a fixed period in time. If necessary, outliers can be discarded by introducing a threshold which defines the maximum deviation from the temporal mean of the transformation parameters. Outliers can occur if the degree of distortion is too high for a particular frame or the initial match is inadequate.

2.2 Gradient Method

Determining the image transformation parameters by minimizing $MSE(\mathbf{a})$ is a non-linear optimization problem which can be solved by a gradient method. In [10], it is shown that any function $f(\mathbf{x})$ can be approximated by its truncated Taylor series. (8) depicts the general approximation of $f(x)$ truncated after the second derivative, where $\mathbf{P}$ denotes the origin of the coordinate system.

$$f(\mathbf{x}) \approx f(\mathbf{P}) + \sum_i \frac{\partial f}{\partial x_i}\bigg|_{\mathbf{P}} \cdot x_i + \frac{1}{2}\sum_{i,j} \frac{\partial_2 f}{\partial x_i \partial x_j}\bigg|_{\mathbf{P}} \cdot x_i x_j \qquad (8)$$

The approximation can be rewritten in our case as follows in (9) where $\mathbf{d}$ is the negative gradient $-\nabla MSE(\mathbf{a})|_{\mathbf{a}_i}$ and $\mathbf{D}$ the Hessian matrix of $MSE(\mathbf{a})$ at $\mathbf{a}_i$. c is a scalar and denotes $MSE(\mathbf{a}_i)$. $\mathbf{a}$ is a M-vector with the unknown parameters a_k ($k \in \{1,...,M\}$).

$$MSE(\mathbf{a}) \approx c - \mathbf{d}^T \cdot \mathbf{a} + \frac{1}{2}\mathbf{a}^T \cdot \mathbf{D} \cdot \mathbf{a} \qquad (9)$$

In case of a good approximation, the final match $\mathbf{a}_{best}$ immediately follows from the current parameter set $\mathbf{a}_{cur}$ by the "Inverse-Hessian Method" (IHM) in (10). This is equivalent to the gradient tending towards zero at the minimum of $MSE(\mathbf{a})$.

$$\mathbf{a}_{best} = \mathbf{a}_{cur} + \mathbf{D}^{-1} \cdot \mathbf{d} \qquad (10)$$

However, if the approximation is poor, we have to step down along the gradient according to a constant g with the "Steepest Descent Method" (SDM):

$$\mathbf{a}_{next} = \mathbf{a}_{cur} + g \cdot \mathbf{d} \qquad (11)$$

The Levenberg-Marquardt (LM) method uses both IHM and SDM, starting with IHM and fading to the latter one if the convergence basin near the minimum is reached. Optimizing due to LM is more robust than the Gauß-Newton approach [11] and guarantees convergence even for inaccurate start parameters.

According to [10], the IHM in (10) can be rewritten as a set of linear equations, where $\beta_k = -\frac{1}{2}\frac{\partial MSE(\mathbf{a})}{\partial a_k}$, $\alpha_{kl} = \frac{1}{2}\frac{\partial^2 MSE(\mathbf{a})}{\partial a_k \partial a_l}$ and δ is the step size:

$$\sum_{l=1}^{M} \alpha_{kl} \cdot \delta \cdot a_l = \beta_k \qquad (12)$$

Similarly, the SDM in (11) can be rewritten with a non-dimensional factor λ ([10]):

$$\lambda \cdot \alpha_{ll} \cdot \delta \cdot a_l = \beta_l \quad \forall \; l \in \{1,...,M\} \qquad (13)$$

Marqardt found, that (12) and (13) can be combined by introducing a new matrix α' defined by $\alpha'_{jk} = \alpha_{jk}$ for all $j \neq k$ and $\alpha'_{jj} = \alpha_{jj}(1-\lambda)$. This leads to a single formula characterizing the optimization of MSE($\mathbf{a}$):

$$\sum_{l=1}^{M} \alpha'_{kl} \cdot \delta \cdot a_l = \beta_k \qquad (14)$$

For large λ, (14) goes over to (13) so the SDM is applied. In case of λ tending to zero, (14) goes over to (12) near the convergence basin and the IHM is in use. The LM method is iteratively performed in four steps:

Figure 2: Truncated (left) and full (right) projection of sequence *crew* (resolution *QVGA*)

1. Determine $MSE(\mathbf{a})$ with initial match $\mathbf{a}_i$
2. Define a moderate λ (for example $\lambda = 10^{-3}$ in [10])
3. Solve the linear equations given in (14) for $\delta \cdot \mathbf{a}$ and evaluate $MSE(\mathbf{a} + \delta \cdot \mathbf{a})$
4. If the error function has ...
   (a) ... grown (i.e. $MSE(\mathbf{a} + \delta \cdot \mathbf{a}) \geq MSE(\mathbf{a})$), increase λ (e.g. by 10, see [10]) and go back to 3.
   (b) ... declined (i.e. $MSE(\mathbf{a} + \delta \cdot \mathbf{a}) < MSE(\mathbf{a})$), decrease λ (e.g. by 10, see [10]) and set $MSE(\mathbf{a} + \delta \cdot \mathbf{a})$ as new trial solution and go back to 3.

3. SIMULATION RESULTS

In this section, we present simulation results for inter-sequence concealment in a multi-broadcast-scenario of both DVB-T and DVB-H or DVB-T and T-DMB. For the HRS, a resolution of 720x576 pixels is used which is typical for a DVB-T video transmission. In a DVB-H or T-DMB network, the LRRS is supposed to have resolution *CIF* or *QVGA*. All used sequences have progressive format. If necessary, deinterlacing has been applied. As the LRRS is a projection of the HRS, corresponding frames have equal image content. In Section 2.1, we discussed the possible transformation properties. Let us consider the edge cases 1 and 2 again meaning exclusive non-uniform scaling or both uniform scaling and cropping. In case of a truncated projection, image content is lost at the margins. A truncated projection of a frame with 720x576 pixels resulting in resolution *QVGA* can be seen on the left side in Fig. 2 on example of the sequence *crew*. There, about 6% of the image information is lost by cropping which is applied symmetrically at the top and bottom margins (case 2). The lost image parts are marked red in the corresponding full projection (case 1) on the right side of Fig. 2.

The objective restoration quality of the proposed ISEC-algorithm is compared to those of state-of-the-art methods belonging to IASEC. For temporal concealment, BMA and DMVE are taken as reference. The spatial methods are represented by H.264 Intra. As DVB-T sequences are typically compressed with high bit rates, the visual quality is excellent. Therefore, we consider the HRS being uncompressed. This is a best case scenario for IASEC methods and therefore allows a more than fair comparison. The LRRS is compressed with the reference implementation (JM 13.0) of the H.264/AVC standard. To stay DVB-H or T-DMB compliant, we use the main profile for encoding. The bit rate is chosen between 0.001 and 0.400 bit per pixel. The lost samples comprise 5% of each image. The positions of lost macroblocks are randomly chosen.

In Fig. 3 the objective restoration quality of intra- and inter-sequence error concealment is compared in terms of mean luminance $PSNR_Y$ for 40 frames of the sequence *crew*. As can be seen, the proposed algorithm for ISEC outperforms the reference IASEC techniques even for low bitrates used for the LRRS. Only in case of bit rates below 0.001 bpp, the objective image quality for DMVE (green) is higher compared to ISEC. BMA is marked black and achieves about 4.2 dB less $PSNR_Y$ compared to DMVE. Fig. 3 shows the results for both edge cases of the LRRS being available in resolution *CIF* (red) or *QVGA* (blue). In general, the ISEC method depends on the bit rate used for the LRRS whereas IASEC does not. Although, the curves are monotonely increasing for ISEC,
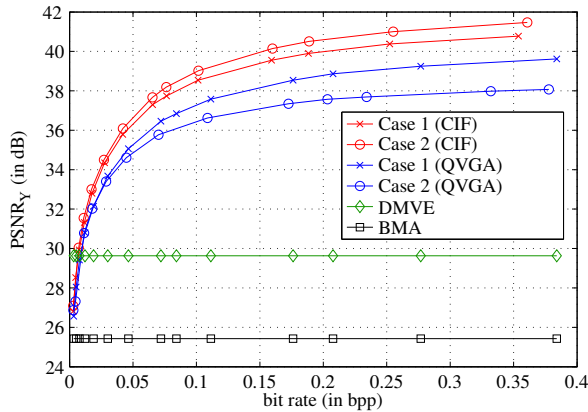
Figure 3: Objective image quality against bit rate of the LRRS for sequence *crew*
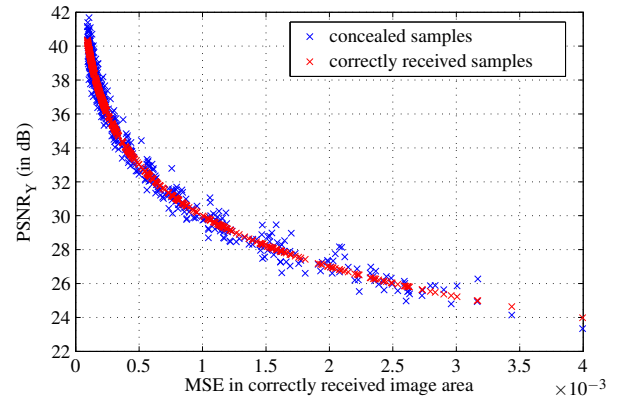

Figure 4: Objective image quality per frame in concealed and correctly received parts for sequence *crew* (resolution *CIF*, case 1)
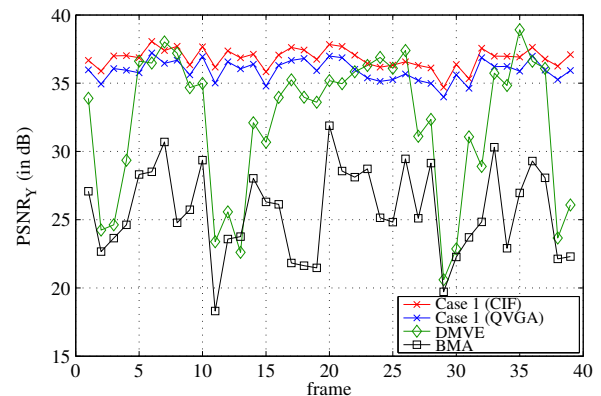

Figure 5: Objective image quality against frame number for sequence *crew* (bit rates: 0.012 bpp for resolution *CIF* and 0.010 bpp for resolution *QVGA*)

here, rate-distortion optimality is not ensured in case of suboptimal image transformation parameters.

The restoration quality of ISEC depends on the spatial resolution of the LRRS as well as the specific projection case (Fig. 3). In general, ISEC performs better based on resolution *CIF* than on *QVGA*. Assuming projection case 2 for both resolutions, ISEC has a gain of up to 3.8 dB in terms of $PSNR_Y$ for CIF. Additionally, it is noticeable that the results of projection case 1 and 2 are complementary in case of LRRS-resolution *CIF* and *QVGA*. The reason can be found in the generation process of the LRRS which influences the final restoration quality of ISEC significantly. In Sec. 2.1, we assumed the LRRS being generated based on LP filtering and subsampling of the HRS. Let the subsampling factor be an integer value. Then, the sample positions of the LRRS are on the original grid of the HRS. In case of non-integer subsampling, however, subpixels are determined for the HRS by interpolation and taken as LRRS samples. This information loss finally leads to a decrease in restoration quality for ISEC. A brief look at both the best and the worst ISEC result in Fig. 3 clarifies this: Matching resolution 720x576 to *CIF* based on uniform scaling (case 2), the horizontal and vertical subsampling factor are integer values, namely 2. Non-uniform scaling (case 1) and LRRS-resolution *QVGA*, however, leads to horizontal/vertical subsampling of the HRS by 2.25 and 2.4. As a result, the objective image quality in the concealed areas is the lowest in comparison to the other three ISEC scenarios.

A further aspect has to be evaluated in connection with the influence of subsampling on the performance of ISEC. Assuming the LRRS as a truncated projection of the HRS with cropped horizontal or vertical image margins (case 2), distorted marginal image samples of the HRS can not be concealed with ISEC. A conventional IASEC method has to be used instead which usually performs worse than ISEC (see Sec. 2.1). This loss in terms of $PSNR_Y$ only occurs for projection case 2 and finally leads to the specific performance scenario of ISEC for sequence *crew* (see Fig. 3). Based on a bit rate of 0.3 bpp for the LRRS, further results for IASEC methods and the proposed inter-sequence error concealment technique can be looked up in Tab. 1 on example of sequences *discovery city* and *rugby*.

Fig. 4 shows high correlation between the objective image quality of correctly received samples (red) and the restoration quality in concealed image parts (blue). It can be seen that the minimization of the mean squared error in the known image areas by a gradient method leads to efficient concealment of the lost image parts. The crucial point of error concealment is the approximation of lost image content without knowing the exact sample values. By minimization the MSE of candidate samples and spatial or temporal neighbored image areas of a lost macroblock as typical IASEC techniques do, the probability of a good approximation for lost samples often is insufficient. Using the proposed algorithm for ISEC, however, the reliability of concealed samples is maximized. This is because the approximation is based on the minimization of the MSE in the whole error-free image area of HRS and LRRS due to (4).

In Fig. 5, the objective image quality is shown for 40 concealed frames of the sequence *crew*. The bitrate of the LRRS for resolution *CIF* is 0.012 bpp (resolution *QVGA*: 0.010 bpp). The $PSNR_Y$ values for BMA (black) and DMVE (green) are highly variant due to the temporal dependance. In case of scene cuts, temporal concealment completely fails whereas inter-sequence error concealment is fully independent of time. As a consequence, error propagation is completely avoided with ISEC. Therefore, ISEC performs at a high level in terms of PSNR with a low variance in time and is not influenced by scene cuts. That means, that temporal concealment methods only perform well for video sequences with static scenes. For some sequences, the results can be superior than those of spatial and inter-sequence error concealment. However, the performance of temporal concealment techniques depends on the degree of distortion of the reference start frame because the error propagates in time. We used a single error-free start frame for BMA and DMVE. In case of block losses in this first frame, the objective video quality in the concealed image areas would decrease significantly for both temporal techniques.

Fig. 6 shows the visual results of ISEC for the sequence *crew*. Based on the error distribution in the distorted HRS frame (Fig. 6(a)), the visual quality can be subjectively evaluated for an underlying mean bit rate of the LRRS of 0.005 bpp (Fig. 6(b)) and 0.258 bpp (Fig. 6(c)). The reference frames in the LRRS have resolution *CIF* and were generated by non-uniform scaling (case 1). So, pure inter-sequence error concealment is applied in contrast to case 2. In terms of $PSNR_Y$ we obtain a high objective quality of 28.17 dB (Fig. 6(b)) and 39.63 dB (Fig. 6(c)) for both bit rates.

By using temporal concealment methods based on block-matching techniques, blockiness is introduced in the concealed images when motion occurs. This holds especially for scene cuts. Applying ISEC, edges are completely avoided in case of effective

optimization of the image transformation parameters and a moderate compression factor of the LRRS. Only for extremely low bit rates, blocking is also introduced by ISEC as the high frequencies are attenuated with increasing compression factor in H.264/AVC. As a consequence, homogeneous blocks are inserted in the concealed high-resolution frame. Also, motion vectors are quantized roughly for extremely high compression with H.264/AVC. Then, the DC value of the concealed blocks can differ from the spatial neighbours as can be partly seen in Fig. 6(b).

We suggested to discard outliers in Section 2.1. This can be a reasonable step when the number of lost macroblocks reduces spatial correlation of corresponding frames significantly. However, this precaution was not necessary for ISEC in our simulations.

4. SUMMARY AND CONCLUSIONS

A new technique for concealment of arbitrarily shaped loss areas in erroneously-received images was proposed in this work. In contrast to intra-concealment methods, the algorithm was designed to conceal a distorted high-resolution sequence utilizing one or more error-free low-resolution reference sequences which are perfectly synchronized in time. A typical application for this inter-sequence concealment technique could be a multi-broadcast-reception scenario of both DVB-T and DVB-H or both DVB-T and T-DMB. First simulation results based on lost macroblocks show that the proposed method outperforms state-of-the-art intra-sequence error concealment methods for typical sequences. Optimizing the unknown image transformation parameters due to Levenberg-Marquardt guarantees robustness and maximizes the objective restoration quality even for inaccurate start values.

REFERENCES

[1] ETSI, *Digital Video Broadcasting (DVB); Framing structure, channel coding and modulation for digital terrestrial television*. EN 300 744, V.1.5.1, Nov. 2004.

[2] ETSI, *Digital Video Broadcasting (DVB); Transmission system for handheld terminals (DVB-H)*. EN 302 304, V.1.1.1, Nov. 2004.

[3] ETSI, *Digital Audio Broadcasting (DAB); DMB video service; user application specification*. TS 102 428, V.1.1.1, June 2005.

[4] W.-M. Lam, A.-R. Reibman, and B. Liu, "Recovery of lost or erroneously received motion vectors, " in *Proceedings of Intern. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, Minneapolis, USA, April 27-30. 1993, pp. V417–V420.

[5] J. Zhang, J. F. Arnold, and M. R. Frater, "A cell-loss concealment technique for MPEG-2 coded video", *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 10, pp. 659–665, June 2000.

[6] Y.-K. Wang, M. M. Hannuksela, and V. Varsa, "The error concealment feature in the H.26L test model, " in *Proc. on Intern. Conf. on Image Processing (ICIP)*, Rochester, USA, Sept. 22-25. 2002, pp. 729–732.

[7] J.-R. Ohm, *Multimedia Communication Technology*, Berlin/Heidelberg/New York: Springer, 2004.

[8] F. Dufaux and J. Konrad, "Effienct, Robust, and Fast Global Motion Estimation for Video Coding," *IEEE Trans. on Image Processing*, vol. 9, pp. 497–501, Mar. 2000.

[9] P. Salama, N. B. Shroff, and E. J. Delp, "Error Concealment in Encoded Video Streams", Chapter 7, in A. K. Katsaggelos and N.P. Galatsanos (editors), *Signal Recovery Techniques for Image and Video Compression and Transmission*, Boston, USA: Kluwer Academic Publishers, 1998.

[10] W. H. Press, B.O. Flannery, S.A. Teukolsky, and W.T. Vetterling, *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge, U.K.: Cambridge University Press, 1988.

[11] R. Schaback, "Convergence Analysis of the General Gauss-Newton Algorithm" in *Numerische Mathematik*, vol. 46, pp. 281–309, June 1985.

Video	Crew	Discovery City	Rugby
H.264 Intra	30.51	28.86	22.16
BMA	25.43	27.26	20.67
DMVE	29.63	30.42	23.72
ISEC (QVGA)	39.33	40.85	28.27
ISEC (CIF)	41.23	41.33	29.21

Table 1: Mean PSNR values for luminance Y in dB (ISEC: the maximum $PSNR_Y$ value of projection cases 1 and 2 is taken at a bit rate of 0.3 bpp used for the LRRS)

a)

b)

c)

Figure 6: Visual results for ISEC of sequence *crew*. (a) Image with lost macroblocks, (b) Concealed Image ($PSNR_Y$: 28.17 dB, LRRS: resolution *CIF*, case 1, 0.005 bpp), (c) Concealed Image ($PSNR_Y$: 39.63 dB, LRRS: resolution *CIF*, case 1, 0.258 bpp)