On the Effects of Sender-Receiver Concealment Mismatch on Multimedia Communication Optimization

Original
On the Effects of Sender-Receiver Concealment Mismatch on Multimedia Communication Optimization / MASALA, Enrico; De Vito, F; DE MARTIN, JUAN CARLOS. - In: MULTIMEDIA TOOLS AND APPLICATIONS. - ISSN 1380-7501. - STAMPA. - 74:9(2015), pp. 2879-2898. [10.1007/s11042-013-1751-y]

Publisher:
Springer

Published
DOI:10.1007/s11042-013-1751-y

Multimedia Tools and Applications manuscript No.
(will be inserted by the editor)

On the Effects of Sender-Receiver Concealment Mismatch on Multimedia Communication Optimization

Enrico Masala · Fabio De Vito ·
Juan Carlos De Martin

Received: date / Accepted: date

Abstract A large number of performance optimization algorithms for multimedia communications, including rate-distortion optimized schemes, rely on knowing the decoder behavior in case of data loss, i.e., the decoder-side error concealment technique. However, for the specific case of video coding, standards do not specify it, thus different decoders may — and typically do — use different concealment techniques. This work investigates the impact of assuming, in the transmission optimization phase, a concealment algorithm different from the one that is actually used by the decoder, in order to determine which are the best assumptions to use at the transmitter. Firstly, we investigate the typical performance provided by ten concealment techniques belonging to three widely used algorithmic families (spatial, temporal and mixed). Then, we assess the impact that an incorrect concealment assumption causes, in terms of both packet transmission policy changes and video quality degradation, using a simple rate-distortion transmission optimization technique that targets a generic two QoS-level network. Simulation results over several standard video sequences show that the performance impact of incorrectly assuming the decoder-side concealment technique may be significant but it is limited if the two techniques belong to the same algorithmic family. Moreover, the impact on performance caused by incorrect assumptions is strongly mitigated if the

---

Enrico Masala, Fabio De Vito, Juan Carlos De Martin
Control and Computer Engineering Department
Politecnico di Torino
Corso Duca degli Abruzzi, 24 - 10129 Torino, Italy Tel.: +39-011-0907036
Fax: +39-011-0907099
E-mail: masala@polito.it, fabio.devito@polito.it, demartin@polito.it

decoder employs a high-performance concealment algorithm. Finally, the impact on the performance of several parameters such as the encoding pattern, the packet loss statistics (uniform and burst losses) and the amount of high-priority traffic is evaluated, showing that the conclusions can be confidently applied to actual multimedia communication scenarios.

Keywords H.264 · Concealment mismatch · Concealment assumption · Distortion estimation

1 Introduction

Multimedia communications account for a consistent — and growing — share of the Internet traffic. Unlike traditional Internet applications such as file transfer, web browsing and emails, multimedia applications can tolerate a limited amount of data loss, i.e., the imperfectly received content can still be perceptually acceptable or even, in the best case, perceptually indistinguishable from the original. In case of data loss, concealment techniques are generally applied at the decoder to mask the effect of the missing data. Numerous efficient techniques have been proposed for different types of media, including video [8].

However, the intrinsic unreliability of many network technologies, including IP networks, has been motivating many researchers to investigate methods to improve the communication robustness for decades. Several approaches have been proposed to deal with the lack of network level Quality of Service (QoS) guarantees; two possible examples among many: data retransmissions and forward error correction. Since network and processing resources are usually limited, such techniques must be optimized in order not to waste resources.

In the specific case of multimedia, a popular and very effective approach has been based on assigning to each compressed data segment, e.g., a packet, a protection level which is proportional to its perceptual importance. This is the principle underlying the Unequal Error Protection (UEP) techniques [17]. The perceptual importance is usually estimated by measuring the distortion caused by the loss of a given media segment. That value can then be used to estimate the resulting expected quality perceived by the user in case of loss [18,12,3].

Unfortunately, estimating the impact of losses implies simulating the behavior of the decoder in case of missing data, which is often difficult for a number of reasons. For instance, simulating the behavior may be computationally heavy depending on the number of transmission outcome possibilities that must be considered. Moreover, for certain multimedia compression standards (as it is generally the case for video) the decoder error concealment algorithm in case of losses or errors is not normative: each designer of decoding software may autonomously decide how to deal with damaged compressed streams. Finally, the scenario may present complexities. For instance, multicasting data to many different decoders running at the same time do not allow to easily simulate the decoder behavior since, even if known, concealment algorithms may be too numerous. In this scenario the best choice would be to use an assumption, in the transmission optimization algorithm, that in general provides the best average performance.

The knowledge of the concealment technique used by the decoder is also fundamental for all multimedia communications systems that optimize the performance within a rate distortion framework adapted to the case of video communication [5,13]. In such framework it is necessary to compute the expected distortion at the decoder given a certain transmission policy, which indeed requires the knowledge of the concealment technique used by the decoder. Alternatively, one could consider all lost data as well as the data dependent on them as to be concealed with a very basic technique, thus effectively computing an upper bound on the expected distortion at the decoder during the optimization phase, as done in older works [6]; this approach, although interesting, yields suboptimal results. However, nearly all the works in the scientific literature assume that the technique is known, and they act consequently while developing the optimized transmission system [24,4,18,1,12,13, 5].

Very few works, instead, have addressed the issue of investigating how the performance of the communication system is affected when the previous assumption does not hold. In this case it may happen that, during the transmission optimization phase, a concealment technique that differs from the one actually used at the decoder is assumed. In the context of studying the impact of mismatch of many parameters used by the transmitter and the decoder, the work in [7] briefly investigated the effect of using one concealment technique slightly different from the one assumed at the transmitter, showing that in the conditions of their experimental setup the performance difference can be significant, up to 1 dB Peak Signal-to-Noise Ratio (PSNR). The work in [19] briefly discussed the impact of using a decoder concealment technique different from the one assumed during the optimization phase concluding that if the actual technique implemented at the decoder performs better than the one assumed by the transmission optimization algorithm the final performance is not negatively affected. However, both studies are limited to considering only one concealment technique.

More recent works analyzed the concealment issues also for recent standards such as H.264/AVC, as done in [23] for the case of packet losses or for the case of bit errors as in [21, 10]. However, their analysis is limited since they do not consider the transmission optimization part, therefore the impact that a wrong assumption about the concealment technique may have on the optimization process remains unclear.

In this paper we attempt an extensive investigation of the impact of assuming, during the transmission optimization phase, a concealment algorithm different from the one that is actually used by the decoder. Ten concealment algorithms, belonging to three widespread algorithmic families (spatial, temporal and mixed), are considered. The first part of the work focuses on the absolute performance of the concealment algorithms under investigation for a set of standard video sequences. Then, the impact that incorrect concealment assumptions cause on the transmission performance is extensively assessed by investigating both the packet transmission policy variations and the perceptual performance degradation. To do so, we use a simple transmission optimization technique that targets a generic two QoS-level network. Some preliminary results presented in [9] on a non rate-distortion optimized transmission showed the impact of incorrect concealment assumptions for low-resolution video sequences. This work focuses on much more realistic conditions by employing a rate-distortion optimized transmission system. Moreover, we thoroughly investigate the dependency of the performance on several important parameters for practical implementations, such as the packet loss rate, the amount of high priority traffic, the video encoding pattern and the packet loss statistics in terms of average burst length of packet losses. Also, higher image resolutions compared to [9] are used to investigate a more realistic setting. The results of this work could be used, for instance, to improve the planning and optimization of video communication systems before actual transmissions take place providing an indication of the most suitable assumptions to use.

The paper is organized as follows. In Section 2 we describe the concealment algorithms used in this work, their properties and their relative performance. Section 3 highlights the importance of the assumption about the concealment algorithm employed at the receiver to correctly estimate the expected distortion, and presents a rate-distortion transmission optimization framework for a two-class DiffServ network. Extensive simulation results are presented in Section 4, addressing both the pure concealment performance of each algorithm and the effect, in terms of end-to-end performance, of incorrect assumptions about the concealment algorithm during the optimization phase. The impact of several encoding and communication parameters on the performance is also assessed. Conclusions are drawn in Section 5.

2 Video Concealment Techniques

2.1 Overview

For video communications a large number of error concealment algorithms have been proposed, and the number is constantly growing (see, e.g., [8,16,25,20,11,26,27] for a review.)

However, a considerable amount of the video concealment algorithms proposed in recent years can be grouped in three families, according to the concealment principle used to generate the data to replace the missing or corrupted macroblocks. *Spatial* concealment algorithms interpolate the missing information using data belonging to the same video frame, for instance using surrounding macroblocks that have already been decoded. They achieve good performance when there are highly correlated areas within the same frame. *Temporal* concealment algorithms rely on the data present in previously decoded frames to build an estimate of the missing areas in the current frame. Missing areas can be derived from the pixels in the same position in reference frames, or motion compensation can be applied in order to preserve the motion continuity among different frames. *Mixed* approaches combine spatial and temporal techniques. One of the most popular approaches uses spatial concealment on the I-frame and temporal concealment on P- and B-frames to exploit the generally higher performance of temporal concealment techniques while avoiding concealment dependencies between consecutive group of pictures (GOP).

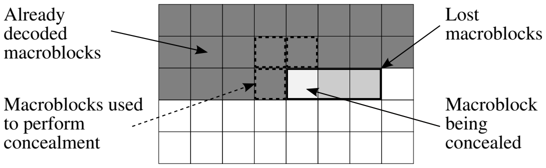


Fig. 1 Macroblocks used to estimate the value of a missing macroblock for the sp_4 concealment technique.

2.2 Considered Techniques

For the purpose of this work, we implemented several concealment techniques that are based on well-known concealment principles (again, see, e.g., [8]) and that we consider — as a whole — representative of a large number of implementations in actual video communication systems.

- Spatial (intra-frame) algorithms:
  - sp_1: each lost macroblock (MB) is replaced by the MB immediately above it; if the loss occurs on MBs of the first row, the concealment data is a 16×16 zero-matrix;
  - sp_2: similarly to sp_1, a missing MB is replaced by the MB immediately on its left; in case the lost MB is on the first column, then a 16×16 zero-matrix is used;
  - sp_3: for each missing MB, each 4×4 block within that MB is filled with the average of the colors in the three left, upper-left and upper 4×4 blocks;
  - sp_4: for each missing MB, the average color of the three upper-left macroblocks is used to fill it, as shown in Figure 1.
- Temporal (inter-frame) algorithms:

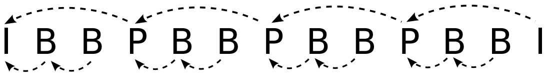


Fig. 2 Concealment algorithm te_1: reference frames used in case of loss for a typical GOP pattern.

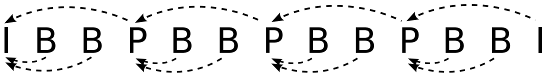


Fig. 3 Concealment algorithms te_2 and te_3: reference frames used in case of loss for a typical GOP pattern.

te_1: missing MBs are replaced by the ones in the same position in the reference frame (*frame copy*);

te_2: for each missing MB a motion vector (MV) is estimated by averaging the MVs of the three left, upper-left and upper MBs, and then used to select a 16×16 area in the reference frame that replaces the missing MB;

te_3: for each missing MB the motion vector of the same MB in the reference frame is used. In case the reference is an I-frame, the MV is set to zero. Then the MV is used to select a 16×16 area in the reference frame that replaces the missing MB.

- Mixed algorithms:

mix_1: use sp_3 on I-frames and te_1 on remaining frames;

mix_2: use sp_3 on I-frames and te_2 on remaining frames;

mix_3: use sp_3 on I-frames and te_3 on remaining frames.

Note that when the concealment technique uses reference frames, they are defined as follows. For the case of te_1, I- and P-frames use the previous I- or P-frame in display order as reference; B-frames always refer to the immediately preceding frame in display order, as shown in Figure 2. For the case of te_2 and te_3, all frames use the most recent I- or P-frame in display order, as shown in Figure 3. Moreover, when a packet loss occurs, several MBs may need concealment. In this case note that many of the above techniques that compute data by interpolating surrounding macroblock information may use, as a starting point, values just computed by the concealment algorithm in a previous step.

2.3 Performance of the Concealment Algorithms

In order to preliminary investigate the performance of the ten concealment techniques previously described, for each one of them the loss of each packet has been simulated, the resulting MSE on the affected frame computed, and then values have been averaged over all the packets which constitute the video sequence. To isolate the effect of the concealment algorithm from the distortion due to compression, first we decode the video sequence without errors, creating an error-free version of the decoded sequence, then we compute the MSE with respect to this reference in all transmission experiments.

Each CIF video sequence has been encoded using the H.264/AVC encoder [14,15] imposing a maximum packet size equal to 1000 bytes. The GOP frame pattern is IBBPBBPBBPBB, thus each GOP is composed of 12 frames. The quantization parameter has been fixed to 28 for I- and P-frames, and to 30 for B-frames, achieving a nearly constant quality at the expense of a variable bitrate over time.

Table 1 Average MSE caused by the loss of a packet assuming the concealment techniques shown in the table.

Concealment algorithm	Average MSE				
	foreman	*tempete*	*mobile*	*paris*	*news*
sp_1	1294.8	213.2	217.0	549.3	749.8
sp_2	2138.2	310.2	474.4	854.2	1083.8
sp_3	1254.0	186.6	159.9	504.2	711.2
sp_4	1251.9	183.6	159.8	497.5	696.5
te_1	120.7	36.4	79.7	34.0	46.6
te_2	91.0	24.5	61.3	27.4	35.3
te_3	95.4	27.2	64.6	28.0	37.1
mix_1	207.1	62.2	75.7	90.0	142.5
mix_2	189.1	54.8	63.8	86.1	135.1
mix_3	191.3	56.7	67.1	86.1	136.2

Table 1 shows the average per-packet MSE of the sequences *foreman*, *tempete*, *mobile*, *paris* and *news*. To evaluate the distortion contribution of each single packet, the MSE is computed considering losses in isolation, i.e., for the results in this table, the video sequence is assumed to be correctly received up to the point of the lost packet. Values show that the MSE of the spatial concealment algorithms are about one order of magnitude higher than the one achieved by the other two families, while temporal and mixed approaches show comparable values. Values provided by algorithms belonging to the same family are close to each other. Given a concealment algorithm, distortion values vary across the sequences, as expected, since each sequence has its peculiar characteristics in terms of content and motion. Note that the results shown here cannot be directly compared with [9] since the frame size is different as well as the packet size, which in this work has been limited to make simulation conditions more realistic.

3 Rate-Distortion Optimization

3.1 Sample Scenario

Let us consider, as an example, a communication scenario where video is multicast to many receivers using video decoders produced by different manufacturers. Each decoder usually implements its own error concealment scheme to deal with packet losses. Although formulations of the rate-distortion optimization problem are quite common in the literature, in the above described multicast scenario it is impossible to select one specific receiver-side concealment algorithm to optimize the transmission simply because there are many of them. Moreover, the concealment algorithms might be not even known at system design time.

The aim of this work is to study the performance that can be achieved when various assumptions are made at the transmitter about the concealment technique used at the receiver. A practical case is, for instance, that of a live video transmitter, which needs to run the rate-distortion optimization framework in real time to adapt to the characteristics of the input signal and must rely on some assumptions about the decoder concealment technique.

3.2 Optimization Framework

If the behavior of the decoder in case of errors is known, it is possible to use a well established rate distortion optimization framework such as [6] in order to dynamically adapt the transmission policy depending on, e.g., the characteristics of the video signal. The rate distortion optimization framework formulates the optimization problem as

$$\begin{aligned} &\min_{\{\Pi\}} \; D(\Pi) \\ &\text{s.t. : system constraints} \end{aligned}$$

where Π is the transmission policy, D is the expected distortion at the receiver as a function of the transmission policy Π, and the system constraints are any constraints imposed by the application and system architecture, e.g., maximum channel rate constraints.

Finding the exact solution of the previous minimization problem by enumerating all possible values of $D(\Pi)$ is generally not possible due to the high number of transmission policies involved as well as to the difficulty in computing the distortion for each one of the policies. Moreover, as already pointed out, computing the distortion implies knowing the concealment technique used at the decoder. Since the technique is often not known in practice, this work aims at investigating and quantifying the impact of an incorrect assumption about the concealment technique used at the decoder, to understand if it is possible to minimize the impact on the performance of the communication.

In order to provide quantitative results, we focus on a simple transmission optimization problem, that is, a video communication over a two-level QoS network, such as, for instance, a Differentiated Services (DiffServ) network [2] that provides a loss free and delay bounded service, that we name *premium* in the following, as well as a free *best effort* service.

A DiffServ network handles packets by forwarding them using different priorities. Packet differentiation is achieved by marking packets, e.g., using the type of service (TOS) field of the IP protocol, so that they can be recognized by routers and treated differently. Thus, the system effectively provides different priorities to each packet (e.g., faster forwarding, lower packet loss probability) depending on the value used for marking.

For the case of two-level QoS network, we can restate the optimization problem as

$$\begin{aligned} &\min_{\{\mathit{\Pi}\}} \; D(\mathit{\Pi}) \\ &\text{s.t. :} \; R(\mathit{\Pi}) < R_{max} \end{aligned}$$

implying that there is a maximum rate (or, equivalently, cost) constraint on the data to be transmitted using the *premium* service. Each possible policy $\mathit{\Pi}$ is an assignment, for each packet, to either the *premium* or *best effort* service.

The previous problem can be recasted as an unconstrained minimization of the Lagrangian cost

$$\min_{\{\mathit{\Pi}\}} J = D(\mathit{\Pi}) + \lambda R(\mathit{\Pi}).$$

Assuming that the expected distortion associated with a certain transmission policy $\mathit{\Pi}$ can be computed as the sum of the individual contributions of each packet that can potentially be lost, as done in many works [4,18], J can be written as $\sum_i d(\pi_i) + \lambda r(\pi_i)$, thus making explicit the contribution of each packet i in terms of the expected distortion d and rate r due to the use of a given policy π_i for that packet (i.e., using the *premium* service or not).

The minimization can be easily carried out by means of the bisection algorithm [22]. The algorithm finds, by means of a successive approximation procedure, the λ value that optimally solves the minimization problem. At each iteration, a λ value is determined, and that λ value is used to perform minimization for each packet independently, thus determining if the packet is to be transmitted using the *premium* service or not. This procedure is possible since the contribution of each packet to the J value is independent of the outcome of the minimization for the others. More details about the algorithm can be found in [22].

However, the entire process relies on the possibility of computing the distortion values for each packet, which clearly depends on the concealment used at the receiver. Therefore, an incorrect concealment assumption might significantly change the policy that solves the optimization problem, i.e., the packet assignment to one of the two classes.

To investigate this issue, Table 2 shows the percentage of packets that, given an incorrect assumption on the concealment, would be sent using a different service. To avoid reporting too many data, we do not report comparisons among every couple of concealment techniques but we only compare concealment families among them. They are indicated as *sp*, *te* and *mix* for the spatial, temporal and mixed concealment families, respectively. All the combinations have been tested and the mean value is shown in the table. Note

Table 2 Percentage of packets that would change service if an incorrect assumption about the concealment technique at the decoder is made. The R_{max} constraint is set to assign 20% of the packets to the premium service.

Sequence name	Encoder family	Decoder family *sp*	*te*	*mix*
foreman	*sp*	10.5	24.7	26.0
	te	24.7	7.2	17.3
	mix	26.0	17.3	7.0
tempete	*sp*	10.9	26.1	29.8
	te	26.1	8.4	22.1
	mix	29.7	22.1	7.7
mobile	*sp*	14.0	19.7	19.6
	te	19.7	7.4	14.0
	mix	19.6	14.0	6.8
paris	*sp*	7.3	29.4	35.6
	te	29.4	7.0	34.9
	mix	35.6	34.9	3.4
news	*sp*	4.7	16.9	32.8
	te	16.9	6.5	30.5
	mix	32.8	30.5	2.4

that the table is symmetrical because, given two different assumptions at the transmitter, the number of packets assigned to a different class (from premium to best effort and vice versa) is the same while the direction of the change is the opposite.

The number of packets that would be assigned to a different service if an incorrect concealment technique is assumed ranges from 2.4% (*news* sequence) up to 14.0% (*mobile* sequence) in case the concealment technique belongs to the same family of the one actually used at the receiver (values on the diagonal of each sequence). The average value is about 7.4%. However, in the case of different families (values not on the diagonal of each sequence), the average increases to 25.3%, and values range from 14.0% up to 35.6%, showing that different assumptions on the concealment algorithms used at the decoder significantly affect the policy used to transmit the data. Similar results have been obtained varying the R_{max} constraint.

4 Simulation results

The percentage of packets that would be assigned to a different QoS class due to different assumptions about the concealment algorithm at the decoder suggests that the quality of the communication may be strongly affected. However, those values do not allow precisely quantifying the impact on the quality of the videocommunication, therefore transmission simulations are needed.

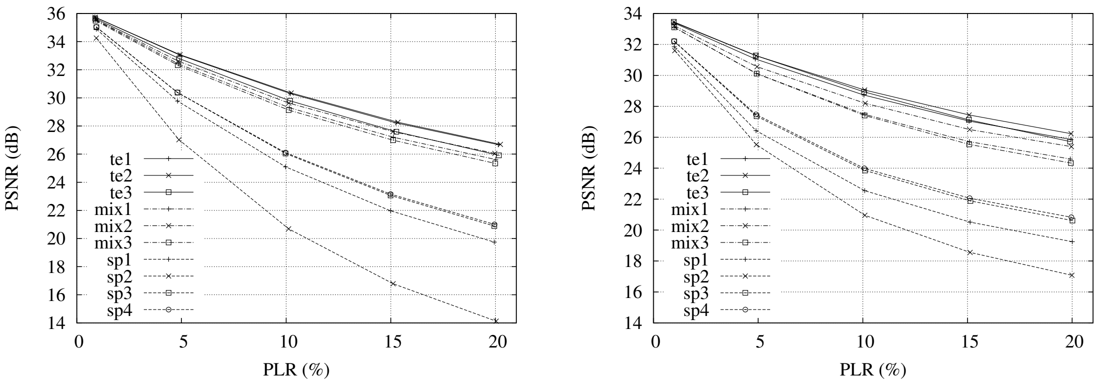
Coding parameters have already been described in Section 2.3. For each sequence, the distortion that would be caused by the loss of each packet has been computed using all the considered concealment algorithms. A video transmission has been simulated over a two-level QoS network in which the *premium* class receives a loss-free service whereas the *best effort* class is subject to uniform random packet losses.

The results have been obtained by using the premium service for an amount of packets corresponding to about 20% of the total channel bandwidth, but similar results have been obtained using different values as detailed in Section 4.5. Several transmission simulations have been performed, assuming all the various concealment techniques at the sender, thus computing, each time, a different transmission policy. For each one of these cases, the transmission has been decoded several times, each time using one of the considered concealment techniques. Therefore, 100 combinations of assumed and actual concealment techniques have been considered for each sequence. Each transmission and decoding experiment has been repeated 40 times in order to achieve statistically significant results.

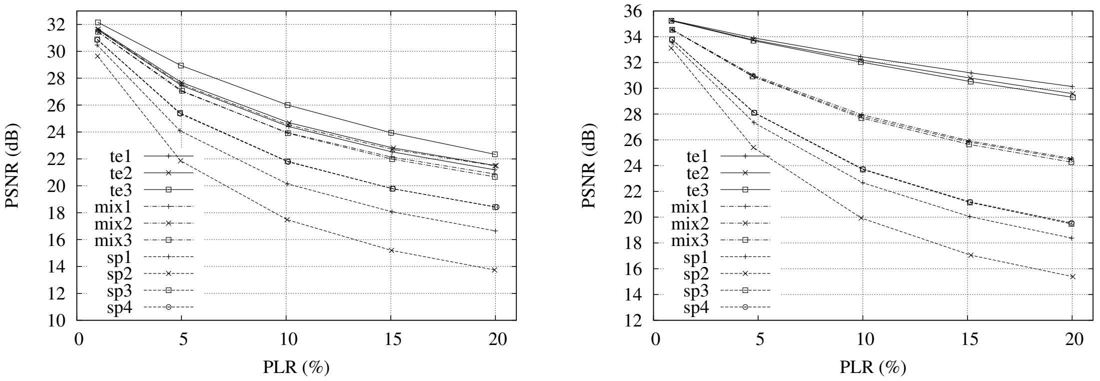
4.1 Correct Concealment Assumption

This section investigates the case when the transmitter uses, for optimization purposes, the same concealment algorithm that has been used at the decoder. This is the ideal condition in which every system would like to operate. Thus, the performance depends only on the ability of the considered concealment algorithm to recover packet losses and it can be considered as an upper bound for the concealment mismatch case. As usual in the multimedia research community we evaluate the results for a video sequence using the PSNR averaged over all frames. The PSNR (in dB) of a frame is computed by means of the formula $10\log_{10}\frac{255^2}{MSE}$ where MSE is the mean squared error between the frame under test and the reference one and 255 is the maximum value of the luminance of a pixel.

The PSNR values have been plotted in Figure 4 for the four considered sequences. As it could be expected from Table 1, spatial concealment algorithms provide the lowest performance, sp_2 being the worst one, whereas sp_3 and sp_4 are the best ones among them. Algorithms based on the mixed approach yield PSNR gains with respect to the spatial approach, up to 5 dB PSNR at 20% Packet Loss Rate (PLR). However, the best performance is usually obtained using algorithms based on temporal concealment techniques, especially with



(a) *foreman* (b) *tempete*



(c) *mobile* (d) *paris*

Fig. 4 PSNR performance as a function of the PLR when the concealment algorithm is the same assumed in the transmission optimization phase (ideal condition.)

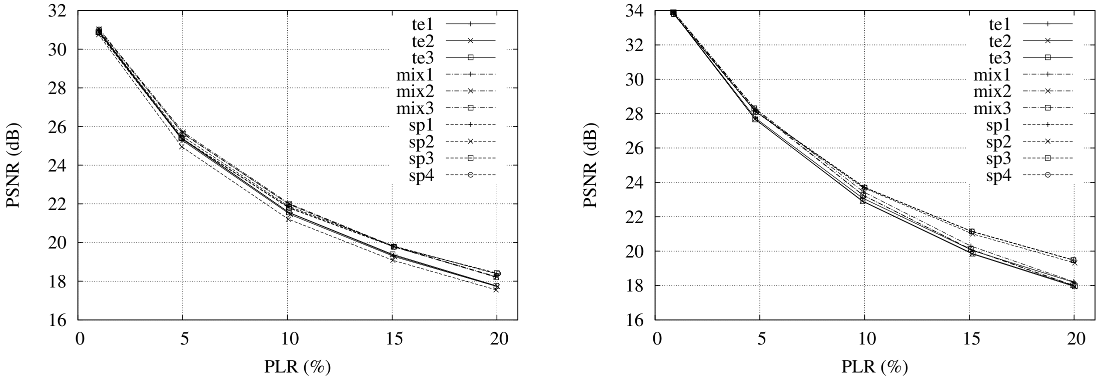
sequences such as *mobile* and *paris*. In the latter case, gains are up to 5 dB. This is probably due to the particular video content in which many fine details are present in the static background of the picture. In such conditions, spatial concealment techniques provide poor performance since they are generally able to effectively recover only the low frequency components in the video content.

4.2 Incorrect Concealment Assumption

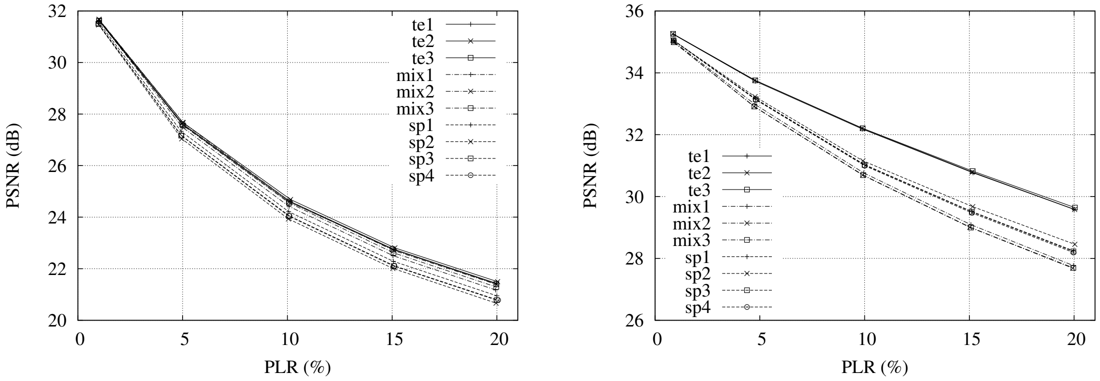
This section investigates the case in which the transmission policy is optimized while assuming a concealment algorithm different from the one that is actually used at the decoder. In these conditions, the final PSNR performance is the combination of two effects. First, packets are assigned to QoS classes in a suboptimal way, thus the expected performance is not maximized. Second, the ability of the concealment algorithm itself to recover lost data influences the performance.

Figure 5 shows the PSNR performance as a function of the PLR for the *mobile* and *paris* sequences when various concealment assumptions are made at the transmitter. For the *mobile* sequence, in case a spatial or mixed concealment technique such as sp_3 or mix_2 are used at the decoder, making the correct assumption at the encoder is important in order to achieve a good performance. When a temporal concealment technique (e.g., te_2) is employed at the decoder, the assumptions at the encoder have a smaller impact on the final PSNR performance. Similar results are achieved with the *paris* sequence. For this sequence, when a temporal concealment technique is used at the decoder it is important to use a similar assumption at the encoder, otherwise the performance degrades significantly especially at high PLR.

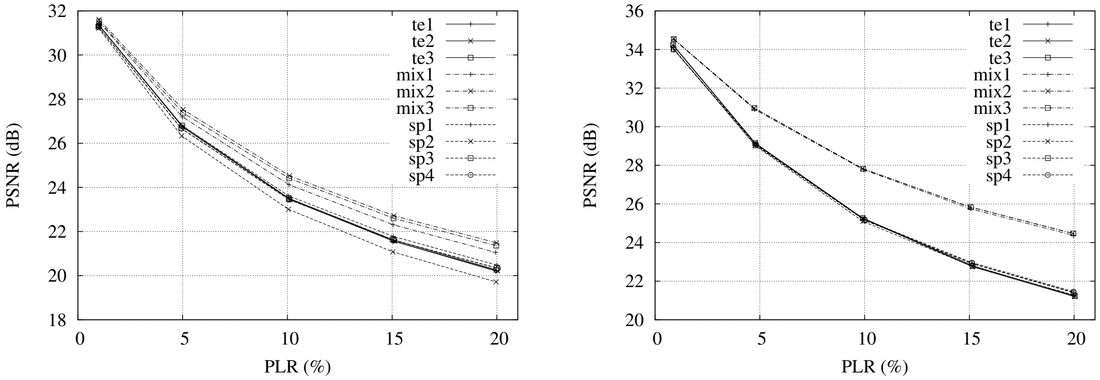
Table 3 summarizes, for the case of 10% PLR, the PSNR performance when each one of the considered concealment techniques is assumed at the transmitter. To reduce the amount of data, results are shown only for the best



(a) *mobile*, dec. algorithm: sp_3 (b) *paris*, dec. algorithm: sp_3



(c) *mobile*, dec. algorithm: te_2 (d) *paris*, dec. algorithm: te_2



(e) *mobile*, dec. algorithm: mix_2 (f) *paris*, dec. algorithm: mix_2

Fig. 5 PSNR performance as a function of the PLR for the *mobile* (left) and *paris* (right) sequences when various concealment assumptions are made at the transmitter and three different concealment algorithms are used at the decoder.

Table 3 PSNR performance at 10% PLR.

Sequence	Encoder assumption	PSNR (dB)		
		Concealment at decoder		
		sp_3	te_2	mix_2
foreman	sp_1	26.05	29.12	27.77
	sp_2	26.11	29.75	27.99
	sp_3	26.05	29.12	27.77
	sp_4	26.05	29.12	27.77
	te_1	25.42	30.30	27.88
	te_2	25.19	30.35	27.79
	te_3	25.34	30.34	27.90
	mix_1	26.23	29.99	29.30
	mix_2	26.03	30.08	29.63
	mix_3	25.98	29.99	29.55
tempete	sp_1	23.93	27.90	25.86
	sp_2	24.35	28.51	26.41
	sp_3	23.85	27.85	25.80
	sp_4	23.88	27.87	25.81
	te_1	23.75	28.96	26.13
	te_2	23.55	29.05	26.06
	te_3	23.43	28.96	25.96
	mix_1	24.34	28.91	27.73
	mix_2	24.00	28.85	28.20
	mix_3	23.98	28.84	28.11
mobile	sp_1	21.88	24.21	23.61
	sp_2	21.20	23.93	23.01
	sp_3	21.82	24.04	23.46
	sp_4	21.77	24.05	23.44
	te_1	21.52	24.61	23.45
	te_2	21.44	24.68	23.45
	te_3	21.56	24.60	23.51
	mix_1	21.97	24.42	24.12
	mix_2	22.01	24.62	24.53
	mix_3	21.98	24.55	24.42
paris	sp_1	23.71	31.03	25.16
	sp_2	23.70	31.16	25.10
	sp_3	23.70	31.02	25.15
	sp_4	23.70	31.00	25.19
	te_1	23.09	32.18	25.27
	te_2	22.91	32.20	25.25
	te_3	22.91	32.21	25.27
	mix_1	23.44	30.77	27.77
	mix_2	23.27	30.69	27.83
	mix_3	23.25	30.69	27.81
news	sp_1	27.35	34.53	28.73
	sp_2	27.19	34.50	28.52
	sp_3	27.38	34.54	28.81
	sp_4	27.34	34.53	28.79
	te_1	27.11	35.00	28.86
	te_2	26.91	35.00	28.85
	te_3	27.01	34.95	28.93
	mix_1	27.27	34.05	30.82
	mix_2	27.14	33.97	30.86
	mix_3	27.10	33.94	30.83

performing concealment techniques at the decoder for each considered family (spatial, temporal or mixed).

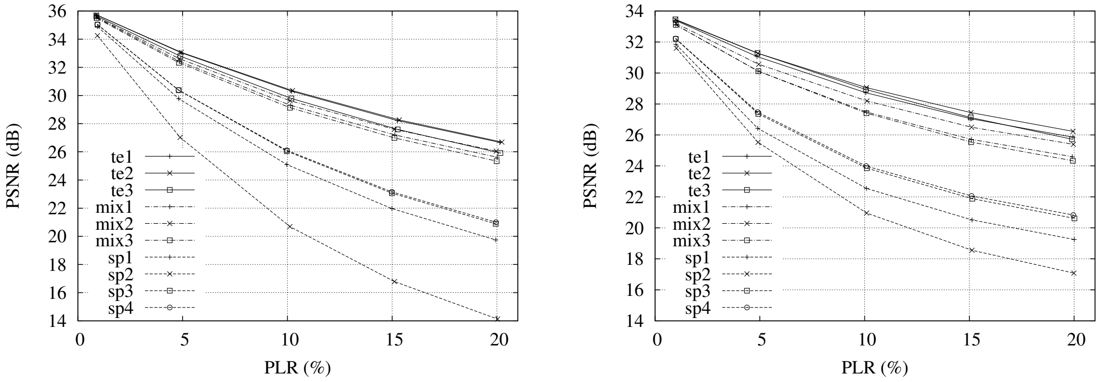
Results show that the best performance is nearly always achieved using a temporal algorithm at the decoder. Moreover, assuming a temporal or, as a second choice, a mixed concealment technique at transmission optimization time yields the best performance.

The table also includes the results for the *news* sequence, which has not been shown in the previous graphs due to lack of space. The sequence is characterized by some very static content (background) and limited motion (the two speakers and a small virtual screen in the middle). However, the performance trend for the *news* sequence is similar to that of the *tempete* sequence, since the latter is characterized by a static scene but with a slow camera zoom-out.

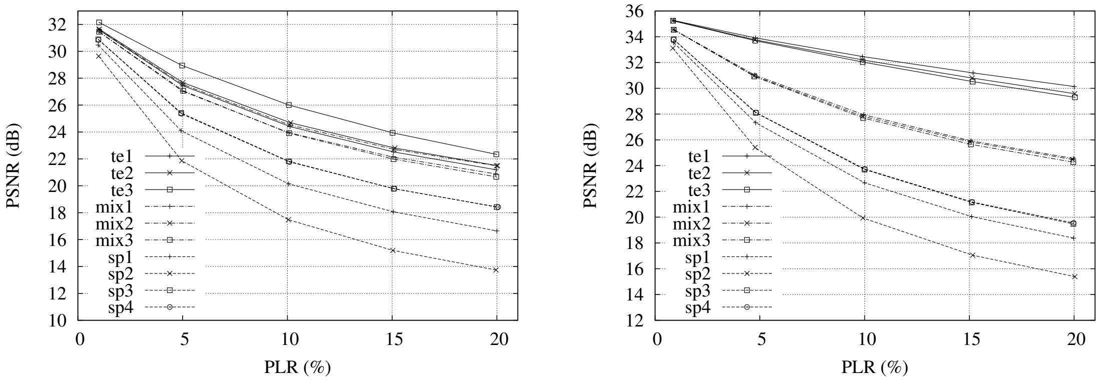
4.3 The Usual Assumption for Optimization

This section investigates the performance for the case in which the best temporal concealment technique is assumed in the optimization phase, as it is done by the majority of the works in literature which, however, rarely analyze the impact of this choice on the performance. This assumption can represent a very common scenario in which the video is broadcast to many receivers thus the transmission policy cannot be tailored to each one of them.

Figure 6 shows that the best performance is achieved when a temporal concealment technique is employed at the decoder, as expected, even though for many sequences the mix_2 algorithm assumption may be a valid alternative since it provides performance similar to that of temporal algorithms.



(a) *foreman* (b) *tempete*

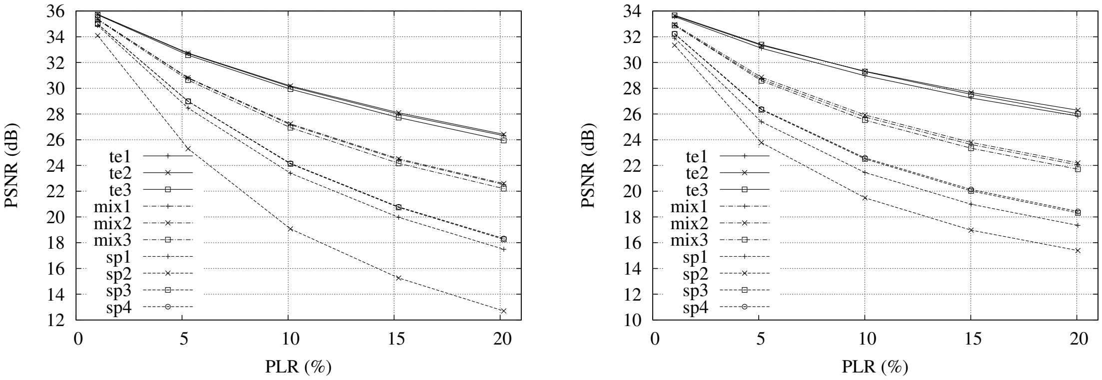


(c) *mobile* (d) *paris*

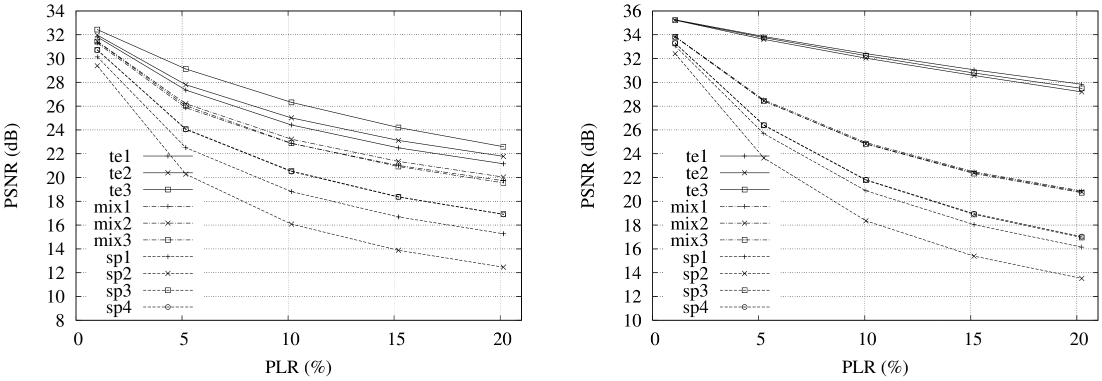
Fig. 6 PSNR performance as a function of the PLR when the te_2 algorithm is assumed in the transmission optimization phase.

Using concealment algorithms belonging to the temporal family both during the transmission phase and at the decoder significantly outperforms other combinations, ranging from 1 dB up to 5 dB PSNR in the case of the *paris* sequence at 20% PLR.

In the following we study how the performance achieved by assuming a temporal concealment algorithm in the transmission optimization phase is affected by several parameters, i.e., different encoding patterns, amount of



(a) *foreman* (b) *tempete*



(c) *mobile* (d) *paris*

Fig. 7 PSNR performance as a function of the PLR for the IBPBP... encoding scheme when the te_2 algorithm is assumed in the transmission optimization phase.

premium bandwidth and channel statistics, in terms of average burst length of packet losses.

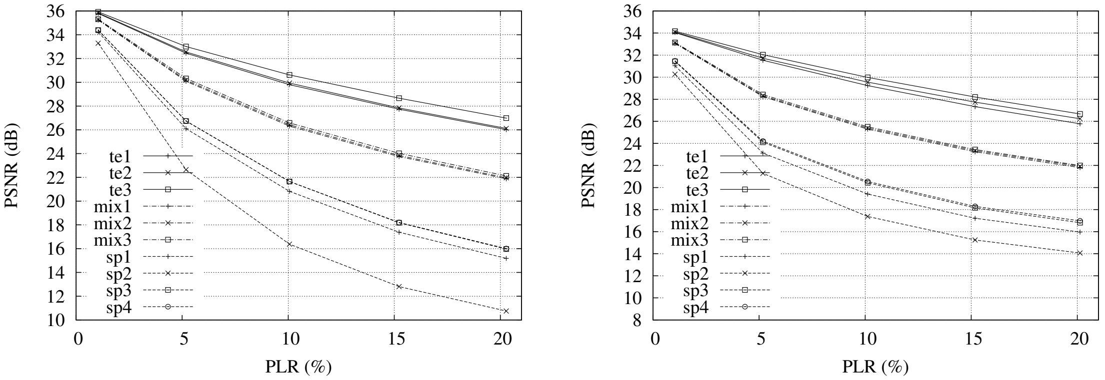
4.4 Influence of Encoding Parameters

First, the GOP structure has been changed by varying the number of B frames between two consecutive P or I frames as well as using an IPPP... coding structure with no B frames. Results are presented in Figure 7 for an

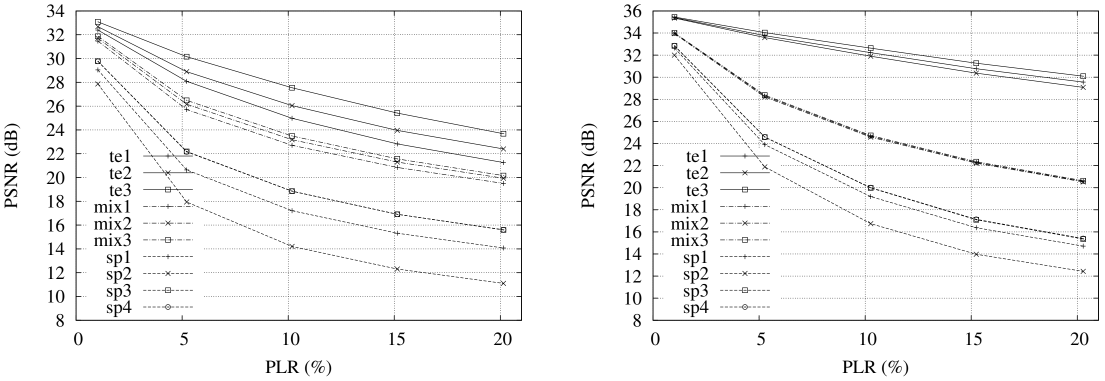
IBPBPBPBPBPB GOP coding structure. Performance is slightly worse than the previous case presented in Figure 6. This can be attributed to the fact that in the previous case the IBBPBBP... coding structure with a higher number of B frames probably helps to reduce the amount of error propagation since B frames do not propagate errors to other frames. This is further confirmed by the performance, show in Figure 8, for the case of the IPPP... coding structure, with eleven P frames between two I frames. In this case there is a clear difference between the various concealment families, since each frame is used to form the prediction of the next frame in the GOP, therefore it is more important to have reliable estimates at the encoder side than in the previous cases.

4.5 Influence of the Bandwidth Constraint

This section investigates the impact on performance of the value of the bandwidth constraint on the premium traffic. Four different R_{max} values, corresponding to the case of 10%, 20%, 30% and 40% of packets sent as premium traffic, have been tested, and results are shown in Figure 9. Clearly as the R_{max} value increases performance improves. However, it is important to note that the performance of the various concealment techniques remains consistent across the whole range of considered values. Note also that values ranging from 30% to 40% are a good representation of a DiffServ network which achieves the QoS guarantees for the premium class by means of strongly limiting the amount of premium traffic allowed in the network.



(a) *foreman* (b) *tempete*

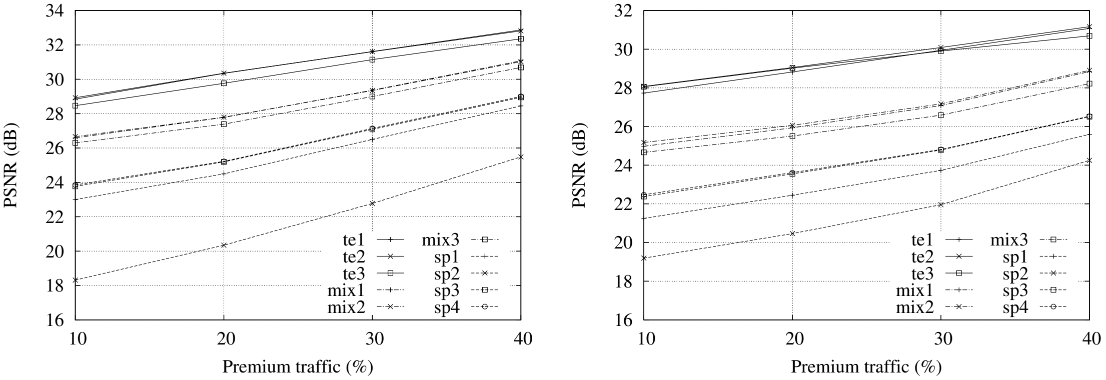


(c) *mobile* (d) *paris*

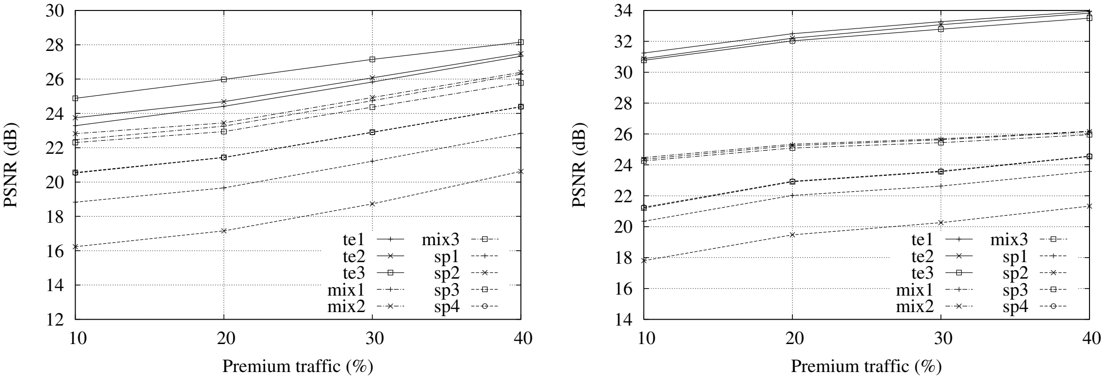
Fig. 8 PSNR performance as a function of the PLR for the IPPP... encoding scheme when the te_2 algorithm is assumed in the transmission optimization phase.

4.6 Influence of Channel Statistics

In actual scenarios losses often happen in bursts. Therefore, we modeled the best effort service as a bursty packet loss channel using the Gilbert-Elliott model, imposing different values of average burst length. The results are shown in Figure 10. Note that the packet loss rate remains constant while the average burst length increases. However, the relative performance of the various



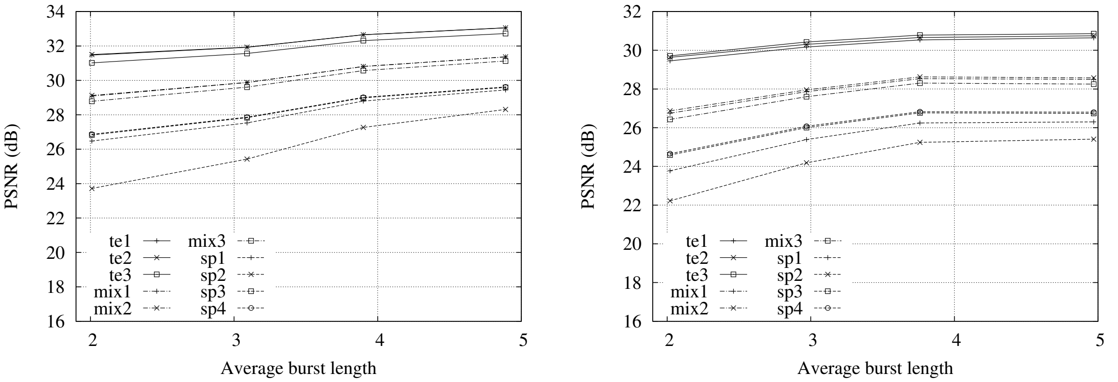
(a) *foreman* (b) *tempete*



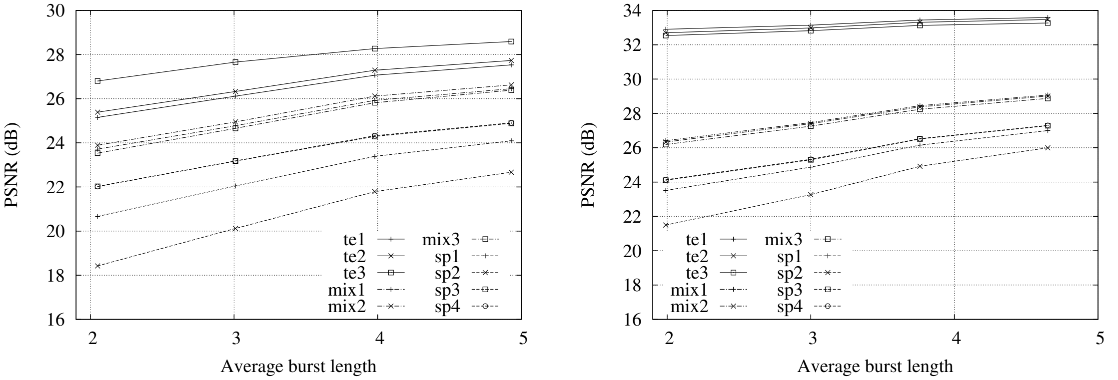
(c) *mobile* (d) *paris*

Fig. 9 PSNR performance as a function of the bandwidth constraint on the premium traffic for the *mobile*, *tempete* and *paris* sequences (from left to right) when the te_2 algorithm is assumed in the transmission optimization phase (10% PLR).

concealment techniques is consistent also when the average burst length is varied, thus showing that the conclusions can be confidently applied to actual communication scenarios.



(a) *foreman* (b) *tempete*



(c) *mobile* (d) *paris*

Fig. 10 PSNR performance as a function of the average burst length when the te_2 algorithm is assumed in the transmission optimization phase (10% PLR).

5 Conclusion

This work investigated the impact of assuming a concealment algorithm different from the one that is actually used by the decoder and the consequences on the performance of a rate-distortion optimized video communication. We considered ten concealment techniques, representing a wide variety of algorithms present in literature, equally subdivided into three widely used concealment families: spatial, temporal and mixed concealment algorithms. We showed that concealment techniques present a substantially similar performance if they belong to the same family, whereas large differences can be seen otherwise. Then, this work investigated the effect of an incorrect assumption, during the transmission optimization phase, about the concealment technique employed at the decoder. Rate-distortion optimized transmission simulations of H.264/AVC compressed video sequences over a two-class differentiated services network have been performed testing all concealment combinations. In our test scenario we found that transmission policy variations are limited if the considered algorithms belong to the same family (on average, policy changes for 7.4% of the packets) while they are much more pronounced otherwise (25.3%). Moreover, the performance strongly depends on the family to which the concealment techniques belong, rather than the algorithms themselves assumed at the transmitter and actually employed at the decoder. If the assumed concealment algorithm at the transmitter and receiver side belong to the same family the negative impact is limited. The best performance is of course achieved when the assumption matches the actual algorithm used at the decoder, however the performance decrease caused by a generic incorrect assumption in the transmission optimization phase is strongly mitigated by employing a high-performance concealment algorithm at the decoder. The common case of the temporal concealment assumption usually maximizes the performance, with gaps up to 5 dB for particular sequences, compared to the other concealment techniques. To further extend the applicability of the conclusions to actual scenarios, extensive simulations tested the communication performance with different coding and transmission parameters, e.g., packet loss rate and average burst length, amount of high-priority traffic and video encoding pattern. The consistency of the results across the various settings show that the conclusions can be confidently applied to actual communication scenarios.

References

1. Baldi, M., De Martin, J., Masala, E., Vesco, A.: Quality-oriented video transmission with pipeline forwarding. IEEE Transactions on Broadcasting **54**(3), 542–556 (2008)
2. Blake, S., Black, D., Carlson, M., Davies, E., Wang, Z., Weiss, W.: An architecture for differentiated services. RFC 2475 (1998)
3. Bucciol, P., Masala, E., Filippi, E., De Martin, J.: Cross-layer perceptual ARQ for video communications over 802.11e wireless networks. Advances in Multimedia (2007, Article ID 13969)
4. Chakareski, J., Apostolopoulos, J., Wee, S., Tan, W., Girod, B.: Rate-distortion hint tracks for adaptive video streaming. IEEE Transactions on Circuits and Systems for Video Technology **15**(10), 1257–1269 (2005)
5. Chen, Z., Pahalawatta, P., Tourapis, A., Wu, D.: Improved estimation of transmission distortion for error-resilient video coding. IEEE Transactions on Circuits and Systems for Video Technology **22**(4), 636–647 (2012)
6. Chou, P., Miao, Z.: Rate-distortion optimized streaming of packetized media. IEEE Transactions on Multimedia **8**(2), 390–404 (2006)
7. Côté, G., Shirani, S., Kossentini, F.: Optimal mode selection and synchronization for robust video communications over error-prone networks. IEEE Journal on Selected Areas in Communications **18**(6), 952–965 (2000)
8. Cui, Z., Gan, Z., Zhan, X., Zhu, X.: Error concealment techniques for video transmission over error-prone channels: A survey. Journal of Computational Information Systems **8**(21), 8807–8818 (2012)
9. De Vito, F., De Martin, J.: On the effects of encoder-decoder concealment mismatch on video distortion estimation. In: Proceeding of European Signal Processing Conference (EUSIPCO), pp. 1588–1591. Antalya, Turkey (2005)

10. Demirtas, A., Reibman, A., Jafarkhani, H.: Performance of H.264 with isolated bit error: Packet decode or discard? In: Proc. IEEE Intl. Conf. on Image Processing (ICIP), pp. 949–952. Brussels, Belgium (2011)
11. Gao, S.: Boundary matching based spatial interpolation for consecutive block loss concealment. In: Proc. IEEE International Symposium on Multimedia (ISM), pp. 128–132. Irvine, CA (2012)
12. Hand, J., Melkote, V., Rose, K.: A spectral approach to recursive end-to-end distortion estimation for sub-pixel motion-compensated video coding. In: Proc. IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP), pp. 825–828. Prague, Czech Republic (2011)
13. Hand, J., Melkote, V., Rose, K.: A unified estimation-theoretic framework for error-resilient scalable video coding. In: Proc. IEEE Intl. Conf. on Multimedia & Expo (ICME), pp. 67–72. Melbourne, Australia (2012)
14. ITU-T & ISO/IEC: Advanced video coding for generic audiovisual services. Std. H.264 & 14496-10 (2005)
15. Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG: Joint Model Number 8.2 (JM-8.2). URL: http://iphome.hhi.de/suehring/tml/download (2011)
16. Koloda, J., Østergaard, J., Jensen, S., Sánchez, V., Peinado, A.: Sequential error concealment for video/images by sparse linear prediction. IEEE Transactions on Multimedia **15**(4), 957–969 (2013)
17. Maani, E., Katsaggelos, A.: Unequal error protection for robust streaming of scalable video over packet lossy networks. IEEE Transactions on Circuits and Systems for Video Technology **20**(3), 407–416 (2010)
18. Masala, E., De Martin, J.: Analysis-by-synthesis distortion computation for rate-distortion optimized multimedia streaming. In: Proc. IEEE Int. Conf. on Multimedia & Expo, vol. 3, pp. 345–348. Baltimore, MD (2003)
19. Masala, E., Vesco, A., Baldi, M., De Martin, J.: Optimized H.264 video encoding and packetization for video transmission over pipeline forwarding networks. IEEE Transactions on Multimedia **11**(5), 972–985 (2009)
20. Midya, A., Sengupta, S.: Hybrid temporal/spatial error concealment strategy robust to scene transitions. In: Proc. IEEE Pacific RIM Conference on Communications, Computers, and Signal Processing, pp. 416–421. Victoria, Canada (2011)
21. Purandare, R., Kshirsagar, S., Koli, S., Gohokar, V.: Impact of bit error on video transmission over wireless networks and error resiliency. In: Proc. Intl. Conf. on Image Information Processing (ICIIP). Himachal Pradesh, India (2011)

22. Shoham, Y., Gersho, A.: Efficient bit allocation for an arbitrary set of quantizers. IEEE Transactions on Acoustics, Speech, and Signal Processing **36**(9), 1445–1453 (1988)
23. Superiori, L., Nemethova, O., Rupp, M.: Error concealment analysis for H.264/advanced video coding encoded video sequences. Elektrotechnik & Informationstechnik **129**(6), 387–399 (2012)
24. Yang, H., Rose, K.: Advances in recursive per-pixel end-to-end distortion estimation for robust video coding in H.264/AVC. IEEE Transactions on Circuits and Systems for Video Technology **17**(7), 845–856 (2007)
25. Yang, Q., Wang, Y., Jiang, X.: Error concealment algorithms and quality evaluation for H.264 videos. In: Proc. 8th Intl. Conf. on Wireless Communications, Networking and Mobile Computing (WiCOM). Shanghai, China (2012)
26. Zhang, Y., Xiang, X., Zhao, D., Ma, S., Gao, W.: Packet video error concealment with auto regressive model. IEEE Transactions on Circuits and Systems for Video Technology **22**(1), 12–27 (2012)
27. Zhou, J., Yan, B., Gharavi, H.: Efficient motion vector interpolation for error concealment of H.264/AVC. IEEE Transactions on Broadcasting **57**(1), 75–80 (2011)