Unequal Packet Loss Resilience for Fine-Granular-Scalability Video

Mihaela van der Schaar and Hayder Radha, *Member, IEEE*

***Abstract*—Several embedded video coding schemes have been recently developed for multimedia streaming over IP. In particular, fine-granular-scalability (FGS) video coding has been recently adopted by the MPEG-4 standard as the core video-compression method for streaming applications. From its inception, the FGS scalability structure was designed to be packet-loss resilient especially under unequal packet-loss protection (UPP). However, since the introduction of FGS, there has not been a comprehensive study evaluating its packet-loss resilience under unrecoverable packet losses that are common in Internet streaming applications. In this paper, we evaluate two important aspects of FGS packet-loss resilience. First, we study the impact of applying UPP between the base- and enhancement-layers on FGS-based streams, and we compare equal packet-loss protection (EPP) with UPP scenarios. Second, we introduce the notion of fine-grained loss protection (FGLP), which is suitable for the FGS enhancement-layer, and we develop an analytical framework for evaluating FGLP bounds. Based on these bounds, we show the impact of applying fine-grained protection to the FGS enhancement-layer for different types of video sequences and over a wide range of bit-rates and packet-loss ratios. As illustrated by our extensive simulation results, applying 1) UPP between the base- and enhancement-layers and 2) FGLP for the FGS enhancement-layer can provide significant resilience under moderate-to-high packet-loss ratios (e.g., 5–10%). Furthermore, the merits of this new packet-loss protection technique go beyond the FGS coding scheme, because FGLP can be successfully applied to improve the resilience to packet-losses of other embedded video coding techniques.**

***Index Terms*—Error resilience and control techniques over IP, multimedia coding for IP transmission.**

I. Introduction

INTERNET video streaming applications usually suffer from high packet-loss ratios due to the underlying "best-effort" model of the Internet protocol (IP) [5], [15]. Therefore, many packet-loss recovery mechanisms have been used in conjunction with unreliable transport protocols (e.g., UDPs) that support real-time Internet applications. Nevertheless, there are many cases when the packet-loss recovery mechanism cannot guarantee 100% recovery [1], [4]. Consequently, when unrecoverable packet losses occur, it is very desirable to have a video coding scheme that is resilient to such losses. Moreover, for streaming applications, the selected video scheme should be capable of adapting to the unpredictable variation in bandwidth over the Internet [15], [19], [22]. Fine-granular-scalability (FGS) has been recently developed to meet these requirements [1], [4], and has also been adopted by MPEG-4 as the video-coding tool for streaming applications [3]. Although the bandwidth-adaptability features of FGS have already been covered [4], its resilience to unrecoverable losses over the Internet has not yet been analyzed or studied. Furthermore, the relatively small number of papers that concentrate on the performance of coding techniques on unreliable channels, focus mostly on wireless and ATM environments, and employ traditional scalability schemes which are different from FGS. For example, in [23], Aravind *et al.* compared the performance of an MPEG-2 nonscalable (NS) coder to that of MPEG-2 SNR, spatial and temporal scalability for unreliable ATM-networks. However, in the Internet video streaming case, where larger size packets are lost and the real-time adaptability to the variable bandwidth characteristics is essential, the results mentioned in [23] are no longer valid. Therefore, it becomes clear that new coding and error concealment techniques are required for Internet video streaming.

In [7] and [19], Girod *et al.* investigated the packet-loss resilience of a hierarchical spatio-temporal scalable coder that uses an H.263 compatible base-layer for Internet video streaming. An unequal packet-loss protection (UPP) strategy has been used for transmission, by fully protecting the base-layer using Reed-Solomon (RS) codes. Experimental results showed that this scalable coder has a more gradual degradation in picture quality as the packet loss probability increases, than the NS H.263 codec. However, this method has several drawbacks compared with the MPEG-4 FGS solution: higher complexity (two motion–compensation loops are required), less resilience to packet losses (inter-frame coding at the enhancement-layer leads to error propagation), and lower adaptability to changing bandwidth characteristics. Nevertheless, the efficient forward error correction (FEC) and packetization mechanism proposed in [7] and [19] can be easily extended to MPEG-4 FGS.

In [22], Tan and Zakhor introduced an embedded three-dimensional (3-D) subband coding scheme for real-time video transmission over the Internet. The compressed video was packetized into individually decodable packets of equal expected visual importance and employed a low-delay TCP-friendly protocol for transmission. The performance evaluation under moderate losses revealed better results then an MPEG-1 NS codec. While the bandwidth scalability and error resilience of this fine-granular 3-D subband technique resemble the MPEG-4 FGS characteristics, its performance without losses is lower than for FGS because no motion-compensation is employed. Also, a larger number of frame-memories is necessary, making

Manuscript received September 4, 2001. The associate editor coordinating the review of this paper and approving it for publication was Dr. K. J. Ray Liu.

The authors are with Philips Research, Briarcliff Manor, NY 10510 USA (e-mail: mihaela.vanderschaar@philips.com; hayder.radha@philips.com).

Publisher Item Identifier S 1520-9210(01)10027-1.

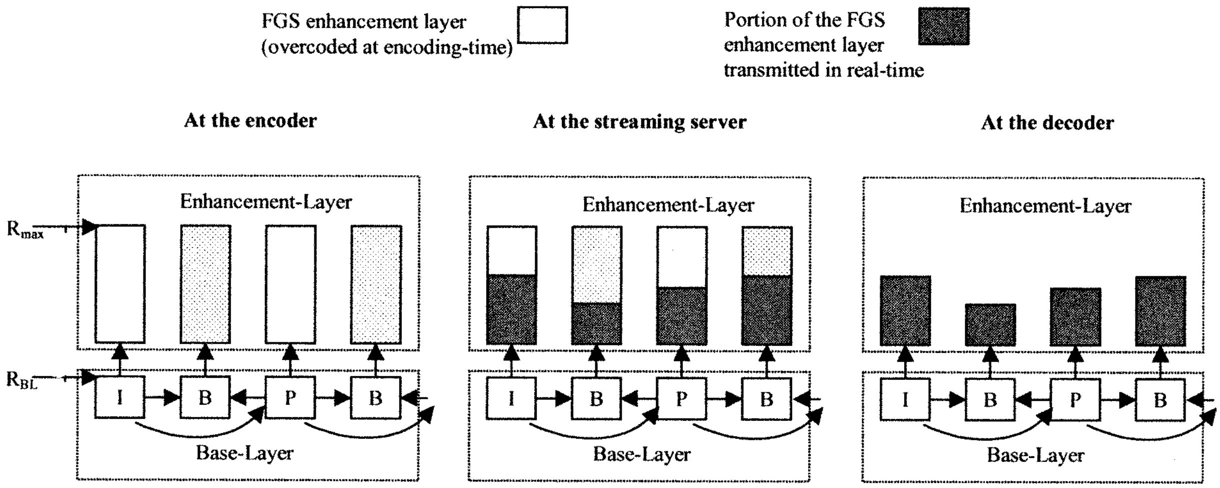

Fig. 1. FGS scalability structure at the encoder, streaming server, and decoder for a typical Internet streaming application.

the solution proposed in [22] more difficult to implement for low-cost and low-power devices. Moreover, for evaluations against NS codecs, the MPEG-4 standard provides a more interesting comparison point then MPEG-1 due to its improved coding performance and more sophisticated resilience tools.

Nevertheless, the 3-D subband coding of [22] as well as the MPEG-4 FGS technique evaluated here, constitute a new class of scalable coding schemes that is able to fulfill two fundamental requirements [15] for video transmission over IP:

1) fast adaptability to changing network conditions by providing bandwidth-scalability, where an *a priori* coded fine-granular (embedded) bit-stream allows decoding at multiple rates;
2) resilience to packet-losses.

In this paper, the MPEG-4 FGS video coding technique is analyzed and evaluated with respect to its robustness for Internet video transmission in the presence of unrecoverable losses. In addition to demonstrating the resilience of FGS over a wide range of bit-rates, we analyze its packet-loss robustness in conjunction with equal packet-loss protection (EPP) and UPP that takes advantage of the scalability structure of FGS. Furthermore, a novel unequal fine-grained loss protection (FGLP) strategy is introduced that is beneficial for the robust Internet video transmission of FGS. The paper is organized as follows. In Section II, the MPEG-4 FGS video coding scheme employed for Internet video streaming is briefly presented together with its packet-loss resilience attributes. The problem statement as well as the main contribution of this paper is also highlighted. Subsequently, in Section III, the FGS performance is determined under both EPP and UPP between the base- and enhancement-layer. Moreover, a comparison between FGS and the NS MPEG-4 codec has been performed from a visual perspective. In Section IV, the impact of UPP within the FGS enhancement-layer is evaluated and the notion of FGLP is introduced together with an analytical framework for evaluating FGLP bounds. Based on these bounds, we show the impact of applying fine-grained protection to the FGS enhancement-layer. In Section V, the conclusions are drawn.

II. FGS and Its Packet Loss Resilience Properties

A. *FGS-Tool for Internet Video Streaming*

FGS has been recently introduced [1], [4], [5] to compensate for the unpredictability and variability in bandwidth between sender and receiver(s) over the Internet. FGS has also been adopted by MPEG-4 as the video-coding tool for streaming applications [3], [12]. The scalability structure of the FGS method is portrayed in Fig. 1. In addition to the base-layer, which is coded with an MPEG-4 compliant NS coder, FGS consists of a single enhancement-layer coded in a progressive (fine granular) manner. Under this framework, the scalable video content can be compressed over any desired bit-rate range $[R_{\min}, R_{\max}]$. The base-layer is coded to a minimally acceptable quality of video. In this case, a bit-rate $R_{\rm BL}$ must be chosen for coding the base-layer such that the available bandwidth (over the time-varying network) is almost certainly higher than $R_{\rm BL}$ at all times ($R_{\rm BL} \leq R_{\min}$). In addition, the base-layer video can be reliably delivered using retransmission or other packet-loss recovery methods [4]. The enhancement-layer improves upon the base-layer video, fully utilizing the available bandwidth at transmission-time. The enhancement-layer is coded progressively (bit-plane by bit-plane[1]) employing either an embedded wavelet or embedded DCT coding scheme [10]. All the results shown in this paper are based on the low-complexity bit-plane embedded-DCT approach that has been adopted by the MPEG-4 standard for coding the FGS enhancement-layer [3], [20]. The enhancement-layer frames are intra-coded, but the coding efficiency from temporal redundancy exploitation is partially retained because the MPEG-4 motion compensated scheme is employed at the base-layer. A more detailed description of the FGS technique can be found in [3] and [4].

The enhancement-layer is over-coded using a bit-rate $(R_{\max} - R_{\rm BL})$. It is important to note that the range $[R_{\min}, R_{\max}]$ can be determined off-line (e.g., for a particular set of Internet access technologies). For unicast streaming, an estimate for the available bandwidth R can be generated in

[1]In a progressive coder, the more significant bit-planes are transmitted prior to the less significant bit-planes.

real-time for a particular session. Based on this estimate, the server transmits the enhancement-layer using a bit-rate $R_{\rm EL} = \min(R_{\max} - R_{\rm BL}, R - R_{\rm BL})$. Due to the fine-granularity of the enhancement-layer, its real-time rate control aspect can be implemented with minimal processing.

For multicast streaming, a set of intermediate bit-rates, $R_1, R_2, \ldots, R_N$, can be used to partition the enhancement-layer into substreams. In this case, N fine-granular streams are multicasted using the bit-rates: $R_{e1} = R_1 - R_{\rm BL}, \ldots, R_{eN} = R_N - R_{N-1}$, with $R_{\rm BL} < R_1 < \cdots < R_N \leq R_{\max}$.

Therefore, in both unicast and multicast scenarios, FGS provides the server with total flexibility in adapting to the network condition without adding significant complexity. Furthermore, when compared with other scalable approaches, FGS has shown to provide a good balance between coding-efficiency and scalability [6]. Nevertheless, when compared with a NS (i.e., single-layer) MPEG-4 codec operating at a particular bit-rate and without packet-losses, FGS performance is lower, due to the limited temporal redundancy exploitation within the FGS enhancement-layer. To adapt to the large bandwidth variations over the Internet, the MPEG-4 NS codec can also be employed to encode and distribute video sequences at several "key" discrete rates corresponding to typical connection speeds. Thus, at encoding time (i.e., off-line), several NS streams are coded at different bit-rates. Subsequently, at transmission time, the server switches among these streams in real-time in response to changing conditions over the network (e.g., changes in available bandwidth or packet-loss events) to best match the client's available bandwidth. In the remainder of this paper, the FGS performance under packet-losses is compared against that of the MPEG-4 NS coder employing the previously described dynamic stream switching mechanism.

B. Error Resilience Properties of FGS

Another important characteristic of the FGS coding scheme, which was designed for Internet video streaming, is its resilience to IP packet-losses. Several characteristics of the FGS coding scheme contribute to its robustness.

1) *No error propagation*. Since the FGS enhancement-layer is intra-frame coded, the impact of packet losses is localized to the particular enhancement-layer picture(s) experiencing the losses.
2) *Unequal packet-loss protection (UPP)*. Since packet-losses are inherent to Internet transmission, unequal protection can help allocate system resources (bandwidth) more efficiently and provide a good quality of service. As FGS is a scalable coding scheme, consisting of a base and an enhancement-layer, different error-protection schemes could be employed for the two layers. In particular, reliable delivery of the base-layer video is of prime importance and consequently, a higher protection level can be supported for this layer.
3) *Lower packet loss probability for the FGS base-layer than for single-layer*. Since the amount of data to be protected is lower for the FGS base-layer than for a NS coder operating at the same bit-rate, for an identical packet loss rate, the number of packets lost from the FGS base-layer is lower than for the NS stream. This implies that if equal error protection or retransmission is employed, a smaller percentage of significant data will be corrupted in the FGS case.
4) *Error resilience over a wide range of bit-rates and packet-losses*. Since the packet-loss rates and bandwidth availability are unknown at encoding time, a robust compression algorithm should be able to adapt to changing bandwidth. As it will be shown in Section III, our experiments revealed that FGS performance steadily increased with the bit-rate if UPP is used.
5) *Better FGS performance compared to conventional SNR scalability*. Due to the fine-granularity in FGS, the different bit-planes covering the entire image could be packetized separately, unlike the case of discrete multi-layer SNR scalability, where several bit-planes are coded together. Therefore, in the FGS case, if a less significant bit-plane in the enhancement-layer is lost, the entire image can still be enhanced using the already received more significant bit-planes. However, with traditional SNR scalability, if a part of the enhancement-layer is lost, the entire corrupted layer and the already decoded enhancement-layers for that frame need to be skipped to avoid annoying artifacts due to varying qualities within the image.

C. Scope of the Paper

FGS has recently been adopted in the MPEG-4 standard as the core video-compression method for streaming applications because it is capable of easily adapting to the unpredictable variation in bandwidth over the Internet. However, the resilience of FGS to unrecoverable packet-losses has not yet been established, and constitutes an equally important component in the evaluation of FGS for video transmission over IP. It is the aim of this paper to establish the resilience of FGS to unrecoverable packet-losses and to determine protection strategies that increase its robustness. Furthermore, all previous comparisons between FGS and the NS MPEG-4 coder have been performed in the absence of packet-losses. However, for video transmission over IP, the impact of unrecoverable packet-losses on the performance of both codecs needs to be considered for a realistic comparison. Therefore, we compared the newly introduced MPEG-4 FGS tool with the MPEG-4 NS codec for different types of video sequences and over a wide range of bit-rates and packet-loss ratios. The comparison has been performed for both EPP and UPP.

Another contribution of this paper is that the UPP concept is not only applied between base- and enhancement-layers, like in most alternative scalable coding techniques [7], [19], but also prioritization within the enhancement-layer is employed. For example, taking advantage of the fine-granularity of FGS, a higher level of protection can be assigned to the more significant bit-planes. Consequently, two UPP-based aspects were addressed regarding the packet-loss resilience of FGS.

1) The impact of providing (virtually) 100% protection for the base-layer while allowing the enhancement-layer to suffer significantly higher packet loss ratios. In this case, although different levels of protection are used for the

base- and enhancement-layers, the entire data within the enhancement-layer is treated equally.

2) The impact of exploiting the embedded nature of the enhancement-layer by providing UPP within the fine-granular FGS layer.

These two aspects of the FGS packet-loss resilience are investigated in-depth within this paper. First, in Section III, we compare equal and unequal packet-loss-protection scenarios over a wide range of bit-rates. Subsequently, in Section IV, we address the second aspect of the FGS packet-loss-resilience analysis: the impact of UPP within the FGS enhancement-layer. We exploit the fine granular nature of the FGS enhancement-layer and introduce the notion of FGLP. Under FGLP, in addition to providing a higher level of protection for the base-layer, the enhancement-layer is partitioned into an arbitrary number of fine-grained sublayers, to which different levels of protection can be assigned. The merits of this new packet-loss protection technique go beyond the FGS coding scheme, because FGLP can be successfully applied to improve the resilience to packet-losses of other embedded video coding techniques [13], [14].

III. FGS Resilience Under Equal and Unequal Packet-Loss Protection Between Base- and Enhancement-Layers

A. The Effective Packet-Loss Ratio (EP) Parameter

Many extensive studies performed over the Internet revealed that the packet loss ratio can vary widely (e.g., 1%–20%) [3], [5], [15]. Moreover, the amount of lost-packet recovery can also be considerably different (e.g., 0%–100%) depending on the network condition and the recovery scheme employed [3], [7]. In that regard, an important parameter, called the effective packet-loss-ratio (EP), is introduced to express the relation between the actual packet-loss-ratio (PLR) and the recovery ratio (RR), which is the percentage of lost packets that are recovered. The EP is

$$\mathrm{EP} = \mathrm{PLR}(1 - \mathrm{RR}).$$

For example, if a stream experiences a PLR of 10%, and due to a good packet-loss recovery mechanism 90% of those lost packets are recovered on time (i.e., $\mathrm{RR} = 0.9$), then the effective packet loss ratio is only 1% ($\mathrm{EP} = 0.01$). Therefore, EP represents the amount of unrecoverable packets.[2]

The EP parameter also provides a simple tool for emulating different levels of UPP. In the case of scalable streams consisting of a base-layer and an enhancement-layer (such as FGS), UPP is achieved by providing a higher level of protection[3] for the base-layer which leads to a higher RR (or a lower EP). This translates into different values $\mathrm{EP}_{\mathrm{BL}}$ and $\mathrm{EP}_{\mathrm{EL}}$ of EPs for the base and enhancement-layers, respectively. To evaluate the performance of different levels of UPP between the base and enhancement-layers, we express the overall EP as a weighted sum of $\mathrm{EP}_{\mathrm{BL}}$ and $\mathrm{EP}_{\mathrm{EL}}$

$$\mathrm{EP} = \mathrm{EP}_{\mathrm{BL}}\, p_{\mathrm{BL}} + \mathrm{EP}_{\mathrm{EL}}\, p_{\mathrm{EL}}$$

where p_{BL} (p_{EL}) is the probability that a packet chosen at random from a scalable stream is a base-layer (enhancement-layer) packet. If $R_{\mathrm{T}} = R_{\mathrm{BL}} + R_{\mathrm{EL}}$ is the total bit-rate, then the above expression can be estimated as

$$\mathrm{EP} = \mathrm{EP}_{\mathrm{BL}} \left(\frac{R_{\mathrm{BL}}}{R_{\mathrm{T}}}\right) + \mathrm{EP}_{\mathrm{EL}} \left(\frac{R_{\mathrm{EL}}}{R_{\mathrm{T}}}\right).$$

For given values of EP and bit-rates (R_{BL} and R_{EL}), the resilience of a scalable stream can be evaluated by varying the EPs for the base and enhancement-layers (i.e., $\mathrm{EP}_{\mathrm{BL}}$ and $\mathrm{EP}_{\mathrm{EL}}$). One key advantage of this formulation is that it provides an abstract framework for evaluating the resilience of scalable streams independent of the particular mechanism(s) used for delivering and protecting the data. In Section III-B, we use this framework to evaluate the packet-loss resilience of FGS streams under both EPP and UPP between the base- and enhancement-layer. In Section IV, we derive a general expression for the EP under the framework of FGLP.

B. Simulation Results for FGS and Nonscalable Streams Under EPP

In this section, we show the simulation results obtained by applying one extreme case of EPs to the FGS base-layer: $\mathrm{EP}_{\mathrm{BL}} = \mathrm{EP}_{\mathrm{EL}} = \mathrm{EP}$. This case represents the scenario when both the base- and enhancement-layer streams are treated equally [i.e., equal packet-loss protection (EPP)]. Therefore, this represents the worst-case scenario for the FGS packet-loss resilience. For comparison, we apply the same EP to NS streams coded using MPEG-4. Since both the FGS base-layer and the NS streams suffer from packet-losses, a robust and consistent approach needs to be adopted for concealing the lost video data. The Appendix provides a brief description of the video concealment method we used for the FGS base-layer and NS streams.

For the experiments, several video sequences have been evaluated that contain different variations of textures and motion, thereby providing a good representation of the content transmitted over the Internet. All sequences were coded at CIF-resolution in progressive 4:2:0 YUV format at 10 frames/s. Each sequence was coded using MPEG-4 at a bit-rate R_{BL} to provide a common base-layer stream for both FGS and the NS case. For FGS, a single enhancement-layer was generated to cover a bandwidth range starting from the base-layer bit-rate R_{BL} until a maximum bit-rate $R_{\max}$ (e.g., for cable-modem Internet access technologies, $R_{\mathrm{BL}} \cong 100$ kbit/s and $R_{\max} \cong 1$ Mbit/s). For the NS streams, in addition to the common base-layer (obtained by employing a fixed quantization parameter (QP) of 28), the MPEG-4 NS encoder was used to generate several single-layer streams coded with a fixed QP of 15, 10, 6, and 3, respectively. Subsequently, FGS has been used to cover the same bandwidth range as the NS streams. The FGS rate-control is extremely simple, involving simply "cutting" the embedded coded layer at the desired rate. In the experiments, a GOP structure

[2]The terminology "unrecoverable" packets refer to packets that are not received by the time needed for decoding. In some cases, packets are eventually received by the client but they arrive after their decode time.

[3]Higher protection for the base layer can be realized either through transmission over reliable channels (e.g., with quality-of-service guarantees), by employing priority retransmission, or by using more robust FEC codes.

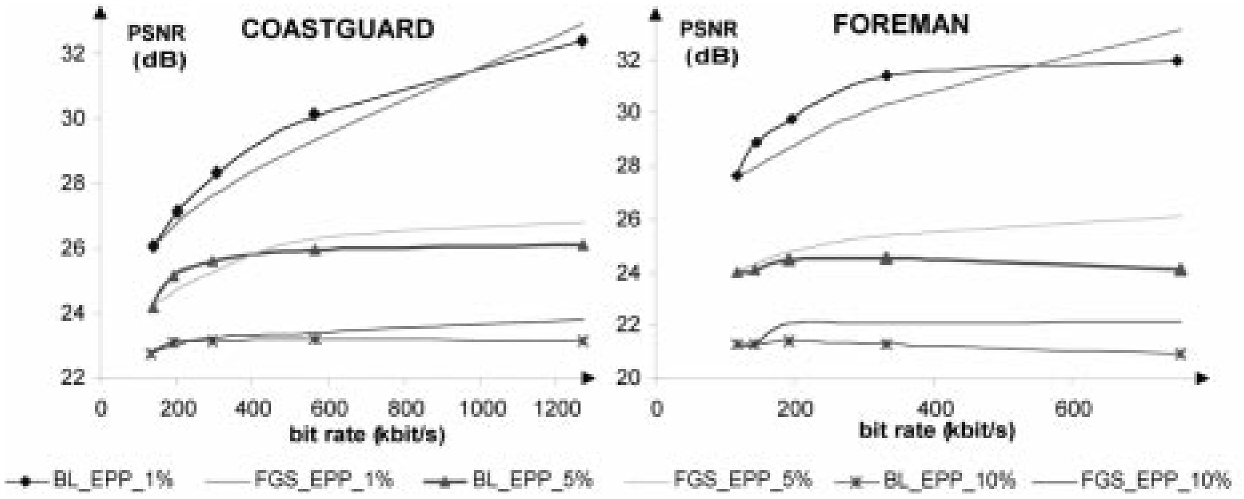

Fig. 2. Packet-loss resilience of NS and FGS under EPP.

of 21 frames has been used (i.e., $N = 21$). The NS and FGS base-layer streams also included B-frames ($M = 3$), which improved the overall resilience of these streams significantly. (Experiments performed with a GOP-structure of $N = 21$ and $M = 1$ revealed that if no B-frames are used, the performance of both FGS and the NS coder decreases. The better resilience of GOP-structures with $M > 1$ was expected because packet-losses in B-frames are not propagating, since they are not used as a reference for motion-compensation. However, since the packet loss probability for the FGS base-layer is lower than for the single-layer at the same transmission rate, FGS performance is less unprotected by the absence of the B-frames.)

Due to the random nature of Internet packet losses and their impact on compressed video, at each tested bit-rate and for each EP value, 50 different runs of the experiments were conducted for each FGS and NS single-layer-stream. For each experiment, a different packet-loss pattern was generated by a random number generator initialized each time with a different seed. The SNR value computed from each experiment was used to determine the average SNR plotted in the following figures. Furthermore, the objective measurements were verified by subjective evaluations performed on an extensive set of sequences containing various motion-characteristics and textures. Finally, it is important to mention that in our experiments, each IP packet carries a maximum of 492 bytes of video payload.[4]

We conducted the equal-packet-loss protection tests for several EP values. Fig. 2 shows the results of two MPEG-4 test sequences—*Foreman* and *Coastguard*—under three typical EP values [15]: 1%, 5%, and 10%. The figure shows that even under this worst-case scenario for FGS (i.e., equal protection for the base and enhancement layers), FGS streams are more packet loss resilient than NS streams, in particular at 5% and 10%. In addition, while FGS performance either improves or stabilizes as a function of the bit-rate, the performance of NS coding could actually degrade as the bit-rate increases. This is because for a given effective-packet-loss ratio, losses impact a larger number of pictures at higher than at lower bit-rates, leading to more degradation for the NS streams as the bit-rate increases. The FGS structure, by design, prevents the propagation of the corrupted image-areas and limits the impact of packet losses to the unprotected enhancement-layer frames only.

Also, it is very important to mention that due to its fine-granularity, FGS can be truncated at each particular bit-rate. Thus, the FGS performance plots of Fig. 2 are truly continuous. However, for the NS streams, the continuous performance plots depicted in Fig. 2 represent an ideal case that can be approached only if an extremely large number of NS streams at a multitude of bit-rates were *a priori* coded and stored at the server side.

Another expected observation which can be made from Fig. 2 is that sequences with higher motion (*Foreman*) are more unprotected by packet-losses than lower motion sequences (*Coastguard*).

C. Simulation Results for FGS Streams Under UPP

In this section, we show the simulation results obtained by applying another extreme case of EPs: $\mathrm{EP_{BL}} = 0$. Therefore, this case simulates the scenario of 100% protection of the FGS base-layer. Hence, all the losses are clustered in the FGS enhancement-layer and $\mathrm{EP_{EL}} = \mathrm{EP}(R_{\mathrm{T}}/R_{\mathrm{E}})$. It is important to note that in this case we treat all enhancement-layer packets as equally important. Therefore, although this case is an extreme (best-case) scenario in differentiating between the base-layer and the enhancement-layer, it is a worst-case scenario in differentiating among the enhancement-layer packets. Fig. 3 shows the simulation results[5] comparing the performance of FGS under this scenario (UPP) and the previous scenario (EPP). The figure clearly shows the significant improvements that can be obtained by using unequal packet loss protection between the FGS base- and enhancement-layer. These results are in line with the conclusions of [7] and [19], which evaluated the packet-loss resilience of a hierarchical spatio-temporal scalable

[4]This number is selected to adhere to certain Internet networking constraints that limit the size of an IP packet. Choosing large packet sizes could result in higher packet loss ratios, while very small size packets result in high overhead (Section IV-B).

[5]The simulations have been performed under the same testing conditions as employed for EPP (see Section III-B).

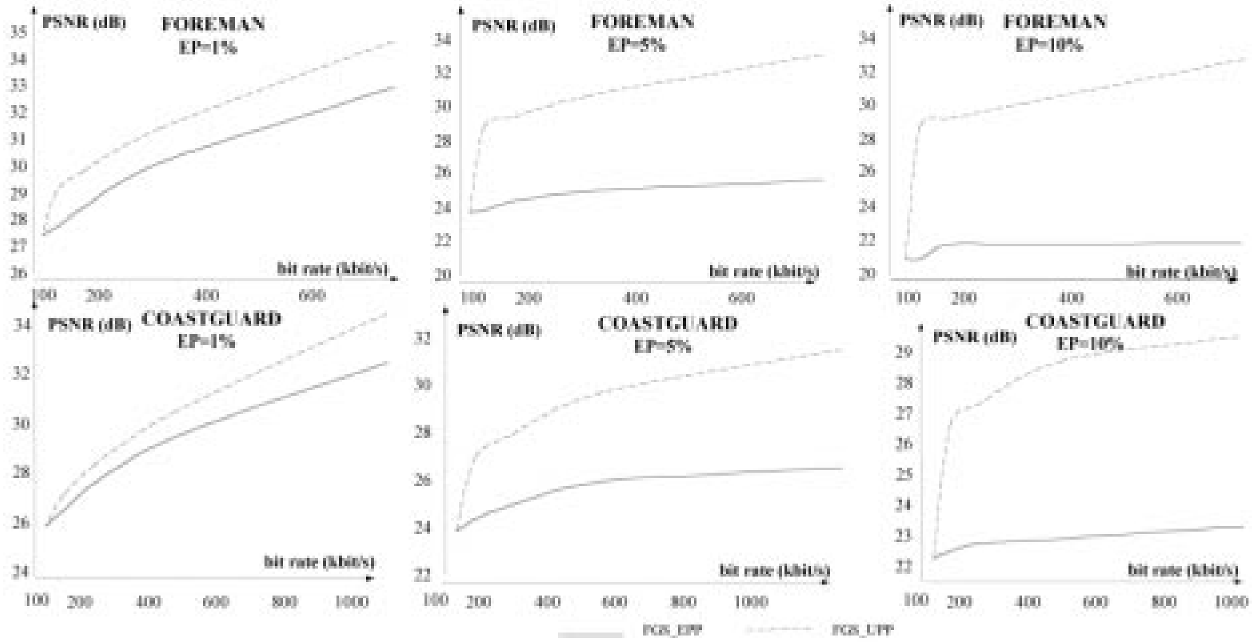

Fig. 3. Packet-loss resilience of FGS under UPP and EPP. The abrupt initial slope for FGS under UPP is because initially (at low bit-rates) the transmitted stream mainly consists of the base-layer (i.e., there is only enough bandwidth for the base-layer and a small portion of the enhancement layer). Consequently, the base-layer is subjected to packet losses, since it constitutes a higher percentage of the whole stream than the packet-loss ratio. However, as the bit-rate increases, a larger part of the enhancement-layer is transmitted, and thus UPP can be more effective in protecting the base-layer.

coder using better packet-loss protection mechanisms for the base-layer.

For the NS case, we employed a UPP strategy that assigns different priorities to the various frame-types. (Different UPP-strategies can be employed for the NS codecs, but they would require an increased server complexity compared with the FGS scenario.) This is because each frame-type has a different impact upon the overall-image quality of the video sequence.

- The I-frames have the highest priority since all the frames in the GOP are predicted based on them.
- The P-frames have a lower priority than the I-frames and higher than B-frames, because they are used for the prediction of subsequent P- and B-frames. Also, the closer a P-frame is to the beginning of a GOP, the higher its priority is. This is because subsequent P- and B-frames rely on it.
- The B-frames have the lowest priority because no other frames use them for prediction, and thus an error in a B-frame is limited to the frame experiencing the loss.

For FGS, only the base-layer data, which represents 10%–50% of the entire decoded stream (depending on the transmission bit-rate), needs to be protected. However, for the NS streams, the I- and P-frames represent more than 50% of the data within the GOP independent of the transmission bit-rate (if at least two B-frames are coded between each subsequent P-frames). Therefore, in order to provide a fair comparison between FGS and the NS streams under UPP, the same amount of data needs to be protected in both cases, namely, $R_{\rm BL}$—the size of the FGS base-layer. At higher bit-rates, the number of packets carrying I- and P-frames data is larger than the number of packets for $R_{\rm BL}$.

Hence, if $np^{\rm I}$ and $np^{\rm P}$ are the number of packets carrying I- and P-frames respectively, then at higher bit-rates, $np^{\rm I}+np^{\rm P} \geq$

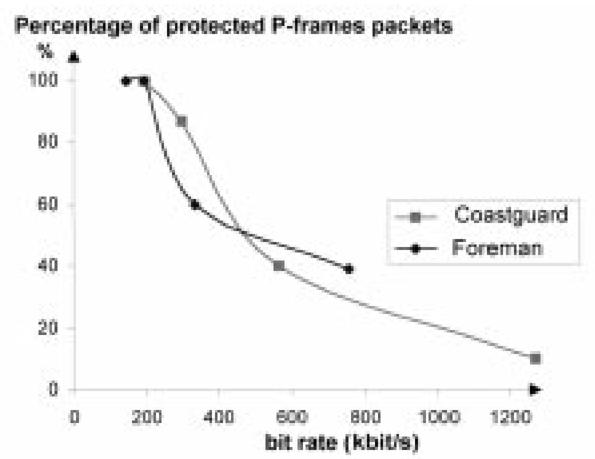

Fig. 4. Percentage of protected P-packets by the UPP mechanism. (The bit-rate range for the two sequences is different since the simulation was based on constant QP rate-control).

$np^{\rm BL}$, with $np^{\rm BL}$ being the total number of packets for the base-layer. Consequently, only the I-frames and the first k_1 packets of P-frames within each GOP can be protected (see Fig. 4), while the remainder k_2 packets of the P-frames $(np^{\rm P} = k_1 + k_2)$ and the B-frames are left unprotected

$$\begin{aligned} np^{\rm protected} &= np^{R_{\rm BL}} = np^{\rm I} + k_1 \\ \text{and } np^{\rm unprotected} &= k_2 + np^{\rm B}. \end{aligned}$$

The unprotected packets of a NS stream under UPP have then an EP of

$$\mathrm{EP}_{\rm UPP}^{\rm NS} = \mathrm{EP} \times \frac{np^{\rm I} + np^{\rm P} + np^{\rm B}}{np^{R_{\rm BL}}}.$$

Fig. 5. Frame #14 from the *Foreman* sequence coded at 332 kbit/s, with NS $\mathrm{QP} = 6$ (left) and FGS with UPP (right) both after one packet-loss event. For FGS, the entire enhancement-layer of this frame is lost (i.e., a worst case scenario), and the image quality is reduced to the base-layer quality. However, in the NS case, due to a packet-loss in a previous frame, the error propagated, leading to a distorted image.

From Fig. 4 it can be concluded that many NS coded frames remain unprotected for packet-losses. Whenever a packet is lost in a NS coded P-VOP, the decrease in quality becomes noticeable, since the frame-rate is relatively low (10 frames/s) and the image is often considerably distorted (see Fig. 5). The perceptual image quality is especially affected for sequences with a high-degree of motion, since the concealment techniques are not very effective in this case.[6]

For the FGS coding scheme under UPP, the impact of packet-losses on the image quality is considerably lower. The reason is that in the FGS UPP case, if a packet is lost, the image quality of the affected frame does not drop below the base-layer quality. Then, the subjective evaluation of the FGS video-clips under UPP revealed that even if an entire enhancement-layer picture is lost, the human visual system is capable of filtering this loss between consecutive frames, such that the perceptual image quality loss is negligible.

These experiments revealed that if UPP is employed between the base- and enhancement-layers of the FGS streams, significant resilience can be obtained independent of the bit-rate and packet-loss ratio. Therefore, in Section IV, we will focus on improving the FGS performance under UPP by applying different levels of protection within the FGS enhancement-layer using FGLP.

It is important to mention that the FGS coding scheme employed in the previous evaluation does not address frame-rate scalability, such that the reconstructed video from an embedded FGS bit-stream has the same frame-rate regardless of the bit-rate. Recently, in MPEG-4, a combination of FGS with temporal-scalability, referred to as hybrid SNR-temporal scalability [24], has been introduced to allow trade offs between individual image quality (SNR) and motion-smoothness (increased frame-rate). This mechanism further improves the error resilience of FGS under UPP by allowing the B-frames to be coded as part of the enhancement-layer [25]. In this way, only the I- and P-frames of the FGS base-layer need to be protected.

[6]Note that the packets used for Internet video streaming are large (i.e., contain a considerable part of an image) and thus the error concealment techniques are less effective than in the case of bit-errors or smaller packets.

IV. FGS Resilience Under Fine Grained Loss Protection (FGLP) of the Enhancement Layer

A. Framework for the Computation of EP Under FGLP

As shown in Section III, providing higher levels of protection for the base-layer of FGS streams (i.e., UPP) leads to significant packet-loss resilience. In this section, we build on this result by evaluating FGS resilience under UPP within the enhancement-layer.

Under FGLP, the enhancement-layer is partitioned into an arbitrary number n of embedded sublayers. Therefore, there are a total of $n+1$ layers (one base plus the enhancement sublayers), where each sublayer k has its own EP, EP_k.[7]

In the sequel, an analytical framework is developed for the computation of EP_k under different FGLP-strategies. EP_k is a function of the transmission bandwidth, number of packets in each sublayer, the overall EP value under consideration and last but not least, the error protection strategy adopted for that layer. As mentioned above, we first extend the EP analysis tool to the FGLP strategy.

The generic equation employed for the effective packet-loss computations is $\mathrm{EP} \times R_{\mathrm{tot}} = \sum_{k=0}^{n}(\mathrm{EP}_k \times R_k)$, with

$R_{\mathrm{tot}} =$ the total bandwidth available at transmission time;
$\mathrm{EP} =$ the overall EP;
R_k, $k > 0 =$ the rate of the kth enhancement sublayer;
$R_0 = R_{\mathrm{BL}} =$ the base-layer rate;
$\mathrm{EP}_k =$ the EP of sublayer k.

Then, if R_{enh} is the FGS enhancement layer rate that can be sent at transmission time

$$R_{\mathrm{tot}} = \sum_{k=0}^{n} R_k = R_{\mathrm{BL}} + R_{\mathrm{enh}} = R_{\mathrm{BL}} + \sum_{k=1}^{n} R_k.$$

For simplicity, the EPs of the different layers can be expressed in terms of the EP of the first enhancement sublayer EP_1, $\mathrm{EP}_k = \alpha_k \times \mathrm{EP}_1$, if $k > 0$, and $\mathrm{EP}_{\mathrm{BL}} = \mathrm{EP}_0 = \alpha_0 \times \mathrm{EP}_1$ and $\alpha_1 = 1$.

[7]It is important to note that in this section, the enhancement sublayer k does not represent an FGS bit-plane, but a portion of the enhancement layer containing a certain number of packets, to which the same priority has been assigned. Thus, depending on the size of the various (coded) bit-planes, data from one or more bit-planes is included in each enhancement sublayer k, and different portions of the same bit-plane can have different priorities if they lie in different enhancement sublayers.

Then, the generic equation above becomes

$$\mathrm{EP} \times R_{\mathrm{tot}} = \left(\alpha_0 \times R_{\mathrm{BL}} + \sum_{k=1}^{n} \alpha_k R_k\right) \times \mathrm{EP}_1.$$

This equation provides a generic framework for expressing the EPs of the different FGS sublayers under a wide range of fine-grained protection strategies. Below, we derive special cases from this generic expression.

- **The Equal Packet-loss Protection (EPP) scenario:** In the EPP-case, all layers (including the base-layer) are equally protected (i.e., $\alpha_k = 1$, $k \geq 0$), and thus:

$$\mathrm{EP}_0 = \mathrm{EP}_1 = \cdots = \mathrm{EP}_k = \cdots = \mathrm{EP}_n = \mathrm{EP}.$$

- **The Unequal Packet-loss Protection (UPP) scenarios:** In the UPP case, we distinguished three different scenarios.

  Case UPP_A (EPP within the enhancement-layer—Worst protection case for the FGS enhancement-layer): The base-layer is fully protected ($\alpha_0 = 0$), and the enhancement sublayers are equally protected ($\alpha_k = 1, k > 0$). In this case, $\mathrm{EP}_k = \mathrm{EP}_1 = \mathrm{EP}_{\mathrm{enh}}$ and $\mathrm{EP} \times R_{\mathrm{tot}} = R_{\mathrm{enh}} \times \mathrm{EP}_{\mathrm{enh}}$, leading to $\mathrm{EP}_{\mathrm{enh}} = \mathrm{EP} \times R_{\mathrm{tot}}/R_{\mathrm{enh}}$. This particular UPP case was described in detail in Section III-C.

  Case UPP_B (Example of realistic FGLP strategy). The base-layer is fully protected ($\alpha_0 = 0$) and the protection of each more significant enhancement sublayer increases with a factor of e.g., $p\,(\alpha_k = p^{k-1}, k > 0)$, leading to $\mathrm{EP}_k = p^{k-1}\mathrm{EP}_1$ and $\mathrm{EP} \times R_{\mathrm{tot}} = \mathrm{EP}_1 \times \sum_{k=0}^{n-1} p^k R_{k+1}$. Thus, $\mathrm{EP}_1 = (\mathrm{EP} \times R_{\mathrm{tot}})/\sum_{k=0}^{n-1} p^k R_{k+1}$.

  If $k = n = 1$, the enhancement-layer has just one protection strategy, and in this case UPP_A and UPP_B are identical. It is important to remember here that EP_k is used as a measure of the packet-loss protection that should be provided for a given sublayer k. The higher EP_k is, the lower protection is given to the corresponding sublayer k.

  Case UPP_C (Ideal protection strategy for FGS): The base-layer is fully protected and all the errors are clustered in the least-significant enhancement sublayers ($\alpha_0 = \cdots = \alpha_{l-1} = 0$, and $\alpha_l = \cdots = \alpha_n = 1$), leading to $\mathrm{EP}_0 = \cdots = \mathrm{EP}_{l-1} = 0$ and $\mathrm{EP}_l = \cdots = \mathrm{EP}_n = \mathrm{EP}_{\mathrm{loss}}$.

Consequently, $\mathrm{EP} \times R_{\mathrm{tot}} = \mathrm{EP}_{\mathrm{loss}} \times \sum_{k=l}^{n} R_k$ and $\mathrm{EP}_{\mathrm{loss}} = (\mathrm{EP} \times R_{\mathrm{tot}})/\sum_{k=l}^{n} R_l$. l in the previous formula is defined as being the largest integer smaller or equal to n, for which the resulting $\mathrm{EP}_{\mathrm{loss}}$ is smaller than 100%. Moreover, it can also be seen that UPP_A is a subcase of UPP_C, when $l = 1$.

In the sequel of this section, we want to determine from a theoretical standpoint the efficiency of the previously described protection strategies. For this analysis, we assumed for simplicity[8] that each sublayer has the same bit-rate, i.e., $R_1 = \cdots = R_k = \cdots = R_n = R_{\mathrm{enh}}/n$. In this case, the enhancement sublayer k does not represent an FGS bit-plane, but a portion of the enhancement layer containing a certain number of packets, to which the same priority has been assigned. Thus, depending on the size of the various (coded) bit-planes, data from one or more bit-planes is included in each enhancement sublayer k, and different portions of the same bit-plane can have different priorities if they are in different enhancement sublayers. Furthermore, for the UPP_B protection strategy, p is set to a value of 2 in the experiments presented in Section IV-C, i.e., a doubling protection strategy is implemented. For $p = 2$ and $R_1 = \cdots = R_k = \cdots = R_n = R_{\mathrm{enh}}/n$, EP_1 in the UPP_B case becomes $\mathrm{EP}_1 = \mathrm{EP} \times R_{\mathrm{tot}}/R_{\mathrm{enh}} \times n/(2^n - 1)$ and $\mathrm{EP}_k = 2^{k-1}\mathrm{EP}_1$, while in the UPP_C scenario, $\mathrm{EP}_{\mathrm{loss}} = \mathrm{EP} \times R_{\mathrm{tot}}/R_{\mathrm{enh}} \times (n/n - k + 1)$.

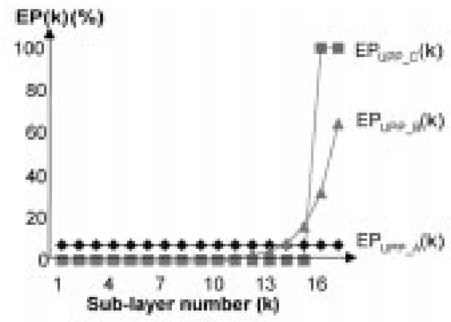

Fig. 6. Packet-loss ratios of each sublayer $\mathrm{EP}(k)$ for the three UPP-strategies.

In Fig. 6, the EP_k for each sublayer k is portrayed for all three UPP strategies introduced above—$\mathrm{EP}_{\mathrm{UPP\_A}}(k)$, $\mathrm{EP}_{\mathrm{UPP\_B}}(k)$, and $\mathrm{EP}_{\mathrm{UPP\_C}}(k)$. The plot is given for a number of enhancement sublayers n equal to 17. It is important to mention that independent of the UPP-strategy adopted, the area below its corresponding function—$\mathrm{EP}_{\mathrm{UPP\_A}}(k)$, $\mathrm{EP}_{\mathrm{UPP\_B}}(k)$, or $\mathrm{EP}_{\mathrm{UPP\_C}}(k)$, is the same and equals $\mathrm{EP} \times R_{\mathrm{tot}}/R_{\mathrm{enh}} \times n$. Moreover, for the two extreme limits —$n = 1$ or n is very large, the UPP_B case approaches UPP_A and UPP_C, respectively.

In Section IV-C, the theoretical performance under losses for these three UPP strategies is determined. In practice, while the ideal UPP_C case cannot easily be realized, the doubling priority strategy adopted in the UPP_B scenario can be accomplished by employing for instance the unequal FEC, as proposed in Section IV-D. Therefore, the performance analysis in Section IV-C is meant to demonstrate that the realistic FGLP scenario of UPP_B does lead to a very similar performance (under losses) to that of the ideal UPP_C case. Moreover, in Section IV-B, we determine the level of granularity necessary to obtain a good performance under FGLP (for the practical UPP_B scenario).

B. Optimal Granularity for FGLP

Let us assume that in the UPP_B case, each sublayer is made of a single packet of size P. Each packet includes 1) the FGS payload (P_{FGS}), and 2) the packet overhead[9] (P_{OH}). Therefore, $P = P_{\mathrm{FGS}} + P_{\mathrm{OH}}$. Then, if an FGS enhancement-layer frame

[8]However, other protection strategies can also be envisaged. For example, to each bit-plane a different EP_k can be assigned.

[9]The packet overhead typically consists of the total IP, UDP, and RTP packet overhead.

has the size S_{FGS}, its corresponding number of packets equals $n = S_{\mathrm{FGS}}/P = S_{\mathrm{FGS}}/P_{\mathrm{FGS}} + P_{\mathrm{OH}}$. Consequently, n also represents the number of FGS sublayers.

It becomes therefore clear that depending on the packet-size P, and the corresponding n value, the level of protection of the various sublayers (packets) in the UPP_B case also varies. To illustrate this, in Fig. 7, the $\mathrm{EP}_{\mathrm{UPP\_B}}(k)$ function is plotted for various FGS payload sizes—$P_{\mathrm{FGS}}^1 = 984$ bytes, $P_{\mathrm{FGS}}^2 = 492$ bytes and $P_{\mathrm{FGS}}^3 = 246$ bytes and their corresponding number of sublayers—n_1, n_2, and n_3, respectively. The packet overhead (P_{OH}) is the same in all cases and equals 40 bytes. To be able to compare the $\mathrm{EP}(k)$ functions corresponding to the three packet-sizes, we plotted them with respect to the same coordinates system. Therefore, the unit used for each sublayer on the horizontal axis is equal to the smallest packet-size considered, i.e., $P_{\mathrm{FGS}}^3 = 246$ bytes. Hence, since e.g., P_{FGS}^2 has twice the size of P_{FGS}^3, it corresponds to two units (i.e., two sublayers of size P_{FGS}^3) on the horizontal axis. Thus, for P_{FGS}^2, two consecutive units of size P_{FGS}^3 (on the horizontal axis) have the same EP-value, since they appertain to a single packet. For $P_{\mathrm{FGS}}^1 = 4P_{\mathrm{FGS}}^3$, four consecutive units of size P_{FGS}^3 have the same EP-value.

Furthermore, the areas below the $\mathrm{EP}(k)$- functions of the three shown FGS payloads are the same. From Fig. 7, the following observations can be made.

1) A smaller packet-size (i.e., larger n) leads to very good protection of the first packets that contain the most significant bitplane(s) and to very limited protection for the last packets containing the least significant bitplane(s).
2) Conversely, a larger packet-size (i.e., smaller n) leads to more uniform protection between the packets.

Another observation that can be made is that even though $P_{\mathrm{FGS}}^3 = P_{\mathrm{FGS}}^2/2$, and S_{FGS} is constant, n_3 is smaller than $2n_2$, due to the packet overhead P_{OH}. Therefore, a smaller packet-size leads to more granularity, but also to increased overhead which reduces the effective FGS payload data.

As will be seen in Section IV-C, for a given bit-rate and EP, an optimal packet-size can be determined for which the best FGLP performance under UPP_B can be obtained.

C. Simulation Results

To evaluate the packet-loss resilience of FGS under the fine-grained protection framework derived above, the same conditions as the ones described in Section III-B were employed. Fig. 8 shows the performance of FGS under the three cases of fine-grained UPP scenarios outlined above (UPP_A, UPP_B, and UPP_C) for EPs of 1%, 5%, and 10%. From the plots, it is clear that FGS packet loss-resilience can be improved by employing UPP within the enhancement-layer. Depending on the protection strategy employed, the FGS rate-distortion performance under heavy losses (i.e., high packet-loss ratios) can vary considerably. In many instances, the UPP_B case has a similar performance to the ideal UPP_C case. This makes the protection strategy of UPP_B an attractive alternative since it is relatively easy to realize, and it also provides a very good performance. The FGLP strategy is especially valuable for high packet-loss ratios, where one can observe a large difference between the two extreme cases: UPP_A (worst-case for FGS enhancement-layer) and UPP_C (ideal-case for FGS enhancement-layer).

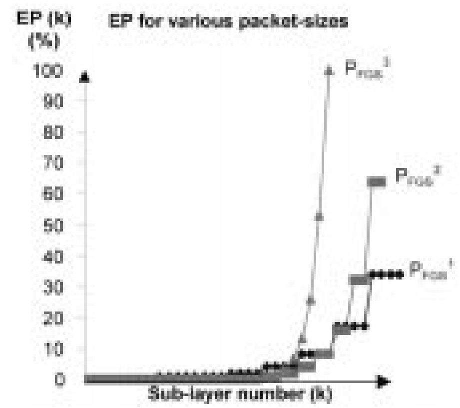

Fig. 7. Effective packet-loss ratios of each sublayer $\mathrm{E}(k)$ dependent on the packet-size.

Based on these results, an important conclusion of our study is that UPP between the base and enhancement-layer and also within the enhancement-layer represent an important mechanism for improving the resilience of the FGS video codec. This is further emphasized in Table I, which summarizes the results obtained with the various protection mechanisms. From Table I, it becomes clear that the proposed FGLP strategies are especially beneficial for moderate to high packet-losses, i.e., 5–10%.

To determine the extent to which the fine granularity in loss protection leads to an improved quality (see Section IV-B), simulations with various packet-sizes have been performed. A subset of the results is presented in Fig. 9, where the performance for the sequence *Foreman* coded with enhancement-layer frames of size $S_{\mathrm{FGS}} = 8004$ bytes is shown for various packet-sizes at various EP ratios. (The packet-size employed for the base-layer is not relevant in this context, since for UPP, the base-layer is entirely protected.)

From Fig. 9, two interesting observations can be made. At low packet-loss ratios (e.g., $\mathrm{EP} = 1\%$), a larger packet-size has the best performance and conversely, at higher packet-loss ratios, smaller packet-sizes give better performance. This trend could be expected based on the conclusion drawn in Section IV-B, since for low packet-losses, FGS benefits from the relatively reduced EP of the less significant (i.e., higher) layers. Conversely, at high packet-loss ratios, the chance of losing packets is high, and thus the better protection provided by a smaller packet-size to the first packets (see Fig. 7 and Section IV-B) is beneficial. However, due to the associated packetization overhead, the optimal performance is not always obtained for the smallest packet-size. Also, a very large packet-size could lead to unwanted packet fragmentation events by the network layer [11]. Therefore, as illustrated by Fig. 9, a packet-size of 492 bytes $(= P_{\mathrm{FGS}}^2)$, like that used for the simulations in Sections III and IV, is a reasonable size for FGS packetization.

At this point it is important to mention that the proposed FGLP strategy provides a good solution also from a resource allocation perspective. The complexity of the FGS decoder increases as a function of the bit-rate: as the number of decoded bit-planes increases, the computational complexity and

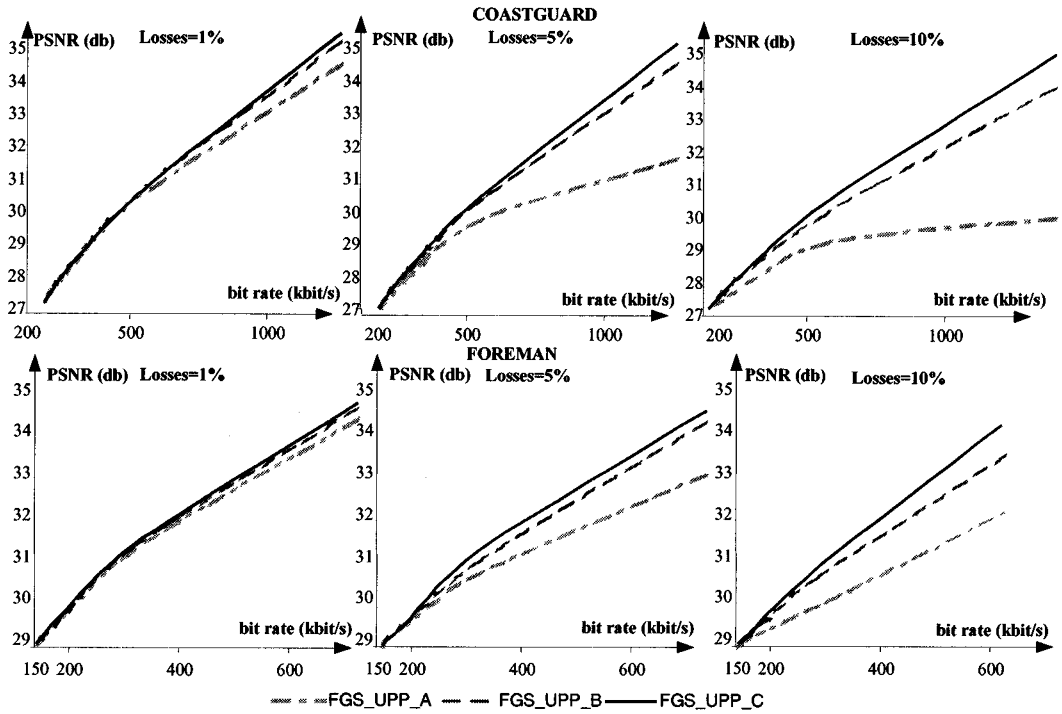

Fig. 8. Performance results of FGS packet-loss resilience with various UPP protection strategies.

TABLE I
PERFORMANCE OF THE VARIOUS PROTECTION STRATEGIES FOR "COASTGUARD"-SEQUENCE AT 1000 KBIT/S

Protection strategy	PSNR (dB) EP=1%	PSNR (dB) EP=5%	PSNR (dB) EP=10%
EPP	32.3	26.7	23.8
UPP_A	33.1	30.8	29.3
UPP_B	33.8	33.2	32.0
UPP_C	34.1	33.8	32.9

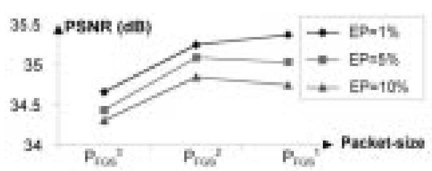

Fig. 9. Impact of the packet-size on the performance.

memory bandwidth access of the decoder increase more than linearly. This is because the less significant bit-planes have higher entropy and thus they need to be compressed using more bits. Hence, the FGS structure has also the inherent property of being complexity scalable with the number of decoded bit-planes. This complexity scalability when combined with the proposed FGLP protection mechanism provides a good resource allocation trade off—the most significant bit-planes are decoded more reliably and require less decoding complexity.

D. FGLP Protection Using FEC

In Section IV-C, the packet-loss resilience of FGS under the fine-grained protection framework has been theoretically derived. However, to provide various levels of protection for the various FGS sublayers, different techniques can be employed, such as retransmission, unequal FEC, distribution through channels with different packet-loss characteristics, or the packet replication technique proposed in [22]. If the proposed FGLP strategy is implemented using retransmission, only a negligible amount of overhead is necessary to ensure various levels of protection between the different layers and thus, to a large degree, the results portrayed in Fig. 8 still hold. Although retransmission based packet loss recovery can be easily employed for unicast applications, it may not be feasible to use retransmission for IP multicast applications (due to the potential implosion of NACK messages at the server). Therefore, it is important to evaluate the FGLP strategy under FEC packet loss protection. When the FGLP strategy is implemented using FEC, the transmission bandwidth is shared between the video data payload and the FEC, and thus the theoretical performance displayed in the plots of Fig. 8 can no longer be achieved.

Hence, in the sequel, we evaluate the performance of the FGLP implementation using FEC. In the Internet engineering task force (IETF), both generic [17] and RS [18] FEC have been proposed for enabling different levels of protection. In this paper, RS codes have been used for the realization of the FEC due to their flexibility. For our simulations, we used the RS design and corresponding two-state Markov channel model described in [7] and [26]. RS codes can be employed for protecting packets against loss if the codewords are formed across q information (i.e., containing video payload) and $m-q$ redundancy packets, as portrayed in Fig. 10. The resulting m packets are called *block of packets (BOP)*. It is important to note that the video payload packets can be entirely reconstructed from any subset of at least q correctly received packets using erasure decoding. Further, depending on the number of packets per BOP,

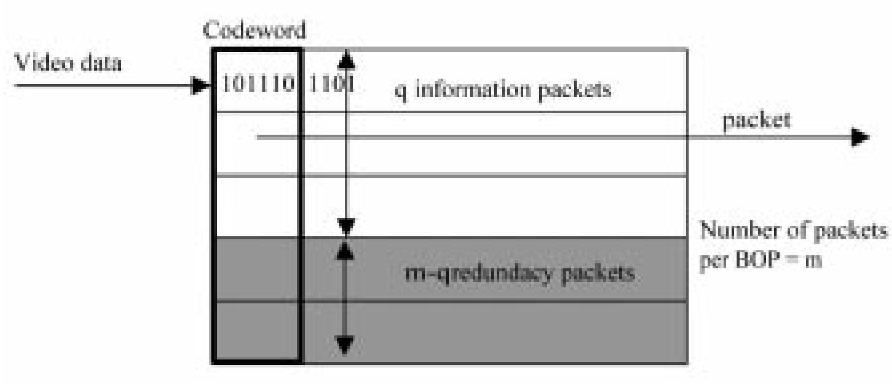

Fig. 10. FEC across packets according to [7], [26]. The video bitstream is written line-wise into q-packets. Additionally $m - q$ packets are generated by forming codewords column-wise. All m packets together form one BOP.

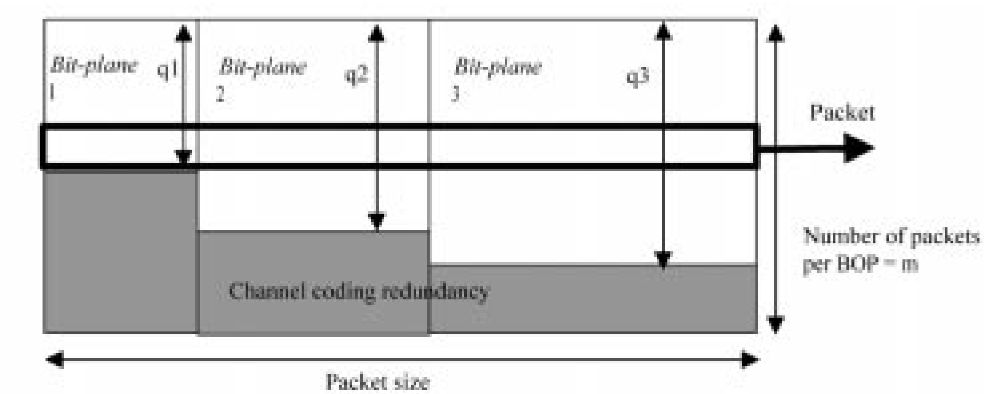

Fig. 11. Transmission of an FGS stream containing $n = 3$ bit-planes.

m, the packet-size and the transmission bit-rate and frame-rate, a different number of frames will be grouped together in a BOP.

The performance of an (m, q) RS code is based on the probability that the lost video packets cannot be reconstructed, which is given by the probability that more than $m - q$ packets are lost. This probability can be computed based on the probability $P(m, t)$ of t packets being lost within a block of m packets. (A derivation of the *block error mass function* $P(m, t)$ for a two-state Markov packet-loss model of the Internet can be found to [26].) For the packetization, we multiplex the FGS enhancement-layer bit-planes for a set of frames into a single BOP, as shown in Fig. 11. Then, each bit-plane k is filled into q_k packets, and $(m - q_k)$ packets are filled with channel redundancy. Therefore, q_k determines the protection level of bit-plane k, and $r_k = q_k/m$ is the code rate for this bit-plane. Hence, for a given BOP size m, the code rate r_k determines the level of protection needed for each sublayer k (i.e., in this case bitplane k). Below, we briefly describe how to determine the code rate r_k for a given bitplane k. This enables us to evaluate the performance of the different packet loss protection strategies. (For more details regarding the probabilistic model for packet losses, the reader is referred to [26].)

Let us now assume that the total number of bit-planes to be transmitted is BP. Then, each bit-plane k, with $k = 0, \ldots, \mathrm{BP} - 1$, will include one or more unrecoverable packets if the number of lost packets t in the BOP satisfies the following relationship: $t > m - q_k$. Hence, the probability P_k that bit-plane k will encounter one or more unrecoverable losses can be computed as $P_k = 1 - \sum_{t=0}^{m-qk} P(m, t)$. Although not explicitly shown, it is important to note that $P(m, t)$, and consequently P_k, is a function of the packet loss ratio PLR. (In [26] PLR is referred to as the average packet loss probability P_{B}.)

Assume now that we apply the FGLP protection strategy described in Section IV-A by assigning to each bit-plane a different protection. Hence, in the case of the UPP_B protection strategy, if we let each bit-plane k has ten times as good protection as the next less significant bit-plane $k + 1$, then $P_{k+1} = 10^k P_1$. Then, we need to determine for each bit-plane k, the relative code rate r_k that provides the desired FGLP protection strategy. Thus, for a particular enhancement-layer protection (represented by the parameter p) and a given average packet-loss probability P_{B}, we determine q_k such that $P_k = 1 - \sum_{t=0}^{m-qk} P(m, t; P_{\mathrm{B}}) = p^k P_1$, where P_1 is the probability of having one or more unrecoverable lost packets within the most significant bitplane (i.e., sublayer 1 in this case). Consequently, from q_k we can determine the code rate r_k.

We compare the various enhancement-layer protection strategies under the assumption that the base-layer is fully protected in all cases. Assume that the total enhancement-layer rate equals R_{enh} and the enhancement-layer contains BP bit-planes of rate R_{enh}^k, $0 \le k < \mathrm{BP}$. To enable the UPP_B protection strategy of Section IV-A, a different level of protec-

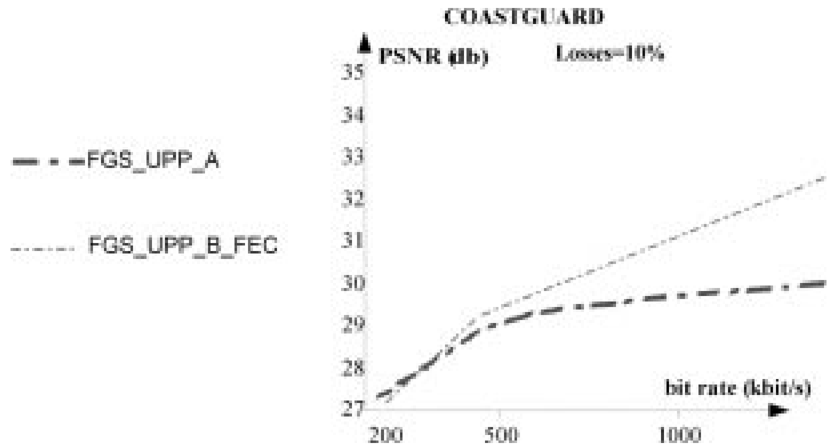

Fig. 12. Performance results of the UPP_B FGLP protection strategy using FEC for its implementation. In the UPP_A case, there is no protection for the enhancement-layer.

tion is assigned to each bit-plane k. The level of protection is determined by the amount of channel coding assigned to each bit-plane, given by the code rate r_k (see Fig. 11). Consequently, $R_{\rm enh} = \sum_{k=0}^{\rm BP-1} R_{\rm enh}^k = \sum_{k=0}^{\rm BP-1} (R_{\rm enhv}^k / r_k)$, where $R_{\rm enhv}^k$ is the video payload bitrate for each bit-plane and the code rates r_k are determined for the UPP_B protection strategy as previously explained. For the UPP_A case, where the entire FGS-layer is unprotected, no channel coding is added and thus $r_k = 1$, $0 \leq k < {\rm BP}$, and there is no overhead due to channel coding. To perform a fair comparison between UPP_A and UPP_B, the average packet-loss probability for UPP_B, i.e., $P_{\rm B}$, should be equal with the packet-loss probability of UPP_A.

The PSNR plots for the UPP_A and UPP_B FGLP cases are portrayed in Fig. 12. The plots of Fig. 12 have been determined for a BOP of $m = 40$, a probability of having one or more unrecoverable lost packets within the most significant bitplane $P_1 = 10^{-4}$ and an average packet-loss probability $P_{\rm B} = 10\%$. (These parameters have been chosen according to the typical packet-loss rates over the Internet.) Due to the random nature of Internet packet losses and their impact on compressed video, at each tested bit-rate 50 different runs of the experiments were conducted for each FGS stream. The results portrayed in Fig. 12 show that FEC is indeed a useful mechanism for achieving FGLP within the enhancement-layer. The FEC-based FGLP strategy is especially useful for increasing the robustness of FGS streams under high packet-losses. It is important to note that an improved PSNR performance can be obtained by determining the FEC codes that lead to the best overhead versus robustness trade off. However, the best trade off needs to be determined individually for a particular transmission bit-rate and average packet loss rate and is beyond the scope of this paper.

Further, it is important to mention that improved results can be obtained by grouping video payload packets from nonadjacent frames, i.e., interleaving, in a single BOP packet. In this way, the visual disturbance introduced by loosing a bit-plane for the entire set of frames within a BOP is limited. However, this interleaving procedure results in an additional delay and increased buffer requirements at the receiving end. Nevertheless, this forms an interesting topic for further research.

V. Conclusions

In this paper, the robustness of the recently adopted MPEG-4 FGS coding scheme under Internet packet losses has been evaluated. Extensive simulations[10] performed at a variety of bit-rates and packet-loss ratios demonstrated clearly the error resilience of FGS. Comparisons conducted with NS MPEG-4 streams under EPP revealed the FGS performance superiority under moderate-to-high packet-loss ratios. Moreover, under UPP, the FGS streams are less visually affected by packet-losses than the NS streams. This conclusion is of extreme importance because it shows that despite the fact that FGS has a lower performance than the NS MPEG-4 coder under ideal network conditions without losses, it has a much better resilience to Internet packet-losses especially under UPP.

To exploit the excellent performance of FGS under UPP as well as its fine-granularity, we introduced the FGLP strategy for the enhancement-layer. In addition, a generic analytical framework has been developed for evaluating FGLP bounds. The results indicate that FGLP for the FGS enhancement-layer can provide significant resilience under moderate-to-high packet-loss ratios (e.g., 5–10%). Also, it was established that for a given bit-rate and EP, an optimal packet-size can be determined for which the best FGLP performance can be obtained.

It is important to mention that the FGLP framework introduced in this paper can also be successfully employed for improving the Internet packet-loss resilience of alternative scalable coding schemes [13], [14]. The proposed FGLP scheme is especially useful for Internet multicast applications where different users are subject to different channel qualities and no feedback channel can be employed.

The combination of the novel FGLP protection strategy with FEC mechanisms that provide different levels of protection for the various layers has also been evaluated. The simulation results indicate a good performance for the FGLP strategy under high packet-losses.

[10]Our study involved more then 5000 simulation runs.

APPENDIX

There are many techniques for the design and implementation of packet-loss resilient FGS base-layer and single-layer NS decoders. For example, the simplest alternative is to freeze the last correctly decoded picture until a new intra-coded frame is received. One major disadvantage of this alternative is that it ignores all of the correct packets received after the lost packet, and therefore the decoder does not take advantage of a large amount of uncorrupted data. This could lead to freezing a picture for up to several seconds. Since the impact of packet-loss events on video quality is greatly influenced by the level of protection supported by the video stream and the corresponding decoder, in our approach we provided three levels of loss-protection and concealment for the FGS base-layer and NS streams. In addition to the sequence header (i.e., sequence-layer resilience), protection at the following levels is provided:

1) *GOP (Group Of Pictures)*—the propagation-effect of a packet-loss event is stopped by periodically intra-coded pictures in the stream.
2) *VOP (Video Object Plane)*—by allowing packets to contain data from only one VOP, the synchronization is always regained at the next VOP. (VOP is an MPEG-4 video picture.)
3) *Video-Packets* are the data between two resynchronization markers. Error concealment is supported using resynchronization markers at the beginning of every packet. Therefore, when a packet is lost, the following packets can be used to decode the remainder of the picture. The lost part of the picture is copied from the corresponding area of the previous frame. This very simple error concealment technique was determined to have the best tradeoff between enhanced visual quality and decoder complexity [15]. For FGS, if a packet within an enhancement-layer frame is lost, the remainder packets associated with that frame are also discarded.

In [26], the same protection levels were used for comparing the performance of the NS H.263 and the spatio-temporal scalable coder. However, it is important to note that H.263 adopts a resynchronization marker after a fixed number of blocks, referred to as groups of blocks (GOBs). This procedure leads to a variable length per GOB, depending on the activity within the GOB, bit-rates etc., and therefore, the resynchronization markers are unevenly spaced throughout the bitstream. On the contrary, the MPEG-4 resynchronization marker is inserted at almost fixed positions in the bitstream. The advantage of this approach is that certain portions of the scene, like high motion areas, which are more susceptible to errors, are easier to conceal with the MPEG-4 approach.

It is also important to note that other MPEG-4 error resilience tools like data partitioning, reversible VLCs, and adaptive intra-refresh methods are not very useful in this study due to the typical size of an IP packet, which can carry a large portion of a VOP. Therefore, these resilience tools were not employed in our study. However, if a feedback channel is available, more sophisticated error resilient coding methods can be used [21].

ACKNOWLEDGMENT

The authors would like to thank C. Dufour from Philips LEP for providing the error concealment mechanism for the NS codec and numerous useful suggestions. Also, they would like to thank J. Lan, L. Boland, and L. Boroczky from Philips Research for their useful comments on a previous version of this manuscript. Last but not least, their sincere thanks to the three anonymous reviewers who provided excellent and thorough feedback that improved the quality of this paper.

REFERENCES

[1] H. Radha and Y. Chen, "Fine-granular-scalable video for packet networks," in Packet Video Workshop, New York, Apr. 1999.

[2] Y. Chen, C. Dufour, H. Radha, R. Cohen, and M. Buteau, "Request for fine granular video scalability for media streaming applications," in Contribution to 44th MPEG Meeting, Dublin, Ireland, July 1998.

[3] Text of ISO/IEC 14 496-2 MPEG-4 Video FGS ver. 4.0, Noordwijkerhout, The Netherlands, 2000.

[4] H. Radha, Y. Chen, K. Parthasaranthy, and R. Cohen, "Scalable internet video using MPEG-4," Signal Process. Image Commun., Sept. 1999.

[5] V. Paxson, "End-to-End internet packet dynamics," in *Proc. ACM SIG-COM*, vol. 27, France, Oct. 1997, pp. 13–52.

[6] M. van der Schaar, H. Radha, and C. Dufour, "Scalable MPEG-4 video coding with graceful packet-loss resilience over bandwidth-varying networks," in Proc. ICME, New York, July 2000.

[7] B. Girod, K. W. Stuhlmüller, M. Link, and U. Horn, "Packet loss resilient internet video streaming," in *Proc. VCIP*, vol. 3653, Proc. SPIE, Jan. 1999, pp. 833–844.

[8] W. Li, "Experiment result on fine granular scalability," in Contribution to 46th MPEG Meeting, Seoul, Korea, March 1999.

[9] G. Carle and E. W. Biersack, "Survey of error recovery techniques for IP-based audio-visual multicast applications," *IEEE Network*, vol. 11, pp. 24–36, Nov./Dec. 1997.

[10] M. van der Schaar, Y. Chen, and H. Radha, "Embedded DCT and wavelet methods for fine granular scalable video: Analysis and comparison," in *Proc. IVCP*, vol. 2974, Proc. SPIE, Jan. 2000, pp. 643–653.

[11] C. A. Kent and J. C. Mogul, "Fragmentation considered harmful,", Internal Compaq Rep. WRL-TR-87.3, 1987.

[12] "Rationale for proposed draft amendment (PDAM4),", Noordwijkerhout, The Netherlands, w3346, 2000.

[13] D. Taubman and A. Zakhor, "Multirate 3-D subband coding of video," *IEEE Trans. Image Processing*, vol. 3, pp. 572–588, Sept. 1994.

[14] B.-J. Kim and W. A. Pearlman, "An embedded wavelet video coder using three-dimensional set partitioning in hierarchical trees (SPIHT)," in Proc. IEEE Data Compression Conf., Mar. 1997.

[15] J. Lu, "Signal processing for internet video streaming: A review," in *Proc. IVCP*, vol. 2974, Proc. SPIE, Jan. 2000, pp. 246–259.

[16] A. Lippman, "Video coding for multiple target audiences," in *Proc. VCIP*, vol. 3653, Proc. SPIE, Jan. 1999, pp. 780–782.

[17] J. Rosenberg and H. Schulzrinne, "An RTP payload format for generic forward error correction,", http://www.ietf.org,, 1998.

[18] ——, "An RTP payload format for Reed-Solomon codes,", http://www.ietf.org,, 1998.

[19] K. W. Stuhlmüller, M. Link, B. Girod, and U. Horn, "Scalable internet video streaming with unequal error protection," in Packet Video Workshop, New York, Apr. 1999.

[20] W. Li, "Bit-plane coding of DCT coefficients for fine granularity scalability," in Contribution to 45th MPEG Meeting, Atlantic City, NJ, Oct. 1998, m3989.

[21] E. Steinbach, N. Färber, and B. Girod, "Standard compatible extension of H.263 for robust video transmission in mobile environments," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, pp. 872–881, Dec. 1997.

[22] W. Tan and A. Zakhor, "Real-time internet video using error resilient scalable compression and TCP-friendly transport protocol," *IEEE Trans. Multimedia*, vol. 1, pp. 172–186, June 1999.

[23] R. Aravind, M. R. Civanlar, and A. R. Reibman, "Packet loss resilience of MPEG-2 scalable video coding algorithms," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 6, pp. 426–435, Oct. 1996.

[24] M. van der Schaar, H. Radha, and Y. Chen, "An all FGS solution for hybrid temporal-SNR scalability," in Contribution to 50th MPEG Meeting, Maui, HI, Dec. 1999, m5552.

[25] M. van der Schaar and H. Radha, "A novel MPEG-4 based hybrid temporal-SNR scalability for internet video," in Proc. ICIP, Vancouver, Canada, Sept. 2000.

[26] U. Horn, K. W. Stuhlmüller, M. Link, and B. Girod, "Robust internet video transmission based on scalable coding and unequal error protection," *Signal Process. Image Commun.*, Sept. 1999.

Mihaela van der Schaar received the M.Sc. degree in electrical engineering in April 1996 from Eindhoven University of Technology, Eindhoven, The Netherlands, where she is currently pursuing the Ph.D. degree.

In 1996, she joined Philips Research Laboratories Eindhoven, where she became a Member of the TV Systems Department. From 1996 to 1998, she was involved in several projects which investigated low-cost very high quality video compression techniques and their implementation for TV, computer, and camera systems. Since 1998, she has been an expatriate at Philips Research, Briarcliff Manor, NY, where she is a Senior Member of Research Staff in the Video Communications Department. She is currently involved in the research of video coding techniques for Internet and wireless video streaming and leads a team of researchers working on scalable video coding, networking, and streaming algorithms. Since 1999, she has been an active participant in the MPEG-4 video standard, contributing to the "Fine-Granularity Scalability"tool. Her research interests include video and graphics coding and video streaming over unreliable channels. She has coauthored more than 30 conference and journal papers in this field and holds several patents.

Hayder Radha (S'85–M'86) received the B.S. degree (with honors) in electrical engineering from Michigan State University, Ann Arbor, in 1984, the M.S. degree in electrical engineering from Purdue University, West Lafayette, IN, in 1986, and the Ph.M. and Ph.D. degrees from the Center for Telecommunications Research (CTR), Columbia University, New York, NY, in 1991 and 1993, respectively.

He joined the faculty at Michigan State University in 2000 as an Associate Professor in the Department of Electrical and Computer Engineering. He is also a Philips Research Fellow. In 1996, he joined Philips Research, Briarcliff Manor, NY, where he worked in the area of video communications and high-definition television. He initiated the Internet Video research program at Philips Research and led a team of researchers working on scalable video coding, networking, and streaming algorithms. Prior to joining Philips, he was a Distinguished Member of Technical Staff at Bell Laboratories. He started his career at Bell Laboratories in 1986, where he worked in the areas of digital communications, signal and image processing, and broad-band multimedia communications. He served as a Co-chair and Editor of the ATM and LAN Video Coding Experts Group of the ITU-T between 1994 and 1996. He is also an Adjunct Professor at the City University of New York. His research interests include image and video coding, multimedia communications and networking, and the transmission of multimedia data over wireless and packet networks. He holds more than 20 patents in these areas (between granted and pending).