### Title

Unequal error protection based on slice visibility for transmission of compressed video over OFDM channels

### Authors

Toni, L
Cosman, P C
Milstein, L B

### Publication Date

2011-07-01

# UNEQUAL ERROR PROTECTION BASED ON SLICE VISIBILITY FOR TRANSMISSION OF COMPRESSED VIDEO OVER OFDM CHANNELS

*Laura Toni*$^*$*, Pamela C. Cosman*$^\dagger$*, and Laurence B. Milstein*$^\dagger$
$^*$ TERA Dept., Italian Institute of Technology (IIT), Genova, Italy
$^\dagger$ ECE Dept., University of California at San Diego, La Jolla, CA
e-mail: `laura.toni@iit.it, pcosman@code.ucsd.edu, milstein@ece.ucsd.edu`.

## ABSTRACT

We address channel code rate optimization for transmission of non-scalable coded video sequences over orthogonal frequency division multiplexing networks. A slice loss visibility (SLV) model is used to evaluate the visual importance of each H.264 slice. Based on both the SLV model and the frequency diversity order available from the channel, we propose a cross-layer technique to allocate video slices within a 2-D time-frequency resource block, and optimize the unequal channel code rate profile, in order to better protect more visually important slices. The proposed algorithm outperforms baseline ones which do not take into account the SLV.

***Index Terms*****—** Slice loss visibility, channel coding, cross-layer design, diversity, multimedia communications, orthogonal frequency division multiplexing (OFDM).

## 1. INTRODUCTION

Since video packet losses have different impacts on the video quality, cross layer techniques, which take into account both physical and application layer conditions, might minimize the video distortion caused by channel impairments. In particular, the physical layer parameters (e.g., channel coding rate, modulation) can be tuned based on both channel conditions and the information about the bitstream to transmit.

When fine-grain scalable video sequences are considered, each bit of the encoded enhancement bitstream within a frame is more important than the subsequent bit. By adopting unequal error protection (UEP), a more reliable transmission is offered to the more important bits, and the distortion of the received information might be reduced compared with an equal error protection (EEP) system [1]. For non-scalable video sequences, assigning priority levels to portions of the compressed bitstream is more challenging. In [2, 3], the authors proposed a bitstream-based metric for slice loss visibility (SLV) for non-scalable compressed video. Bitstream-based metrics predict video quality using packet header information and limited information from the encoded bitstream such as motion vectors. The authors conducted subjective tests in which the viewers' task is to indicate when they observe a packet loss artifact. From these tests, a SLV metric was proposed with the goal of predicting whether an individual packet loss in the video stream is visible to a viewer. In [4], the SLV model was used to optimize the channel code rate for video transmission over additive white Gaussian noise (AWGN) single carrier channels.

In this paper, we aim to minimize the distortion of non-scalable bitstreams transmitted over doubly selective orthogonal frequency division multiplexing (OFDM) systems. Based on the SLV, we propose a technique that jointly groups the encoded bitstream into packets and optimizes the channel code rate for each packet. The remainder of this paper is organized as follows. In Section 2, we describe basics of SLV and OFDM systems. Section 3 discusses the proposed cross-layer diversity approach and the baseline algorithms. In Section 4, we provide simulation results and discussion, and conclude in Section 5.

## 2. PRELIMINARIES

In the following, we provide a brief introduction to the SLV metric and a description of the system model.

### 2.1. Slice Loss Visibility Overview

We consider a non-scalable video encoder (e.g., H.264, MPEG-4) and assume that each frame is divided into $N_{\text{s}}$ slices (each slice consists of a constant number of macroblocks). The $i$th slice of frame $k$ is encoded into $L_k(i)$ bits and has a priority level $V_k(i)$. The $V_k(i)$ values range from $0$ to $1$, and can be interpreted as the probability that the slice, if lost, would produce an artifact detected by the end user. So $V_k(i) = 0$ means that the slice, if lost, would likely not be noticed by any observer, whereas $V_k(i) = 1$ means that the loss artifact would likely be seen by all users. The priority level is determined by the SLV model which estimates the quality degradation the video experiences when that slice is lost [2]. So, each encoded slice is characterized by the pair $(V_k(i), L_k(i))$, for $i = 1, \dots, N_{\text{s}}$ and $k = 1, \dots, N_{\text{F}}$, where $N_{\text{F}}$ is the number of frames per group of pictures (GOP).

---

This research was supported in part by Intel Corp. by Cisco Systems, Inc., and by the UCSD FWGrid Project, NSF Research Infrastructure Grant Number EIA-0303622.

978-1-61284-350-6/11/$26.00

## 2.2. System Model

The video sequences are transmitted over frequency-selective OFDM networks and we use a block fading channel model to simulate the frequency selectivity [5]. In this model, the spectrum is divided into blocks of size $(\Delta f)_{\mathrm{c}}$. Subcarriers in different blocks are considered to fade independently; subcarriers in the same block experience identical fades. As illustrated in Fig. 1, we assume an OFDM system with an overall system bandwidth $W_{\mathrm{T}}$, such that we can define $N$ independent subbands. Each subband consists of $M$ correlated subcarriers spanning a total bandwidth of $(\Delta f)_{\mathrm{c}}$. The total number of subcarriers in the OFDM system is $NM$. Generally, the maximum achievable frequency diversity $\mathcal{D}_{\mathrm{f}}$ is given by the ratio between the overall system bandwidth $W_{\mathrm{T}}$ and the coherence bandwidth $(\Delta f)_{\mathrm{c}}$.

In the time domain, the channel experiences Rayleigh fading. We use the modified Jakes' model [6] to simulate different fading rates, resulting in different time diversity orders. The maximum time-diversity gain $\mathcal{D}_t$ is given by the ratio between the duration of a packet and the channel coherence time $(\Delta t)_{\mathrm{c}}$. For possible diversity and coding gains in the time domain, a concatenation of cyclic redundancy check (CRC) codes and rate-compatible punctured convolutional (RCPC) codes are applied to each transmitted packet. Depending on the algorithm for allocating slices to packets, the packets might end up all having roughly the same importance (in which case EEP would be appropriate for the packets) or they might end up having very different priority levels (in which case UEP would make sense, in order to better protect the more important packets). In the following, we propose an algorithm for the transmission of non-scalable video sequences over slow fading channels. We assume a broadcasting scenario, in which instantaneous channel state information (CSI) is not available at the transmitter. Thus, the allocation of the source bitstream in the resource block (RB) and the channel code rate profile are optimized taking into account the mean SNR values, the frequency and time diversity orders, and the SLV parameters.

## 3. VISIBILITY-BASED ALGORITHM

The algorithm steps are applied to each GOP. Since the number of bits in which a single frame is encoded might be considerably different (e.g., the number of bits for an I-frame will be greater than the number required for a B-frame), assuming a constant RB for each frame would not give good quality. Instead, we adopt a fixed-sized 2D time-frequency RB for each GOP. This cross-layer choice corresponds to a very common approach in application-layer video rate control, in which the number of bits allocated to individual frames is allowed to vary, but the number of bits given to each GOP is held roughly constant.

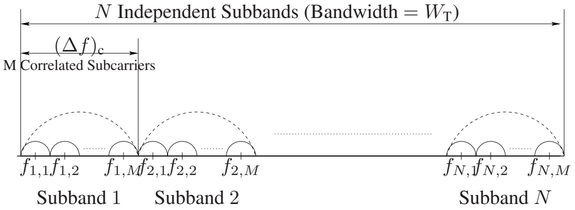

**Fig. 1**. Subcarrier spectrum assignment.

As illustrated in Fig. 2, the current GOP is processed by a joint allocation/coding step. In Fig. 2, $N_{\mathrm{F}}$ frames form a GOP; each frame is divided into $N_{\mathrm{s}}$ slices. After the allocation/channel-coding algorithm, groups of slices are allocated to each packet. Then, each packet will consist of some slices plus the forward error correction (FEC) added by the RCPC code. It should be noted that information bitstream and RCPC parity symbols would be interleaved in an actual system. However, for illustration, we show the de-interleaved version to depict the relative amounts of RCPC parity symbols and information symbols. After channel coding, packets have constant length (equal to $L_{\mathrm{p}}$ modulated symbols) and will be assigned to a subcarrier. Then, for each RB, $N_{\mathrm{t}}$ packets will be transmitted on $N_{\mathrm{t}}$ subcarriers.

Note that slice allocation into the RB and channel code rate optimization are mutually dependent processes. The best EEP or UEP profile for the packets depends on the mean SLV parameter for the slices within each packet. But the mean SLV for a packet depends on how many information bits get allocated to the packet, thus it depends on the RCPC code rate adopted for the packet. This joint allocation/coding step is the focus of our work. We propose an algorithm to allocate the slices of each GOP and evaluate the optimal RCPC code rate by taking into account both the SLV and the channel model parameters.

The proposed method can be described with the following steps, depicted in Fig. 3.

***Step 1.*** We subdivide all the $N_{\mathrm{F}} \times N_{\mathrm{s}}$ slices of the GOP into $K_v$ groups based on the SLV parameter. The first group $(\Lambda_1)$ contains the most visible slices (i.e., the slices which, if lost, are most likely to produce a visible glitch) and the last one $(\Lambda_{K_v})$ the least visible. The $j$th group $\Lambda_j$ is defined as

$$\begin{aligned} \Lambda_j : &\left\{ \text{Slice } V_k(i) \text{ s.t. } V_k(i) \in \left[ V_j^{\star}, V_{j+1}^{\star} \right] \right\}, \qquad (1) \\ &\text{with } j = 1, \ldots, K_v, k = 1, \ldots, N_{\mathrm{F}}, i = 1, \ldots, N_{\mathrm{s}} \end{aligned}$$

where $\{V_j^{\star}\}$ are fixed thresholds such that $V_{j+1}^{\star} > V_j^{\star}$, with $V_1^{\star} = 0$ and $V_{K_v+1}^{\star} = 1$. We consider equally spaced thresholds in the range $[0, 1]$, therefore $V_{j+1}^{\star} = V_j^{\star} + 1/K_v$.

***Step 2***. Since the slices within each group have roughly the same visual importance, they should have the same protection. So we assign a RCPC code rate for each group. We are looking for the rate vector $\mathbf{r} = [r_1, r_2, \ldots, r_{K_v}]$ where $r_j$ denotes the RCPC code rate assigned to the slices within group $\Lambda_j$. That is, all slices in the $j$th visibility group

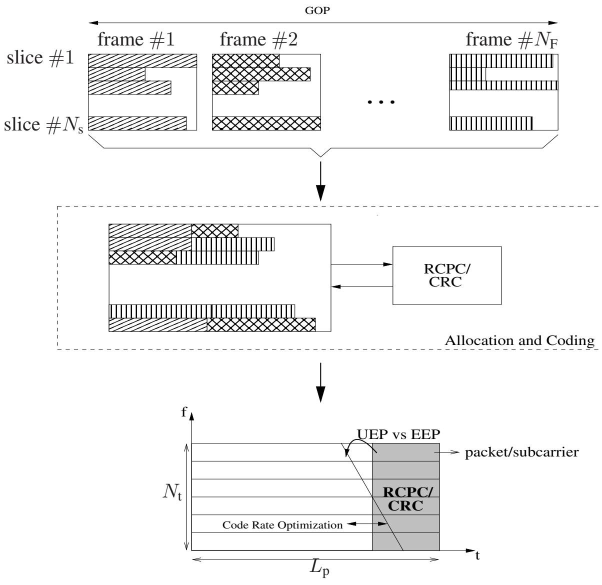

**Fig. 2**. Transmission of the GOP over OFDM mobile wireless networks. Note that the CRC/RCPC parity symbols are interleaved with the information symbols in the actual system.

should be allocated to packets encoded with a code rate $r_j$. We will use the capital letter $R_i$ to denote the RCPC code rate for the $i$th subcarrier or packet. As depicted in Fig. 3, if the $i$th packet contains slices from group $\Lambda_j$, then $R_i = r_j$.

For each rate vector $\mathbf{r}$ to be evaluated, the following step is considered. Based on both the frequency diversity order and the $\mathbf{r}$, the slices of each group will be allocated into subchannels. Assuming that the group $\Lambda_1$ needs to be allocated in the RB, the first $m$ subcarriers will be occupied by the group $\Lambda_1$, and each one of these $m$ packets will be protected with a RCPC code rate $R_i = r_1$ for $i = 1, \ldots, m$. The number of subcarriers in which group $\Lambda_1$ is allocated has to meet the following constraint

$$\sum_{i,k:V_k(i)\in\Lambda_1} L_k(i) \leq m \times L_\mathrm{p} \times r_1$$

where $L_\mathrm{p} \times r_1$ is the number of information bits per subcarrier. Since we do not assume knowledge of the instantaneous CSI, we consider the first $m$ subcarriers for the allocation of the first group, but we could select the subcarriers randomly. This step is considered for all $K_v$ groups. If the number of bits in the GOP is greater than the number of information bits available in the RB, randomly chosen slices from the least important group are dropped.

If the final goal of the proposed method is to choose the RCPC channel code rate profile able to maximize the mean quality of the whole video sequence, once steps $1-2$ are computed for all the GOPs of the sequence, the sequence quality

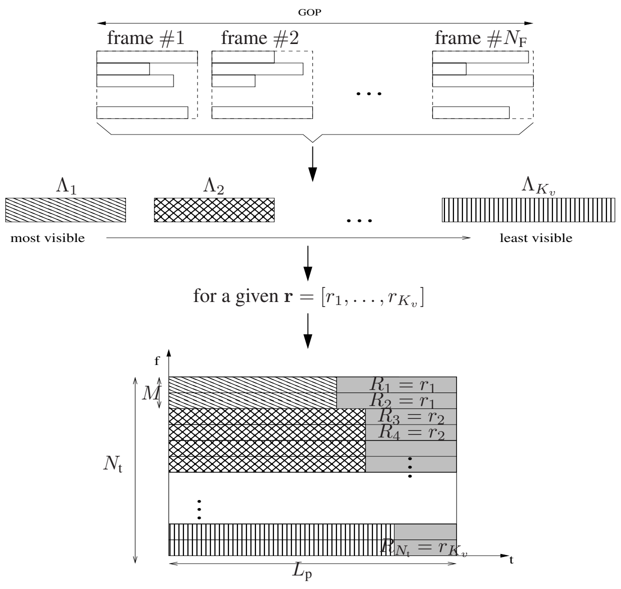

**Fig. 3**. Steps of the visibility-based algorithm.

is evaluated for all possible rate vectors $\mathbf{r}$ and the best $\mathbf{r}$ is selected. Alternatively, if the goal is to choose the best FEC GOP by GOP, then steps $1-2$ are computed for all rate vectors $\mathbf{r}$ and the best RCPC profile is evaluated for each GOP of the video sequence.

### 3.1. Baseline Algorithms

For comparison, we consider two baseline algorithms: A) Sequential, B) Random. In both of these, we assume that slice importance is not known, and so no packet is more important than another. Thus, EEP is considered for the RCPC coding.

The *Sequential* algorithm simply allocates sequentially the slices of each frame into the RB. This means that the first slices of the first frame of the considered GOP will be allocated to the first subcarrier. When no more information bits are available in the first subcarrier, the algorithm starts allocating the current frame to the next subcarriers. Once the slices of the first frame of the GOP are allocated, the second frame is considered. The *Random* algorithm allocates each slice of the GOP in a random position of the RB.

## 4. RESULTS

We carried out simulations on four videos of 10s duration, coded at $R = 600$ kbps using the H.264/AVC JM codec with SIF resolution $(352 \times 240)$, and with Motion-Compensated Error Concealment (MCEC) as used in [7], implemented in the decoder. For brevity, we provide results for three test sequences: “LowMot”, “MedMot” and “HighMot”. “LowMot” is an almost static video, “MedMot” is a mid-level mo-

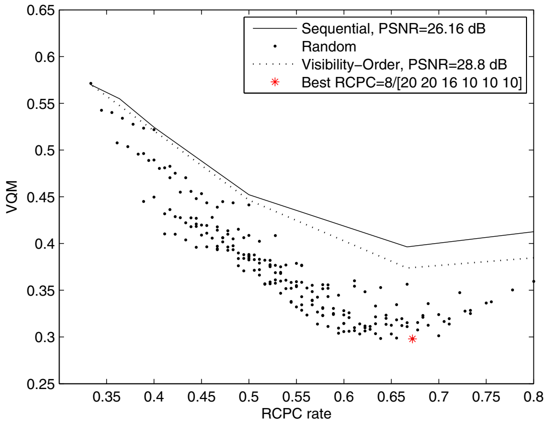

**Fig. 4**. VQM vs. $R_{rcpc}$ for both visibility-based and baseline algorithms optimized for the whole sequence, for systems with SNR $=$ 16 dB, $(N, M) = (32, 4)$, $f_{nd} = 10^{-4}$. "MedMot" video is considered.

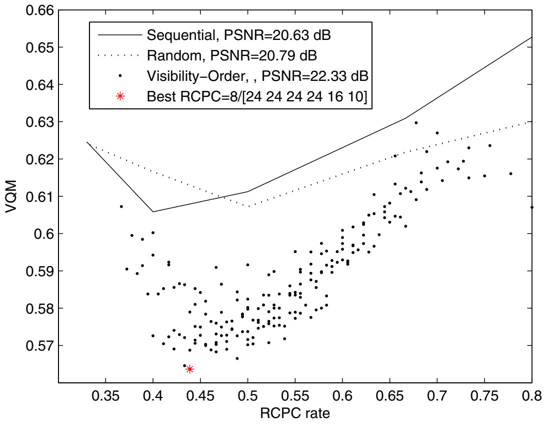

**Fig. 5**. VQM vs. $R_{rcpc}$ for both visibility-based and baseline algorithms optimized for the whole sequence, for systems with SNR $=$ 8 dB, $(N, M) = (32, 4)$, $f_{nd} = 10^{-4}$. "MedMot" video is considered.

tion sequence, while "HighMot" has high motion and several scene changes. We used the IBBP encoding structure with I-frames every 24 frames. There are $N_\mathrm{t} = 128$ OFDM subcarriers in total. The RCPC codes of rates $R_{rcpc} = \left\{\frac{8}{10}, \frac{8}{12}, \frac{8}{16}, \frac{8}{20}, \frac{8}{24}\right\}$, were obtained by puncturing an $R_c = 1/3$ mother code with $K = 7$, $p = 8$ and generator polynomials $(133, 165, 171)_{octal}$ with the puncturing table given in [8]. QPSK modulation is considered. The packet size after the RCPC/CRC coding was set equal to $L_\mathrm{p} = 588$ bytes, such that $L_\mathrm{p} \times 8 \times T_\mathrm{b} \approx 24/30s$ (to respect the constraint of 30 fps), where $T_\mathrm{b}$ is the bit duration time. Results are provided in terms of the Video Quality Metric (VQM) score [9], a full-reference (FR) metric that has much higher correlation with human perception than PSNR or other simple FR video quality metrics [9]. Moreover, the PSNR achieved with the best RCPC scheme (the one that produces the lowest VQM value) is provided for each considered scenario.

We initially compare the visibility-based and the baseline algorithms for whole sequence optimization. Then, we will provide results for the GOP by GOP optimization case. For the visibility-based model, we used 6 visibility groups for the slices (i.e., $K_v = 6$) and considered all possible combinations of RCPC code rates for the 6 groups. In the plots which have RCPC code rate on the x-axis, the plotted value represents the EEP code rate for the random and sequential methods, whereas for the visibility-based method we plot a point at the average rate (that is, it is the ratio of information bits to total bits for the whole sequence). Since a slow fading scenario is considered $\left(\text{i.e., } f_{nd} = 10^{-4}\right)$, the maximum order of diversity in the time domain is $\mathcal{D}_\mathrm{t} = 1$. So, no diversity can be exploited by using RCPC codes, although coding gain can still be obtained. Fig. 4 depicts the VQM vs. the mean RCPC rate when "MedMot" is transmitted over a system with SNR $=$ 16 dB, $f_{nd} = 10^{-4}$, and $(N, M) = (32, 4)$. The diversity order experienced by the system in the frequency domain is $\mathcal{D}_\mathrm{f} = 4$. It is worth noting that the orders of diversity available from the channel, together with the mean SNR, are considered for selecting the packet loss rate able to suitably characterize the channel during the optimization process. We observe that the best RCPC combination of the visibility-based algorithm (the best is the one that produces the lowest VQM value) is better (lower) than the best VQM provided by the Sequential or Random methods. This means that there is a UEP level able to outperform the baseline algorithms. In the literature, a VQM gain of $0.1$ is considered to be a good improvement, and the gain in Fig. 4 is about $0.1$. For the visibility-based method, the best UEP rate vector is $\mathbf{r} = [8/24 \ 8/16 \ 8/10 \ 8/10 \ 8/10 \ 8/10]$. Note that the mean RCPC channel code rate is around $0.67$, which is almost the same as the best RCPC channel code rate of the baseline algorithms. This means that, for the three methods, the amount of FEC inserted within the RB is almost the same, but in the visibility-based method there is a more appropriate use of the redundancy bits.

We now examine a different channel SNR. The same behavior can be observed in Fig. 5, where VQM score vs. mean RCPC code rate is considered for a system with SNR $=$ 8 dB, $f_{nd} = 10^{-4}$, and $(N, M) = (32, 4)$. Compared to the system in Fig. 4, the orders of diversity are the same, while the mean SNR is reduced. This reduction of reliability leads to an increase in the FEC level of the best RCPC code rate for the visibility-based method. In particular, the most visu-

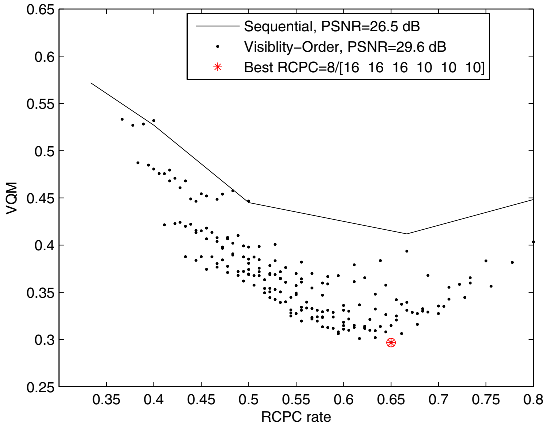

**Fig. 6**. VQM vs. $R_{rcpc}$ for both visibility-based and sequential algorithms optimized for the whole sequence, for systems with SNR $= 16$ dB, $(N, M) = (128, 1)$, $f_{nd} = 10^{-4}$. “MedMot” video is considered.

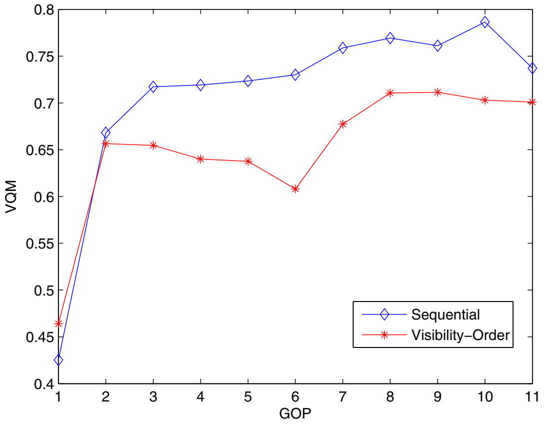

**Fig. 7**. Best VQM for each GOP of the “MedMot” sequence for visibility-based and sequential algorithms optimized for the whole sequence, for systems with $(N, M) = (32, 4)$, $f_{nd} = 10^{-4}$, and SNR $= 8$ dB.

ally important groups $\Lambda_i$ are more protected than they are in the 16 dB case. This increasing FEC level in the RB keeps the slice loss rate due to channel losses roughly the same as it was for the system with mean SNR $= 16$ dB, at the expense of an increase in the number of low-priority slices being discarded prior to transmission. It can be observed that, for such extreme conditions, the gain of the visibility-based algorithm over the baseline ones is limited when the UEP profile is optimized for the whole sequence. The performance gain in Fig. 5 is only about $0.04$ in VQM score.

In Fig. 6, a different order of diversity in the frequency domain is considered. VQM for “MedMot” vs. mean channel code rate is provided for systems with $(N, M) = (1, 128)$, SNR $= 16$ dB, and $f_{nd} = 10^{-4}$. Here, a maximum diversity order of 128 is experienced, meaning that the transmitted packets experience independent fades. In general, as the frequency diversity order increases, the variation in the number of lost packets decreases and thus reduces the need for FEC. For the visibility-based algorithm, the best RCPC rate vector $\mathbf{r} = [8/16\ 8/16\ 8/16\ 8/10\ 8/10\ 8/10]$ outperforms the baseline algorithms, achieving an improvement greater than $0.1$ in VQM score.

Rather than providing results of the whole sequence optimization in terms of mean VQM, in Fig. 7, the VQM score for each GOP is provided for the “MedMot” sequence for $(N, M) = (32, 4)$, $f_{nd} = 10^{-4}$, and SNR $= 8$ dB. Even in these poor channel conditions, where the average improvement for the whole sequence is about $0.04$ in VQM score, the gain of the visibility-based algorithm over the sequential one, for some individual GOPs, is significant (i.e., the gain is up to $1.2$ in VQM score).

We now provide results when the RCPC profile is optimized GOP by GOP. Fig. 8 compares the best VQM for each GOP of “MedMot” achieved from the sequential and the visibility-based algorithm for both SNR $= 8$ and 16 dB. Here the improvement for some GOPs is very much larger than those in the earlier figures. As expected, for both the algorithms, the case of SNR $= 8$ dB achieves a VQM value higher (worse) than for 16 dB, due to the higher packet loss rate. Most importantly, for all the GOPs of the video sequences, the proposed algorithm achieves VQM values lower (better) than the one provided by the sequential algorithm. For example, for the 9th GOP transmitted over a 16 dB channel, the visibility-based algorithm achieves a VQM score equal to $0.3$, despite the VQM of $0.49$ of the sequential method. As already observed by comparing Fig. 4 and Fig. 5, the gain of the visibility-based algorithm for systems with mean SNR $= 16$ dB is greater than the one achieved in scenarios with mean SNR $= 8$ dB. This holds true for the case where the optimization is done GOP by GOP as well. However, the gains for both channel conditions are more substantial in this case. When SNR $= 8$ dB, and the optimization is done GOP by GOP, the visibility-based algorithm outperforms the baseline ones by more than $0.1$ in VQM score, although the gain was only $0.04$ in VQM score when the RCPC profile was optimized for the whole sequence. Similar considerations can be deduced from Fig. 9, where the best VQM is provided for each GOP of the “HighMot” and “LowMot” sequences for systems with SNR $= 16$ dB, $(N, M) = (32, 4)$, and $f_{nd} = 10^{-4}$. For both the sequences, the visibility-based algorithm achieves VQM lower (better) than the sequential method. However, for the quasi-static sequence “LowMot”, which has few visually im-

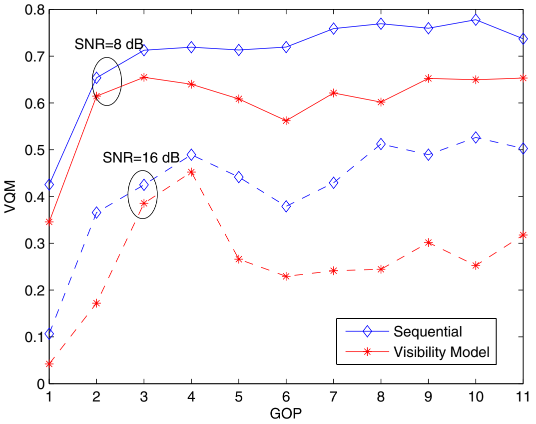

**Fig. 8**. Best VQM for each GOP of the "MedMot" sequence for visibility-based and sequential algorithms optimized GOP by GOP, for systems with $(N, M) = (32, 4)$, $f_{nd} = 10^{-4}$, and SNR $= 8$ and $16$ dB.

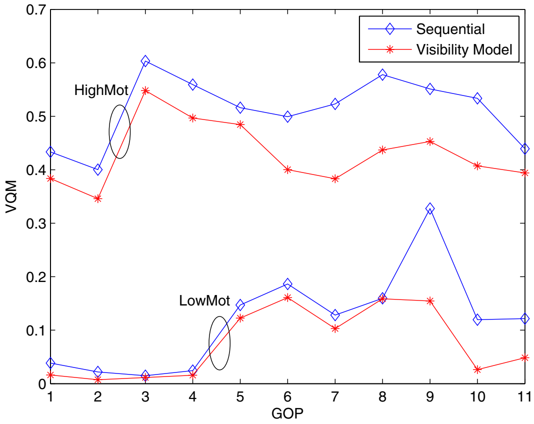

**Fig. 9**. Best VQM for each GOP of the "HighMot" and "LowMot" sequences for visibility-based and sequential algorithms optimized GOP by GOP, for systems with $(N, M) = (32, 4)$, $f_{nd} = 10^{-4}$, and SNR $= 16$ dB.

portant slices, the gain is negligible for most of the GOPs.

## 5. CONCLUSIONS

We studied channel coding in a 2-D time-frequency resource block of an OFDM system for transmission of non-scalable compressed video. We used a network-based slice loss visibility (SLV) model to estimate the visual importance of individual video slices. We proposed a cross-layer algorithm to map slices into time-frequency resource blocks. By taking into account the SLV model and the diversity orders offered by the channel in the frequency domain, the proposed technique provides protection tailored to each video slice. The proposed method significantly outperforms the baseline ones considered in this paper. In poor channel conditions, due to the high packet loss rates and/or the large number of slices that need to be discarded in order to fit the bitstream within the resource block, the gain in terms of mean VQM of the proposed algorithm is almost negligible, but it increases with the improvement in channel conditions. Moreover, for some GOPs of the sequence, the visibility-based algorithm offers a performance improvement greater than $0.1$. The proposed technique is especially useful for video sequences with medium to high motion. Lastly, using GOP by GOP optimization rather than whole sequence optimization, one increases the algorithm complexity and substantially decreases the latency, and the performance gain of the proposed algorithm over the sequential baseline increases. In particular, a noticeable gain is experienced even for poor channel conditions.

## 6. REFERENCES

[1] F. Masoud S. Cho and W. A. Pearlman, "Robust image transmission using a new joint source channel coding algorithm and dual adaptive OFDM," in *Proc. GlobeCom*, San Jose, CA, U.S.A., Jan. 2004.

[2] S. Kanumuri, P. C. Cosman, A. R. Reibman, and V. A. Vaishampayan, "Modeling packet-loss visibility in MPEG-2 video," *IEEE Trans. on Multimedia*, vol. 8, no. 2, pp. 341–355, 2006.

[3] T.-L. Lin, S. Kanumuri, Yuan Zhi, D. Poole, P. C. Cosman, and A. R. Reibman, "A versatile model for packet loss visibility and its application to packet prioritization," *IEEE Trans. on Image Processing*, vol. 19, no. 3, pp. 722 –735, Mar. 2010.

[4] T.-L. Lin and P. C. Cosman, "Efficient optimal RCPC code rate allocation with packet discarding for pre-encoded compressed video," *IEEE Signal Processing Lett.*, vol. 17, no. 5, pp. 505 –508, May 2010.

[5] R. J. McEliece and W. E. Stark, "Channels with block interference," *IEEE Trans. Inform. Theory*, vol. 30, no. 1, pp. 44–53, Jan. 1984.

[6] Y. R. Zheng and C. Xiao, "Simulation models with correct statistical properties for Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 51, no. 6, pp. 920–928, Jun. 2003.

[7] S. Kanumuri, S.G. Subramanian, P.C. Cosman, and A.R. Reibman, "Predicting H.264 packet loss visibility using a generalized linear model," in *Proc. ICIP*, Oct. 2006, pp. 2245 –2248.

[8] P. Frenger, P. Orten, T. Ottosson, and A. Svensson, "Multi-rate convolutional codes, technical report no. 21," Tech. Rep., Department of Information Theory, Chalmers University of Technology, Apr. 1998.

[9] "VQM software," in *[Online]. Available: http://www.its.bldrdoc.gov/n3/video/vqmsoftware.htm*.