Subcarrier Mapping Based on Slice Visibility for Video Transmission over OFDM Channels

Laura Toni, Dawei Wang, Pamela C. Cosman, and Laurence B. Milstein
Department of Electrical and Computer Engineering,
University of California, San Diego
La Jolla, CA 92093-0407 USA
{ltoni,daw017}@ucsd.edu, pcosman@eng.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 for scenarios with arbitrary mobility. A slice loss visibility (SLV) model is used to evaluate the visual importance of each H.264 slice. In particular, taking into account both the visibility scores available from the bitstream and the channel state information, we optimize the channel code rate for each GOP and the mapping of video slices within 2-D time-frequency resource blocks, in order to better protect more visually important slices. Results demonstrate that the proposed algorithm outperforms baseline ones which do not take into account the SLV in the video transmission.**

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

I. Introduction

Since video packet losses have different impacts on video quality, cross layer techniques, which take into account both physical and application layer conditions, can reduce the video distortion caused by channel impairments. In particular, the physical layer parameters (e.g., channel coding rate, modulation) can be tuned based on information about both channel conditions and video content [1]. Unlike fine-grain scalable video sequences, in which each bit of the encoded enhancement bitstream within a frame is more important than the subsequent bits, for non-scalable video sequences, assigning priority levels to portions of the compressed bitstream is more challenging, as it is not the simple case that earlier bits are more important than later ones. In [2], the authors optimized an H.264 flexible macroblock ordering (FMO)-based classification of the macroblocks with a jointly optimal channel rate allocation. Rather than operating at the source encoder, in [3], rate-compatible punctured convolutional (RCPC) code rate allocation was optimized for non-scalable pre-encoded video sequences transmitted over additive white Gaussian noise (AWGN) channels. In both these works, the mean square error (MSE) induced by a packet loss was considered as the metric for the cross-layer optimization.

When dealing with quality assessment for sequences transmitted over error-prone channels, an important aspect is the relation between the distortion metric and the packet losses [4]. In [5], the authors investigated this relationship, proposing 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, but do not involve a full decoding or pixel-level reconstruction of the video source. 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 [6], the SLV model was used to design a policy for perceptual-quality based packet discarding. An intermediate router in a congested network, for example, might employ the SLV metric to decide which packets should be dropped to minimize degradation in the quality of the transmitted video streams.

In [3], the SLV model was used for the optimization of the channel code rate in the time domain for video transmission over AWGN single carrier channels. Channel coding optimization for non-scalable encoded video sequences was also addressed in [7]. In particular, a cross-layer technique to optimize the quality of experience (QoE) metric was proposed for slow-fading scenarios.

In this paper, we extend our previous work on cross-layer mapping to scenarios with arbitrary mobility. We aim at maximizing the decoded quality of a non-scalable bitstream when the compressed video sequences are transmitted over a frequency selective orthogonal frequency division multiplexing (OFDM) network. Based on the information available from the feedback channel and the application layer, we propose a technique that jointly optimizes the channel code rate for each GOP, and maps the encoded slices into 2-D resource blocks (RBs). We consider a point-to-point communication system experiencing fading channels at arbitrary mobility.

The remainder of this paper is organized as follows. In Section II, we describe some technical preliminaries. In Section III we discuss the proposed cross-layer diversity approach. In Section IV, we provide simulation results and discussion, and we conclude in Section V.

II. Preliminaries

We consider a non-scalable video encoder and assume that each frame is divided into N_s slices (each slice consists of a constant number of macroblocks). The ith slice of frame k is encoded into $L_k(i)$ bits and has an SLV score $V_k(i)$. The SLV value of each slice ranges from 0 to 1, and estimates the quality

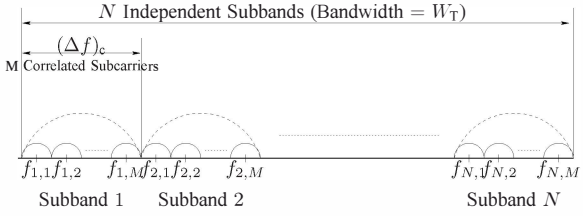

Figure 1. OFDM spectrum.

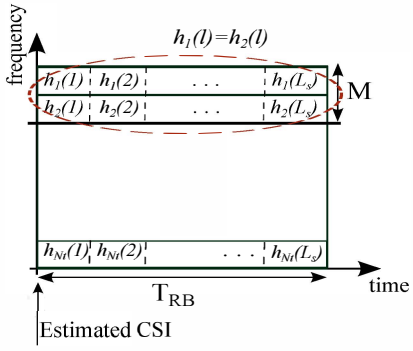

Figure 2. Structure of the OFDM resource block.

degradation the video experiences when that slice is lost. [6]. 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. So, each encoded slice is characterized by the pair $(V_k(i), L_k(i))$, for $i = 1, \ldots, N_s$ and $k = 1, \ldots, N_F$, where N_F is the number of frames per group of pictures (GOP).

The video sequences are transmitted over frequency-selective OFDM channels, and a block fading channel model is used to simulate the frequency selectivity [8]. In this model, the spectrum is divided into blocks of size $(\Delta f)_c$. Subcarriers in different blocks are assumed 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_T, such that we can define N independent subbands. Each subband consists of M correlated subcarriers spanning a total bandwidth of $(\Delta f)_c$. The total number of subcarriers in the OFDM system is $N_t = N \times M$.

In the time domain, we assume the channel experiences Rayleigh fading. We use the modified Jakes model [9] to simulate different fading rates, resulting in different time diversity orders. A concatenation of cyclic redundancy check (CRC) code and RCPC codes, for possible diversity and coding gains in the time domain, are applied to each codeword. We assume that the fading gain $h_j[l]$ experienced by the jth subcarrier at time slot l is distributed as a complex Gaussian random variable with mean zero and variance σ_h^2. Because of the Jakes model, the time correlation function is $\rho_n = \mathbb{E}\{h_j(l)h_j^*(l+n)\} = J_0(2\pi n f_{nd})$, where f_{nd} is the normalized Doppler spread and the superscript * stands for conjugate. We denote by $(\Delta t)_c \propto 1/(f_{nd})$ the channel coherence time.

We denote the instantaneous signal-to-noise ratio (SNR) experienced by the jth subcarrier at time slot l by $\gamma_j(l) = |h_j(l)|^2 E_s/N_0, j = 1, \ldots, N_t$. Note that E_s is the symbol energy, and $N_0/2$ is the two-sided spectral density of the AWGN. The RB will experience a constant mean SNR, which is denoted by $\overline{\gamma} = \mathbb{E}\left\{|h|^2\right\} E_s/N_0 = 2\sigma_h^2 E_s/N_0$, where $\mathbb{E}\{\cdot\}$ denotes the statistical expectation.

As depicted in Fig. 2, each RB consists of L_s QPSK modulated symbols with T_s symbol period, and $T_{RB} = L_s \cdot T_s$ is the duration of the RB, which corresponds to an OFDM symbol duration. Note that the receiver can make CSI estimation more frequently than a RB; however the estimated fading gain is sent to the transmitter and used in the slice mapping optimization only at the beginning of each RB. Due to the time-varying nature of the channel, channel state information (CSI) accurately estimated at the beginning of the RB can become outdated within an OFDM symbol duration. In this work, we assume perfect channel estimation.

III. Cross-Layer Optimization

The main steps of the proposed algorithm are applied to each GOP of the video sequences which we assume to be transmitted within a transmission period denoted by T_{GOP}[1]. Given the OFDM symbol duration, each GOP is transmitted across N_{RB} RBs, where

$$N_{RB} = \left\lfloor \frac{T_{GOP}}{T_{RB}} \right\rfloor .$$

Without losing generality, we assume that N_{RB} is an integer value and $N_{RB} \geq 1$. 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 bit budget for each frame would not make good use of available resources. Instead, we adopt N_{RB} fixed-sized 2D time-frequency RBs for each GOP. This cross-layer choice corresponds to a common approach in application-layer video rate control, in which the number of bits given to each GOP is held roughly constant.

As illustrated in Fig. 3, the current GOP is processed by a concatenation of channel coding and slice mapping steps. In particular, slices from the encoded GOP are processed by the channel encoder. We consider a constant RCPC code rate R_{rcpc} for all the slices of the GOP. Then, each codeword will consist of one or more slices plus the forward error correction (FEC) added by the RCPC code and the bits added by the CRC code.[2] Denoting by Mod the constellation size of the adopted modulation, after channel coding, codewords have constant length [equal to $C_L \cdot \log_2(Mod)$ bits] and will be opportunistically mapped across RBs. In the example of Fig. 3, codewords are allocated across two resource blocks. The proposed method can be described with the following three steps that are processed for each R_{rcpc} code rate. The

[1]Since we impose a frame rate of 30 fps and each GOP consists of 24 frames, $T_{GOP} \leq 24/30s$.

[2]It should be noted that information bits and CRC/RCPC parity symbols would be interleaved in an actual system. However, for illustration, we show the de-interleaved version so that the relative amounts of parity symbols and information symbols can be depicted.

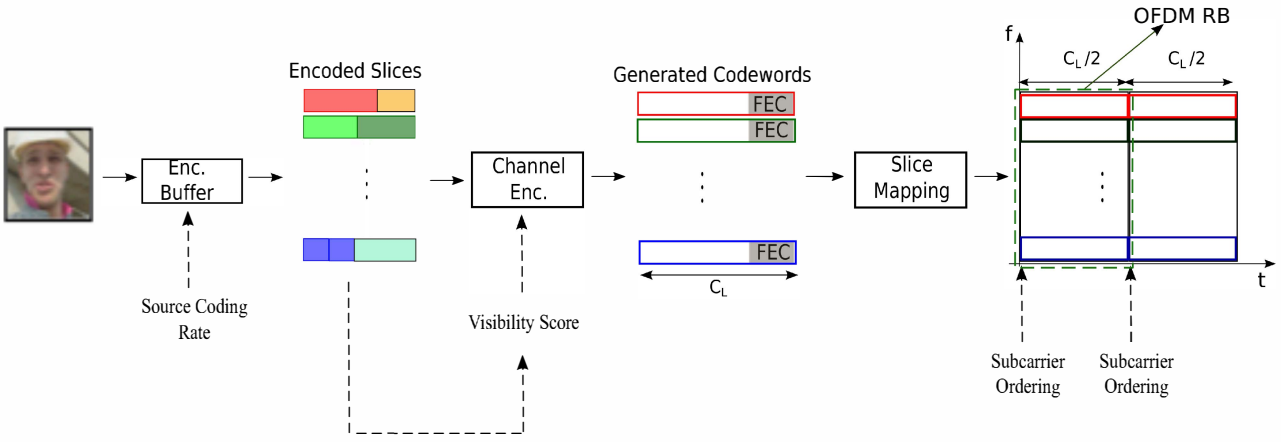

Figure 3. Slice ordering and subcarrier mapping. Note that the FEC parity symbols are interleaved with the information symbols in an actual system.

first step selects the source coding rate at which each GOP needs to be encoded. Note that, in our previous work [7], archival and pre-encoded video sequences were considered for transmission. Thus, no access to the source was possible, and adjustment of the source information rate could only happen by discarding some of the pre-encoded packets. In the current work, we actually consider more general scenarios in which sequences are acquired, encoded and then transmitted. So, having access to the source, we choose the source encoding rate so that there will be minimal discarding of packets. We refer to the previous scenario as "No access to the source", and to the current one as "Access to the source".

Once the source coding rate is selected in the first step, slices are sorted based on the slice priority and encoded in the second step. Finally, in the third step, subcarriers are ordered based on their reliability and the slice mapping is performed. In the following, we describe these steps in detail.

Step 1: Source Coding

Given the RCPC channel code rate, the number of source bits that can be transmitted within a transmission period is given by

$$B_s = \left\lfloor \frac{T_{\text{GOP}}}{T_s} \right\rfloor \cdot \log_2(Mod) \cdot R_{rcpc}\,. \tag{1}$$

In order to reduce the slice dropping rate that is experienced any time the number of bits in which a GOP is encoded is greater than B_s, we would like to encode each GOP in roughly B_s bits. Considering a frame rate of 30 fps and 24 frames per GOP, the source coding rate that we need to impose is given by

$$B_s \left[\frac{bits}{GOP}\right] \times \frac{1}{24} \left[\frac{GOP}{frames}\right] \times 30 \left[\frac{frames}{s}\right] \tag{2}$$

This means that we impose a source coding rate of $B_s \cdot 30/24\, bps$.

Step 2: Slice Ordering & Channel coding

We order all the $N_{\text{F}} \times N_{\text{s}}$ slices of the current GOP from the most visible slices (i.e., the slices which, if lost, are most likely to produce a visible glitch) to the least visible slices. After the ordering, the kth slice will be denoted S_k and will be characterized by the pair (V_k, L_k). Note that $V_k \geq V_{k+1}, \forall k \in [1, N_{\text{F}} \times N_{\text{s}}]$.

The ordered slices are then fed to the channel encoder. In particular, $C_L \cdot \log_2(Mod) \cdot R_{rcpc}$ is the maximum number of source bits allowed per codeword. If one slice is lower than this maximum bit budget, then more than one slice is encoded into a single codeword. In the other case, if the length of a single slice is larger than $C_L \cdot \log_2(Mod) \cdot R_{rcpc}$, the slice is encoded into two or more codewords. We call this phenomenon "slice fragmentation". Moreover, it is worth noting that the transmission period over which a single GOP will be transmitted is constant and the total number of bits that can be sent within this period for a given RCPC code rate is fixed. Thus, if the number of *source bits* in which the GOP is encoded is greater than the number of information bits that can be sent in the transmission period, randomly chosen slices from the least important group are dropped. We use the word "dropped" or "discarded" to describe slices which are pre-emptively dropped by the encoder because the RCPC code rate profile does not allow all the information bits to fit in the RB, whereas we use the word "lost" to describe slices which are allocated to the RB but which fail to be decoded correctly at the receiver. Both dropped and lost slices are concealed at the decoder.

In order to maintain a level of priority between codewords, we sequentially encode the slices (after ordering). This means that the most important slice (or slices) is encoded into the first codeword and so on.

Step 3: Subcarrier Ordering & Slice Mapping

As already mentioned, a given GOP is transmitted through N_{RB} resource blocks. So, $C_L \cdot \log_2(Mod)/N_{\text{RB}}$ is the portion of bits per codeword allocated to each subcarrier of each RB. For each RB, we perform the following two phases:

a) The subcarriers of the 2-D RB are ordered from the most reliable to the least reliable. Denoting by $\boldsymbol{\gamma} =$

$[\gamma_1(1), \gamma_2(1), \ldots, \gamma_{N_t}(1)]$ the estimated SNR information available in the feedback channel, we assume that the subcarriers are ordered in such a way that $\gamma_1(1) \geq \gamma_2(1) \geq \ldots \geq \gamma_{N_t}(1)$.

b) Based on subcarriers and codeword ordering, codewords are allocated to subbands. If $N_{\text{RB}} = 1$, ordered codewords are sequentially allocated to ordered subcarriers such that the most important codewords are sent through the most reliable subcarriers. If $N_{\text{RB}} > 1$ then only a portion of each codeword will be transmitted within a single RB. In particular, subdividing the codewords in chunks of $C_L \log_2(Mod)/N_{\text{RB}}$ bits, the ith chunk of each codeword is mapped to the ith RB. Within each RB, we still allocate the most important chunks (i.e., the one from the most important codewords) to the most reliable subcarriers.

Once the slices are allocated within the N_{RB} RBs, we evaluate the mean Video Quality Metric (VQM) score associated with the considered RCPC code rate by simulating 1000 random realizations of the channel.

Both **Step 2** and **Step 3** are repeated for all possible FEC rates considered in the optimization process. The mean VQM score for each code rate is then compared with the best one, VQM_{best}.[3] If $\text{VQM}(R_{\text{rcpc}}) < \text{VQM}_{\text{best}}$, then the best VQM is updated. With this algorithm, the best channel code rate is chosen for each GOP of the video sequence (GOP-by-GOP optimization). Alternatively, if the final goal of the proposed method is to choose the RCPC channel code rate able to maximize the mean quality of the whole video sequence, all the steps are repeated for all the GOPs and the VQM is evaluated for the whole video sequence. In the following, for simplicity in the optimization, we consider the latter one (i.e., the whole sequence optimization, in which one RCPC channel code rate is selected for the entire video).

IV. Results

We carried out simulations on videos of 10s duration, coded using the H.264/AVC JM codec with SIF resolution (352×240), and with Motion-Compensated Error Concealment (MCEC) implemented in the decoder. For brevity, we provide results for one video sequence, named "HighMot", which has a medium/high motion level and several scene changes. We used the IBBP encoding structure with I-frames every 24 frames. There are $N_t = 128$ OFDM subcarriers in total per RB. The RCPC codes of rates $R_{rcpc} = \left\{\frac{8}{10}, \frac{8}{12}, \frac{8}{16}, \frac{8}{18}, \frac{8}{20}, \frac{8}{24}\right\}$, 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 [10]. We consider QPSK modulation and perfect channel estimation. The codeword size after the RCPC/CRC coding was set equal to 588 bytes, i.e., $L_m = (588 \cdot 4)$ QPSK modulated symbols. Results are provided in terms of the VQM score, which ranges from 0 (best quality) to 1 (worst quality).

For comparison, we consider a baseline algorithm, in which slice importance is not known. In particular, the *Sequential* algorithm sequentially allocates the slices of each frame to the RB.

[3] As an initial value, we set $\text{VQM}_{\text{best}} = 1$.

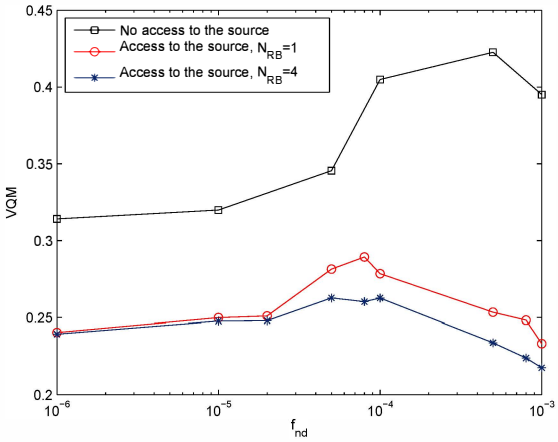

Figure 4. VQM vs. f_{nd} for the cross-layer optimization for systems either with or without access to the source. "HighMot" video is considered for systems with $(N, M) = (32, 4)$, and $\overline{\gamma} = 16$ dB.

We first show a comparison of the current scenario with the one in [7], characterized by archival and pre-encoded sequences. In Fig. 4, the VQM of the decoded sequence as a function of the normalized Doppler has been provided when the cross-layer optimization is performed for systems with and without source coding variation, with $(N, M) = (32, 4)$, and $\overline{\gamma} = 16$ dB. In particular, all the results have been carried out optimizing the slice mapping and the RCPC channel code rate per RB. The "No access to the source" curve corresponds to transmitting a video sequence encoded at $600kbps$, while in the "Access to the source" systems, **Step 1** is performed. As expected, it is seen that the "No access to the source" system is outperformed by the one with varying source coding rate. Note that, in the literature, a VQM gain of 0.1 is considered to be a good improvement, and the gain in Fig. 4 is around 0.1 for most of the Doppler values. Moreover, in the case of "Access to the source", if a subcarrier re-ordering is allowed four times during the codeword transmission (i.e., $N_{\text{RB}} = 4$) the channel outdating effect is reduced and the system outperforms the case of only initial reordering (i.e., $N_{\text{RB}} = 1$). Note that the gain is negligible in slow fading channels, where there is not an outdating effect.

We now show the gain of the cross-layer mapping optimization, i.e., the visibility-based optimization, with respect to the sequential mapping algorithm. In Fig. 5, the VQM is depicted as a function of the available bandwidth (i.e., total number of subcarriers available per RB) for systems with $M = 4$, $\overline{\gamma} = 16$ dB, and $f_{nd} = 10^{-6}$. For most of the bandwidth values, the cross-layer optimization leads to a gain of at least 0.1 in VQM score for a maximum gain of 0.18 achieved with 224 total subcarriers. This cross-layer gain can be interpreted in terms of capacity gain. For example, to achieve a VQM score of 0.3, the cross-layer algorithm needs 102 total subcarriers, with respect to the 224 subcarriers required by the baseline algorithm. This

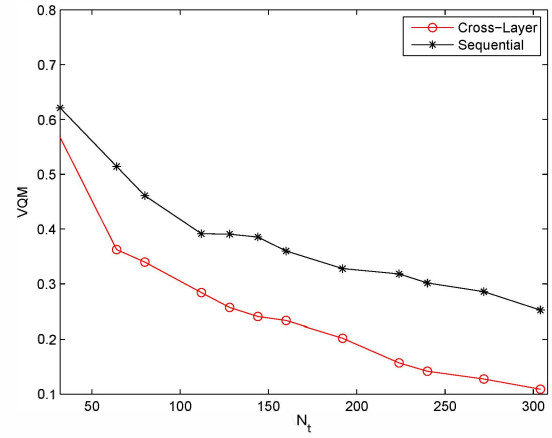

Figure 5. VQM vs. N_{t} for both cross-layer and baseline algorithms for systems with $M = 4, \overline{\gamma} = 16$ dB, and $f_{nd} = 10^{-6}$. "HighMot" video and $N_{\mathrm{RB}} = 1$ are considered.

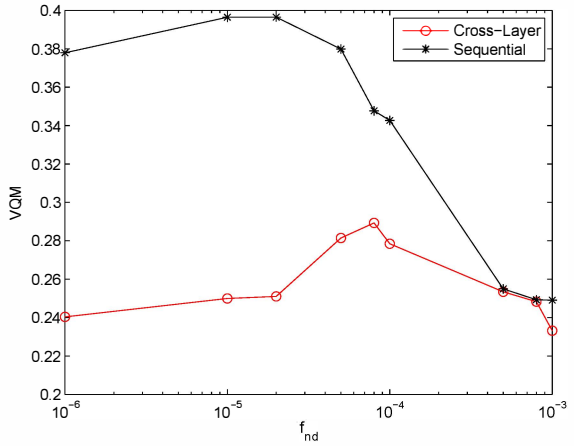

Figure 6. VQM vs. f_{nd} for both cross-layer and baseline algorithms for systems with $(N, M) = (32, 4)$, and $\overline{\gamma} = 16$ dB. "HighMot" video and $N_{\mathrm{RB}} = 1$ are considered.

gain is mainly due to the fact that the cross-layer algorithm makes use of the SLV information, exploiting the most reliable subcarriers. Thus, the best RCPC code rate (i.e., the minimum packet loss rate) required from the cross-layer is higher than the one required from the sequential algorithm. It follows that the baseline optimization requires more FEC and thus more bandwidth in order to send enough source bits such that a 0.3 VQM score is achieved.

Finally, the comparison between cross-layer and baseline algorithms is provided for different Doppler values. In Fig. 6, the VQM vs. f_{nd} is depicted for systems with $(N, M) = (32, 4)$, and $\overline{\gamma} = 16$ dB. Analogous to the previous figure, a gain of the cross-layer scheme w.r.t. the sequential one can be observed for all the Doppler values lower than 10^{-3}. Gains up to 0.15 (for $f_{nd} = 2 \times 10^{-5}$) are experienced in terms of VQM score. For Doppler values around 10^{-3}, in ideal CSI systems, the time diversity is sufficient to balance the outdating effects, leading to a system with high transmission reliability. Then, in this case, both optimization techniques converge to the same VQM scores.

V. Conclusions

We studied the optimization of channel coding in a 2-D time-frequency resource block of an OFDM system, aimed at maximizing the quality of experience when non-scalable compressed video sequences are transmitted. We used a slice loss visibility model to estimate the visual importance of slices to be transmitted over a wireless channel, and to provide the best level of protection to the video slices, opportunistically mapped within the 2D RB. Scenarios with arbitrary mobility and ideal channel estimation have been addressed. Results demonstrated the validity of the SLV-based (cross-layer) model in the optimization process, showing that the proposed algorithm provides a considerable improvement in the system performance in terms of both VQM gain (up to 0.18 in VQM score) and capacity gain.

Acknowledgments

This work has been supported by the Intel-Cisco Video Aware Wireless Networks (VAWN) program, by the office of Naval Research under Grant N00014-11-1-0733, and by the UCSD FWGrid Project, NSF Research Infrastructure Grant Number EIA-0303622.

References

[1] L. Toni, Y. S. Chan, P. C. Cosman, and L. B. Milstein, "Channel coding for progressive images in a 2-D time-frequency OFDM block with channel estimation errors," *IEEE Trans. Image Processing*, vol. 18, no. 11, pp. 2476 –2490, Nov. 2009.

[2] N. Thomos, S. Argyropoulos, N. V. Boulgouris, and M. G. Strintzis, "Robust transmission of H.264/AVC streams using adaptive group slicing and unequal error protection," *EURASIP J. Appl. Signal Process.*, vol. 2006, pp. 120–120, Jan. 2006.

[3] 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.

[4] Y. J. Liang, J. G. Apostolopoulos, and B. Girod, "Analysis of packet loss for compressed video: Effect of burst losses and correlation between error frames," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 18, no. 7, pp. 861 –874, July 2008.

[5] 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.

[6] T.-L. Lin, S. Kanumuri, Y. 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.

[7] L. Toni, P. C. Cosman, and L. B. Milstein, "Channel coding optimization based on slice visibility for transmission of compressed video over OFDM channels," *IEEE J. Select. Areas Commun.*, vol. 30, no. 7, pp. 1172 –1183, August 2012.

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

[9] 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, June 2003.

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