UC San Diego
UC San Diego Previously Published Works

Title

A low complexity model for predicting slice loss distortion for prioritizing H.264/AVC video

Permalink

https://escholarship.org/uc/item/20f1c0fx

Journal

Multimedia Tools and Applications, 75(2)

ISSN

1380-7501

Authors

Paluri, Seethal
Kambhatla, Kashyap KR
Bailey, Barbara A
et al.

Publication Date

2016

DOI

10.1007/s11042-014-2334-2

Peer reviewed

A low complexity model for predicting slice loss distortion for prioritizing H.264/AVC video

**Seethal Paluri · Kashyap K. R. Kambhatla ·
Barbara A. Bailey · Pamela C. Cosman ·
John D. Matyjas · Sunil Kumar**

Abstract The cumulative mean squared error (CMSE) is a widely used measure of distortion introduced by a slice loss. We propose a low-complexity and low-delay generalized linear model for predicting CMSE contributed by the loss of individual H.264/AVC encoded video slices. We train the model over a video database by using a combination of video factors that are extracted during the encoding of the current frame, without using any data from future frames in the group of pictures (GOP). We then analyze the accuracy of the CMSE prediction model using cross-validation and correlation coefficients. We prioritize

S. Paluri (✉)
Ph.D. Candidate, Computational Science Research Center, San Diego State University,
San Diego, CA, USA
e-mail: spaluri@mail.sdsu.edu

K. K. R. Kambhatla
Electrical and Computer Engineering, San Diego State University and University
of California San Diego, San Diego, CA, USA
e-mail: kkambhat@ucsd.edu

B. A. Bailey
Department of Mathematics and Statistics, San Diego State University,
San Diego, CA, USA
e-mail: bbailey@mail.sdsu.edu

P. C. Cosman
Electrical and Computer Engineering, University of California San Diego,
San Diego, CA, USA
e-mail: pcosman@ucsd.edu

J. D. Matyjas
Air Force Research Laboratory, Rome, NY, USA
e-mail: john.matyjas@us.af.mil

S. Kumar
Electrical and Computer Engineering, San Diego State University,
San Diego, CA, USA
e-mail: skumar@mail.sdsu.edu

the slices within a GOP based on their predicted CMSE values. The performance of our model is evaluated by applying unequal error protection, using rate compatible punctured convolutional codes, to the prioritized slices over noisy channels. We also demonstrate an application of our slice prioritization by implementing a slice discard scheme, where the slices are dropped from the router when the network experiences congestion. The simulation results show that *(i)* the slice CMSE prediction model performs well for varying GOP structures, GOP lengths, and encoding bit rates, and *(ii)* the peak signal-to-noise ratio and video quality metric performance of an unequal error protection algorithm using slices prioritized by the predicted CMSE is similar to that of the measured CMSE values for different videos and channel signal-to-noise. We also extend the GOP-level slice prioritization to frame-level slice prioritization and show its performance over noisy channels.

Keywords H.264/AVC · Video coding · CMSE prediction · Slice prioritization · Slice loss · Unequal Error Protection (UEP) · Slice discard

1 Introduction

The demand for real-time video streaming over wireless networks is rapidly growing. In order to efficiently utilize wireless bandwidth, video data is compressed using sophisticated encoding techniques such as H.264/AVC [49], where each video frame is divided into independently coded slices (i.e., network abstraction layer units (NALU)) that consist of a group of macroblocks. However, the transmission of compressed video over wireless channels is highly susceptible to congestion and/or channel-induced packet losses, resulting in error propagation along the motion prediction path, and causing video quality degradation [22, 43, 49]. Loss of some H.264 video slices introduce higher distortion to the video quality than other slices due to the spatio-temporal dependencies and video content.

A commonly used measure of distortion introduced by a slice loss is the cumulative mean squared error (CMSE), which takes into account the distortion in the current frame as well as the temporal error propagation in the future frames of the group of pictures (GOP) [10, 11, 20, 21, 29, 47]. However, computing CMSE for each slice loss introduces computational complexity and delay because it requires decoding the current and subsequent frames of a GOP [20, 21]. Therefore, several schemes for predicting the distortion contributed by the slice loss have been proposed in the literature [3, 23, 24, 30, 39, 41, 48, 52]. Most of them use video features which are extracted from the current as well as the future frames of the GOP, to estimate the distortion introduced by the slices. The accurate prediction of transmission distortion is, however, difficult.

In this paper, we present a novel low-complexity and low-delay generalized linear model (GLM) to predict the CMSE contributed by the loss of individual H.264 AVC video slices. The model uses video factors (such as motion vectors, average integer parts, slice type, initial mean squared error, and temporal duration), which are extracted during the encoding of only the current video frame, without using any data from the future frames in the GOP. The CMSE prediction accuracy of the model is evaluated for test videos encoded using different GOP structures, GOP lengths, and bit rates. This model is also used to design a priority assignment scheme for video slices based on their predicted CMSE values. The performance of the model is then evaluated by applying unequal error protection (UEP) to the prioritized slices over noisy channels and slice discarding for network congestion. Simulation results demonstrate that our model has a very satisfactory performance, given the fact that it uses a limited set of parameters only from the current frame, without using

any information from the future frames in the GOP. Preliminary results of the proposed scheme were presented in [34].

This slice CMSE prediction model can be used in many real-time video streaming applications such as (i) Slice prioritization [13, 15, 16, 20, 21, 44, 53], packet scheduling in a transmitter, and traffic shaping for streaming applications [9]; (ii) Discarding some low priority slices at an intermediate router during network congestion [26]; (iii) Designing UEP schemes where more parity bits are assigned to higher priority slices to protect them against channel errors [21]; (iv) Determining the optimal packet and fragment sizes in cross layer schemes [20, 21].

The remainder of the paper is organized as follows. Section 2 presents past research on modeling video quality and how our scheme is different. Section 3 discusses the video factors used to model the impact of a slice loss, followed by the GLM model development for CMSE prediction in Section 4. Section 5 discusses experimental results, followed by a frame-level slice prioritization scheme in Section 6. We discuss an application of slice prioritization in slice discard in a congested network in Section 7. Finally, Section 8 presents our conclusions.

2 Related work

Video quality is influenced by various network dependent and application oriented factors, such as packet losses, video loss recovery techniques, and encoder configurations [25, 26, 38]. Although peak signal to noise ratio (PSNR) and mean squared error (MSE) do not always reflect perceptual quality well, they have been commonly used to measure video quality [9, 20, 21, 36]. Kanumuri et al. [19] used a tree structured classifier that labeled each possible packet loss in MPEG-2 video as being either visible or invisible, and also developed a GLM to predict the probability that a packet loss will be visible to a viewer. A versatile model was developed by Lin et al. [26] for predicting the slice loss visibility to human observers for the MPEG-2 and H.264 encoded videos, by considering loss of one slice at a time.

The performance of different objective video quality assessment methods was evaluated for the Laboratory for Image and Video Engineering (LIVE) Video Quality Database in Chikkerur et al. in [14] and Seshadrinathan et al. [40]. Both full-reference and reduced-reference schemes were considered in [14] and each scheme was classified based on whether it used the natural or perceptual characteristics. Both papers found that the MultiScale-Structural SIMilarity (MS-SSIM) index, VQM and Motion-based Video Integrity Evaluation (MOVIE) index showed the best performance. Hemami and Reibaman [17] reviewed the necessary steps to design effective no-reference quality estimators for images and video. They outlined a three stage framework for the no-reference quality estimators that permitted factors from human visual system to be incorporated throughout.

Recently, Zhang et al. [51] evaluated video quality by examining the impact of quantization, frame discarding and spatial down-sampling for video transcoding. Further, they proposed a no-reference multidimensional video quality metric by taking into account the per-pixel bitrate, and spatio-temporal activity and showed its effectiveness when the frame rate and frame sizes changed simultaneously. Lottermann and Steinbach [28] presented a model to compute the bit rate of H.264/AVC video by using the quantization parameter, frame rate, the GOP length and structure, and content-dependent parameters. The content dependent parameters were computed by examining temporal and spacial activity which was determined from uncompressed video.

Several schemes also exist for video packet (or slice) prioritization. Typically, these prioritization schemes can be classified into two categories, i.e., heuristic schemes or computationally intensive schemes. Heuristic schemes rely on the type of the frame (or slice), GOP structure and/or its length to determine its frame/slice priority [13, 15, 16, 44, 53]. These schemes do not lend themselves to systems that support adaptive reliability. The more computationally intensive schemes are more accurate with the distortion estimation but have more computational complexity. For example, the recursive per-pixel end-to-end distortion estimate (ROPE) algorithm evaluates the expected distortion at a pixel level by considering the error propagation in intra and inter-coded macroblocks [39, 41, 52]. However this scheme has the drawback of memory requirements as two moments have to be tracked for each pixel.

Liu et al. [27] developed a macroblock prioritization scheme to improve the loss resilience of parallel video streams over a two-class DiffServ network. By jointly exploiting the H.264 flexible macroblock ordering (FMO) tool, a multi-stream macroblock ordering framework was designed to classify all macroblocks of a super-frame into two categories: important macroblocks as high-reliability traffic class and unimportant macroblocks as best-effort traffic class. This scheme effectively reduced the compound transmission distortion of the parallel video streams. Srinivasan et al. in [42] addressed the question of how many priority levels a single layer H.264 encoded bitstream required when the encoded frames were statistically multiplexed in transport networks. The authors conducted simulations with a modular statistical multiplexing structure. They showed that for buffered statistical multiplexing, frame prioritization did not significantly impact the number of support streams. In [21], we designed a cross-layer priority-aware packet fragmentation scheme at the medium access control (MAC) layer. The slices were prioritized based on their measured CMSE values; no CMSE prediction model was used for the slices.

Another approach adopted in [24] determines the video distortion by taking the expectation over all possible channel realizations. The authors assume that each frame is packaged into a single packet, which is not a realistic assumption. Their model also assumes guaranteed reception and decodability of I-frames, which is impractical in real time systems. Wang et al. [48] follow a similar approach to distortion estimation by modeling the temporal attenuation as a function of packet loss ratio and the proportion of intra-coded macroblocks. Similar to [24], Wang et al. assume single slice per frame packetization and derive expressions for distortion estimation. They discuss unequal error protection for single frame loss by assuming an additive distortion behavior. Babich et al. [3] proposed models that use MSE of consecutive frames to estimate the distortion caused by error propagation due to channel loss. The authors define decay constants in their computation, which increase memory constraints as the frame loss history has to be tracked. Masala et al. [30] proposed a scheme called analysis-by-synthesis, which estimates the channel induced distortion of each packet individually. The co-impact of multiple losses is modeled additively but due to the exhaustive nature of all the loss patterns, it is computationally intensive. The authors address this issue by proposing to use their scheme to evaluate distortion in the current frame while error propagation is estimated through another model. Li et al. [23] proposed a transmission distortion modeling scheme for pre-encoded videos in the compressed domain, by taking into account packet losses in the networks and the video content. This approach relies on extracting video features which are then used in a predictive model to evaluate the distortion caused by the video transmission under the current network conditions.

Recently, Schier and Welzl [38] demonstrated how the packet (i.e., NALU) prioritization scheme can be beneficially integrated into systems with tight time constraints. They developed a NALU prioritization scheme to measure the impact on video quality by

considering macroblock partitions, motion vectors, and temporal prediction dependencies between NALUs. Baccaglini et al. [4] proposed a UEP scheme that allocated forward error correction (FEC) codes to video slices according to their impact on the GOP distortion. Zhang et al. [54] developed a hierarchical UEP scheme for H.264/AVC video packets; the level of protection for a slice was determined by considering its frame number in the GOP, the per-frame bitrate, and the data partition type. Perez and Garcia [35] studied the effect of packet losses in video sequences by using a simple prioritization scheme to determine a UEP strategy and demonstrate its effectiveness over a random packet drop. Miguel et al. [32] proposed a multicast distribution system for High-Definition video over wireless networks based on rate-limited packet retransmission. They developed a retransmission scheme based on the packet priority which is assigned based on the packet content and their delay limitation.

The main contributions of our work are: (i) Our model predicts the CMSE distortion introduced by a slice loss in real-time, which is a widely used measure of distortion [9, 20, 21, 36]. (ii) The random forest method [7] is used to determine the importance of video factors, which also helped us in selecting additional factors based on the interactions among the most important factors. (iii) Because the proposed scheme does not use any factors from future frames, our model can be useful in frame-based slice priority assignment for applications with stringent delay constraints. As discussed in Section 1, the proposed CMSE prediction model can be used in many cross-layer network protocols for enhancing video quality over wireless networks.

3 Video factors affecting slice loss distortion

The loss of a slice can introduce error propagation in the current and subsequent frames within the current GOP. In order to build a model that accurately predicts the CMSE, many video factors of the current and future frames (within the current GOP) that affect the distortion introduced by a slice loss could be used in the model. In order to minimize the complexity of our model and avoid the delay introduced by using factors from future frames, we only use the following factors that can be easily extracted during the encoding of the current frame, without depending on any future frame [34]. Let the n-th original uncompressed video frame be $f(n)$, the reconstructed frame without the slice loss be $\hat{f}(n)$, and the reconstructed frame with the slice loss be $\tilde{f}(n)$.

– **Motion Characteristics**: The magnitude of the distortion induced by a slice loss is influenced by motion. Significant motion activity between two successive frames implies that lost slices may be difficult to conceal. For each slice, we define **MOTX** and **MOTY** to be the mean motion vector in the x and y directions over all the macroblocks in the slice.

– **AVGINTERPARTS**: A slice loss in a complex scene, (e.g., a crowded area), can also be difficult to conceal at the decoder, resulting in a high distortion. **AVGINTERPARTS** represents the number of macroblock sub-partitions averaged over the total number of macroblocks in the slice. If the underlying motion is complex, AVGINTERPARTS would be high.

– **Maximum Residual Energy (MAXRSENGY)**: First, Residual Energy (**RSENGY**) is computed for a macroblock as the sum of squares of all its integer transform coefficients after motion compensation. Then **MAXRSENGY** of a slice is equal to the highest

RSENGY value of its macroblocks. If a scene has high motion, its MAXRSENGY would also be high.

– **Signal Characteristics**: We consider mean **SigMean** and variance **SigVar** of the slice luminance. We also consider the slice type **Slice_type**, such as IDR or P or B slice. This is treated as a categorical factor in our model discussed in Section 4.2.

– **Temporal Duration (TMDR)**: It is defined as the maximum possible temporal error propagation length due to a slice loss. A slice loss in a non-reference B frame has a TMDR of 1, while a slice loss in a reference IDR slice could propagate to the end of GOP.

– **Initial Mean Squared Error (IMSE)**: It is computed as the MSE between the compressed frame $\hat{f}(n)$ and the reconstructed frame with slice loss $\tilde{f}(n)$ within the encoder. Assuming that each frame has N x M pixels, the MSE introduced by a slice loss in the n-th video frame is computed as

$$\frac{1}{NM} \sum_{i=1}^{N} \sum_{j=1}^{M} (\widehat{Pel}_{i,j} - \widetilde{Pel}_{i,j})^2 \tag{1}$$

Here, $\widehat{Pel}_{i,j}$ and $\widetilde{Pel}_{i,j}$ represent the pixel intensity at co-ordinate (i, j) in frames $\hat{f}(n)$ and $\tilde{f}(n)$ respectively.

– **Initial Structural Similarity Index (ISSIM)**: It is a measure of the structural similarity between two frames [37]. Like IMSE, ISSIM is computed per slice using $\hat{f}(n)$ and $\tilde{f}(n)$.

Cumulative Mean Squared Error (CMSE): The loss of a slice in a reference frame can introduce error propagation in the current and subsequent frames within the GOP. CMSE is computed by systematically discarding one slice at a time and measuring the distortion introduced by it as the sum of MSE over the current and subsequent frames in the GOP. This is computationally intensive. We use the measured CMSE as the "ground truth" in our model.

4 Development of slice CMSE prediction model

We use a database of sixteen video sequences shown in Table 1. Seven of these sequences have a spatial resolution of 720×480 pixels and the remaining nine have a spatial resolution of 352×288 pixels. These video sequences have a wide range of content, motion, and complexity. The videos were compressed at different bitrates at 30 frames per second

Table 1 Summary of the video sequences used in database

Motion	Resolution	Name of Sequence
Low	720x480	Mother and Daughter, Boat
	352x288	Silent, Hallway, Container
Medium	720x480	Open Ceremony, Woods
	352x288	Tempete, Coastguard, Ice
High	720x480	Whale Show, News, Tiger
	352x288	City, Soccer, Football

(fps), by using the JM reference software of H.264/AVC. We used three different GOP structures, i.e. (IDR B P ... B, IDR P P ... P, and Hierarchical) with a combination of GOP lengths of 20, 40 and 60 frames, a fixed slice configuration, and dispersed flexible macroblock ordering (FMO) with 2 slice groups. The average number of frames per sequence is approximately 300. Since the compression efficiency is higher for the predicted P and B frames than the intra-coded IDR frames, a slice from an IDR frame typically contains fewer macroblocks than a P-frame slice and significantly fewer macroblocks than a B-frame slice. At the decoder, error concealment was performed using temporal concealment and spatial interpolation. The type of error concealment used depends on the frame type and the type of loss incurred. If the entire frame is lost, motion copy is performed, where the motion vectors are copied from the previous reference frame and all the macroblocks are reconstructed from it. When a slice from P or B frame is lost, error concealment is performed by first verifying if the motion vectors of the neighboring macroblocks/macroblock sub-partitions are available and whether the number of available motion vectors is greater than a pre-defined threshold. If sufficient motion information is available, motion copy is performed else the co-located macroblocks from the previous reference frame are directly copied [5, 50]. Any slice loss from an IDR frame is concealed through spatial interpolation. Figure 1 shows our proposed CMSE prediction scheme. Once the database is formed, a model can be developed as discussed in the next section.

4.1 Overview of model development

In this paper, we use a GLM to predict the CMSE contributed by a single slice loss. The GLMs are an extension to the class of linear models [31]. The advantage of using a GLM is that it allows the response variable to follow any distribution as it may be difficult to estimate the true nature of the underlying probability distributions from a sample of observations.
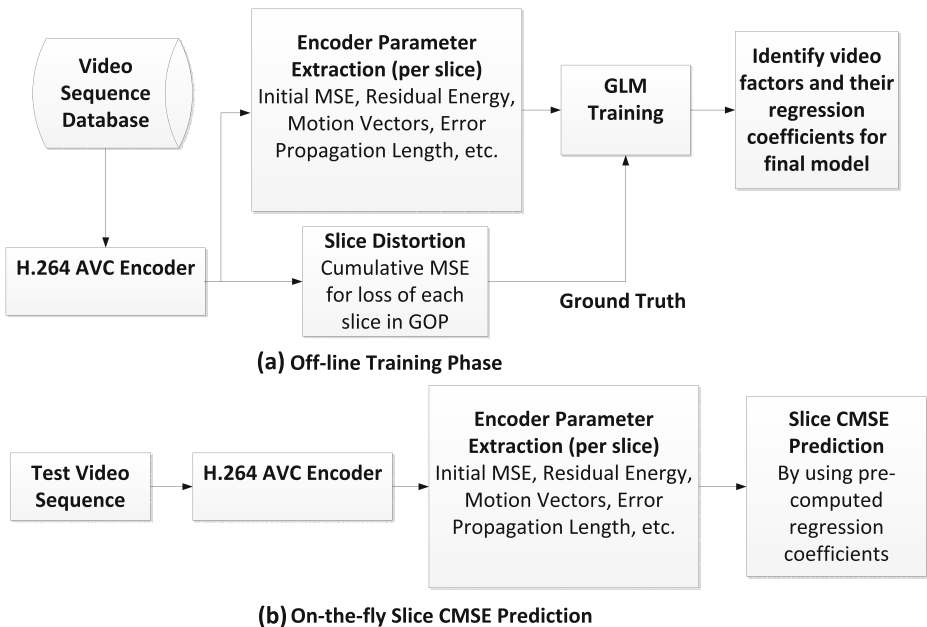

Fig. 1 The proposed CMSE prediction scheme

This facilitates easy modification of the model should the response variable change. Let $\mathbf{Y} = [y_1, y_2, ..., y_N]^T$ be a vector of our response variable, i.e., measured CMSE values. The linear predictor of the response variable is described as a linear combination of the video factors described in Section 3. Here, $\mathbf{X}$ is an $N \times (m + 1)$ matrix of covariates (i.e., video factors), where each row, $[1, x_1, x_2, ..., x_m]$ corresponds to a single observation from $\mathbf{Y}$. Let $\beta = [\beta_0, \beta_1, ..., \beta_m]^T$ be a $(p + 1) \times 1$ vector of unknown regression coefficients that need to be determined from the data. The link function denoted as $g(\cdot)$ describes a function of the linear predictor, $g^{-1}(\mathbf{X}\beta)$. The regression coefficients are estimated through an iteratively re-weighted least squares technique [12]. After estimating β, we use it to predict the response variable vector $\mathbf{Y}$, computed as $\mathbb{E}(\mathbf{Y}) = g^{-1}(\mathbf{X}\beta)$.

4.2 Model fitting

In model fitting, a subset of covariates (i.e., video factors) is chosen for the best fit of the response variable. We use the statistical software R [45] for our model fitting and analysis. The procedure for selecting the order of covariates is given below.

Evaluating the Distribution of the Response Variable: A visual analysis of our response variable (i.e., measured CMSE) shows that low CMSE values occur more frequently than higher CMSE values. The maximum CMSE value in our database is 3500. Figure 2 shows the measured CMSE values. CMSE values greater than 500 are not shown in the figure because their probability density is very low. Non-parametric density estimation was performed using a normal kernel with a bandwidth chosen based on cross-validation. We hypothesize that the measured CMSE follows a distribution from the exponential family of distributions.
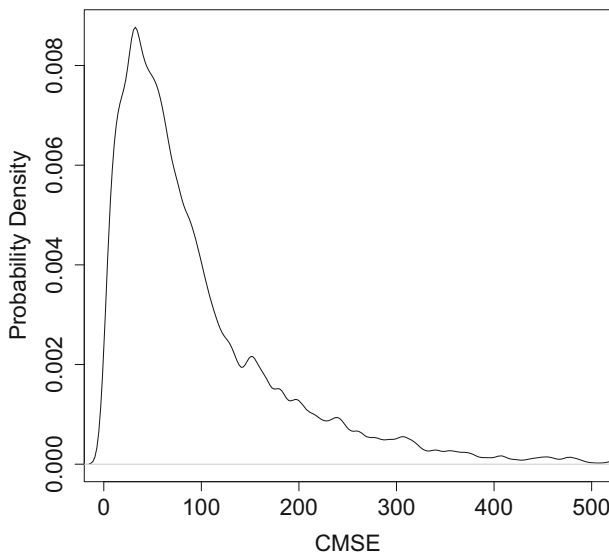
Fig. 2 Probability density estimate of slice CMSE values for videos in database

We used the Anderson-Darling test [2] to examine whether the measured CMSE has a Gaussian-like distribution. Let n be the number of slices in our database and $u_i = F(x_i)$ be the cumulative distribution function (CDF) of the CMSE as denoted by x_i. The test statistic is then calculated as $A_n^2 = -n - \frac{1}{n}\sum_{j=1}^{n}(2j-1)[\log u_i + \log(1 - u_{n-j+1})] = 467.31$. This test statistic has a corresponding p-value of 0.001. This suggests that the exponential family with identity as the link function is a reasonable choice. Figure 3 shows the binned measured normalized-CMSE fitted to a Gaussian distribution. Here, the normalized CMSE values greater than 0.2 correspond to a CDF value of 1.

Choosing Covariates: We use the Akaike information criterion (AIC) [1] index to determine the order of the covariates to be fitted because the dimensionality of the covariate matrix is not high [8]. The decrease in information criterion (IC) value during covariate selection is indicative of a good model fitting process. The IC takes into account the number of potential covariates (m), the number of observations (n), and the log-likelihood of the model (L_{max}). The AIC is computed as $-2\log(L_{max}) + 2m$. A GLM with m parameters has 2^m potential models. We let $\mathbf{Y}^k$ represent the model with a subset of k covariates. The i-th data point in $\mathbf{Y}^k$, y_i^k, where $i = 1, 2, ..., N$ is expressed as:

$$y_i^k = \beta_0^k + \beta_1^k x_{i1} + \beta_2^k x_{i2} + \ldots + \beta_k^k x_{ik} + \epsilon_i \tag{2}$$

Here, β_0 is the intercept, β_j^k, $j = 1, 2, ..., k$ are the regression coefficients, x_{ij} represents the j-th covariate for the i-th observation in $\mathbf{Y}^k$, and ϵ_i the random error term. The response variable vector is computed as $\mathbb{E}(\mathbf{Y}) = g^{-1}(\mathbf{X}\beta)$ with $\mathbb{E}(\epsilon) = 0$. The simplest model is the null model having only the intercept β_0^k whereas the full model has all the
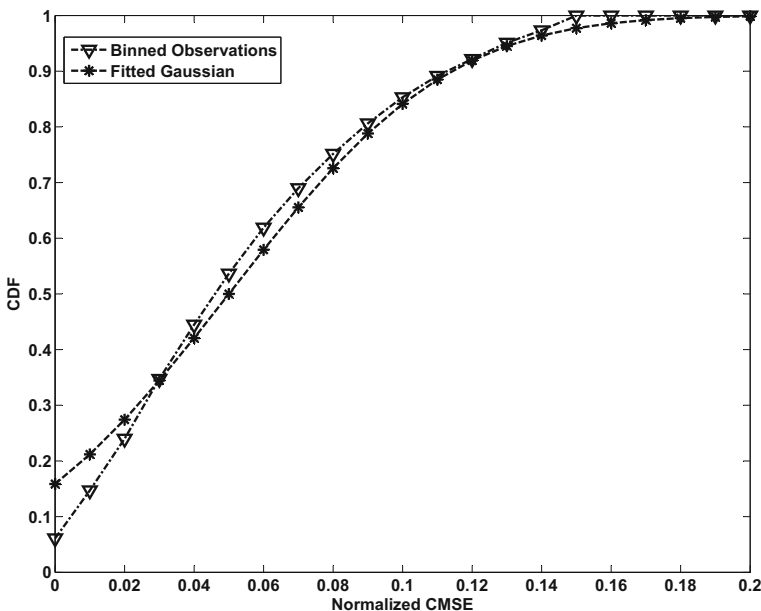

Fig. 3 CDF of the normalized-CMSE for the binned observations and fitted Gaussian distribution

m covariates, $k = m$. We use the forward stepwise approach to choose the covariates [46].

Stopping Criterion: The model is fitted with covariates until the best model with $k + 1$ covariates has a higher AIC value than the model with k covariates. With our data, the stopping criteria was not satisfied even when we fitted the full model (i.e., $k = m$) because our model only relies on factors derived from the current frame and not the future frames of the GOP. As discussed in Section 1, the CMSE also depends on the future frames of the GOP. We improved the performance of our model by introducing interactions between the most important covariates by using the random forest approach as discussed below.

Random Forest: We determine the importance of the factors by using the random forest (RF)[6], a tree structured classifier that estimates the measured CMSE from a subset of video factors. RF consists of decision trees (typically 100,000 or more) that are grown to the full extent using a binary recursive partitioning. We computed the covariate importance using 100,000 trees. Figure 4 shows that the maximum increase in the node purity is achieved by IMSE, followed by TMDR, MAXRSENGY and MOTX. We also introduced two interactions between the first three covariates, i.e., IMSE and TMDR, and IMSE and MAXRSENGY. Adding additional interactions did not improve the performance of our model significantly.

The regression coefficients of our model are reported in Table 2. The magnitude of the prefix of the regression coefficients is indicative of the range of values that the covariate can take. The sign of a regression coefficient does not always provide an insight into its impact on CMSE [33]. However, the order in which the factors are added to the model gives an indication of their importance.
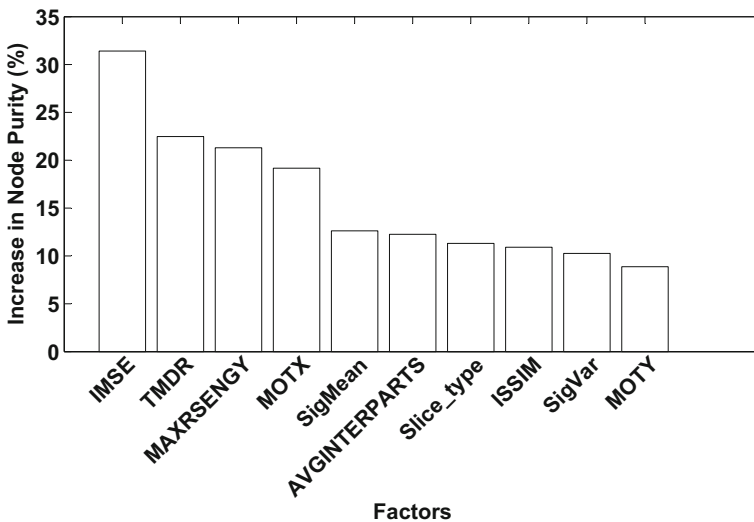
Fig. 4 The importance of covariates determined by the random forest

Table 2 Final model coefficients

Factor No.	Factor	Regression Coefficient
1.	Intercept	1.74×10^1
2.	IMSE	3.93×10^{-1}
3.	TMDR	-9.41×10^{-1}
4.	MAXRSENGY	-1.10×10^{-8}
5.	MOTX	3.32×10^{-1}
6.	SigMean	-2.95×10^{-2}
7.	AVGINTERPARTS	-2.51
8. a	Slice_type.f2	-5.34×10^1
8. b	Slice_type.f3	-6.95×10^1
9.	ISSIM	5.44×10^1
10.	SigVar	3.58×10^{-3}
11.	MOTY	-2.67×10^{-1}
12.	IMSE × TMDR	4.93×10^{-1}
13.	IMSE × MAXRSENGY	7.71×10^{-10}

5 Performance of CMSE prediction model

We compare the performance of our CMSE prediction model with the measured slice CMSE values. Specifically, the performance of our model is evaluated by studying (i) the goodness of fit using the "leave-one-out" cross validation to compute the normalized root mean squared error (NRMSE) of the predicted CMSE, and (ii) the correlation coefficients ρ of the predicted vs. the measured CMSE values. The model is then used to compute the misclassifications in the slice priority assignment with respect to the measured CMSE and to evaluate UEP performance of the prioritized data over an AWGN channel. For this, four test video sequences are used, which were not used during training, each with approximately 300 frames, Akiyo (352×288, slow motion), Foreman (352×288, medium motion), Bus (352×288, high motion), and Table Tennis (720×480, high motion). These video sequences are encoded at various bitrates and GOP structures. The correlation coefficients are evaluated for various GOP structures, whereas the misclassification in the slice priority assignment and UEP performance are evaluated for only one GOP structure.

5.1 Model accuracy

We study the model accuracy by examining the NRMSE as factors are added to the null model. The NRMSE is defined as the ratio of the root mean squared error (RMSE) to the range of the measured CMSE of the training set. The range of measured CMSE is 3500, which corresponds to the maximum CMSE value in our database. The decrease of NRMSE indicates an improved performance of the CMSE prediction. The NRMSE for the null model is 3.56. As factor 2 (from Table 2) (IMSE) is added to the null model, the NRMSE decreases significantly from 3.56 to 0.875. This decrease is further enhanced by adding factors 3 (TMDR) and 4 (MAXRSENGY). Next, as interaction factors 12 (IMSE × TMDR) and 13 (IMSE × MAXRSENGY) are added to the model, a further drop in NRMSE is observed. The decrease in NRMSE becomes more gradual when the last three factors 9 (ISSIM), 10 (SigVar) and 11 (MOTY) are added. Adding additional interactions from other factors gave negligible improvement in NRMSE.

5.2 Slice CMSE prediction accuracy

The Pearson's correlation coefficients of the measured and predicted CMSE values are used to demonstrate the prediction accuracy of our model. The correlation coefficients for the test video sequences Akiyo, Foreman, and Bus (encoded at 1Mbps), and Table Tennis (encoded at 2Mbps) are 0.96, 0.78, 0.89, and 0.88, respectively. For Akiyo, Foreman and Bus encoded at a lower bit rate of 512 Kbps and Table Tennis encoded at 1Mbps, the correlation coefficients are 0.96, 0.83, 0.90 and, 0.91 respectively. Note that these test sequences were not used for model training. Here, a GOP structure IDR B P B... was used, with GOP length of 20 frames and dispersed mode FMO with two slice groups. The performance of our model is also examined for other encoder configurations, as discussed below. When the GOP length was increased from 20 frames to 40 frames, the correlation coefficients for the sequences are 0.92, 0.74, 0.73, and 0.87, respectively. When the slices were formed without using FMO, the correlation coefficients for the four sequences were 0.96, 0.87, 0.94 and 0.93. For the GOP structure IDR P P ... P, and GOP length of 20 frames, the correlation coefficients are 0.96, 0.80, 0.77, and 0.83. For the Hierarchical GOP structure, the correlation coefficients are 0.84, 0.74, 0.82, and 0.78. This indicates that our model performs reasonably well when video characteristics and encoder configurations are changed.

5.3 Accuracy of priority assignment for slices

The slice CMSE prediction allows us to design a flexible scheme for assigning the priorities to different video slices. The slice priorities are beneficial in designing efficient cross-layer network protocols such as routing, fragmentation, assigning access categories for IEEE 802.11e, and for investigating UEP over error prone channels [20, 21]. We divide all the slices of each GOP into four equally populated priority classes ($P1$, $P2$, $P3$, and $P4$), where priority $P1$ corresponds to the highest CMSE values. We call the predicted CMSE based slice priority scheme as Predicted_CMSE_GOP and the measured CMSE based slice priority scheme as Measured_CMSE_GOP. Table 3 shows the percentage of slices contributed by each frame type in the encoded bitstreams. The L after the sequence name indicates a low bitrate of 512 Kbps and the H indicates a high bitrate of 1 Mbps. On average, 25 %, 53 % and 22 % of slices belong to the IDR, P and B frames, respectively. The percentage of IDR slices is higher for the slow motion Akiyo and decreases with increase in motion of a video sequence. Table 3 also shows the contribution of each frame type (as a percentage of the total number of slices in the GOP) towards the four priorities. Approximately 51 % and 30 % of IDR slices belong to the two highest priorities $P1$ and $P2$, respectively. However,

Table 3 Percentage slices contributed by IDR, P, and B frames to different priorities for Akiyo, Foreman, and Bus

Sequence	$IDR(P1/P2/P3/P4)$	$P(P1/P2/P3/P4)$	$B(P1/P2/P3/P4)$
Akiyo(L)	27.7(19.2/6.7/1.8/0.0)	51.0(5.8/17.0/16.5/11.7)	21.3(0.0/1.0/8.5/11.8)
Akiyo(H)	27.4(19.2/6.8/1.3/0.1)	53.0(6.0/16.4/15.5/15.1)	19.6(0.0/2.0/6.8/10.8)
Foreman(L)	25.3(14.8/7.7/2.3/0.5)	52.7(9.6/14.6/15.5/13.0)	22.0(0.4/2.8/7.7/11.1)
Foreman(H)	23.3(15.4/6.3/1.5/0.1)	55.4(9.0/17.2/16.7/12.5)	21.3(0.6/1.7/6.7/12.3)
Bus(L)	23.4(10.8/9.7/2.4/0.5)	54.0(14.1/14.7/9.6/15.6)	22.5(0.0/0.5/13.0/9.0)
Bus(H)	21.1(11.2/6.7/2.9/0.3)	54.3(14.0/16.1/13.0/11.2)	24.7(0.0/2.2/8.9/13.6)

the contribution of IDR slices to the highest priorities decreases with increase in motion of the video sequence. An average of 18 %, 30 %, 27 %, and 25 % of the P slices belong to priorities $P1$, $P2$, $P3$ and $P4$, respectively. The contribution of IDR slices to the lower priorities is much less than that of the P or B frames. Approximately 39 % and 52 % of the B slices lie in the lower priorities $P3$ and $P4$, respectively. IDR slices have the highest contribution to $P1$ priority followed by P slices. The P slices have the highest contribution towards the remaining three priorities.

Next, we discuss the priority misclassification of slices in Predicted_CMSE_GOP compared to Measured_CMSE_GOP. We let $p_j = i$ for $i \in \{1, 2, 3, 4\}$ represent the priority of the j-th slice assigned using measured CMSE. Then we define the first degree (1°) misclassification if that slice were assigned a $p_j = i + 1$ or $p_j = i - 1$ priority using the predicted CMSE. For example, if $p_j = 1$ then a first degree (1°) misclassification would result in $p_j = 2$. Likewise, a second degree (2°) misclassification would result if the slice is assigned a priority of $p_j = i + 2$ or $p_j = i - 2$ using the predicted CMSE. In a third degree (3°) misclassification, a slice with the highest priority is assigned the lowest priority or vice versa. Table 4 shows the percentage misclassification of Akiyo, Foreman, and Bus video slices from each priority. Most of the misclassified slices belong to 1° and approximately half of them are in lower priorities $P3$ and $P4$. Only about 5 % of these misclassified slices belong to the highest priority $P1$. The 2° misclassification is less than 4 %, and 3° misclassification is less than 1 % . We will show in the next section that our slice priority assignment scheme is able to achieve considerable accuracy for our UEP scheme over additive white Gaussian noise (AWGN) channels.

5.4 Performance of UEP scheme for prioritized slices over AWGN channels

We designed a UEP FEC scheme for AWGN channels to compare the performance of slice priority assignment schemes based on predicted and measured CMSE values. We use the following formulation to find the optimal rate compatible punctured convolutional (RCPC) code rate(s) for the UEP schemes.

We formulate the total expected video distortion of our prioritized data as in [20]. Let R_{CH} be the transmission bit rate of the channel. The video is encoded at a frame rate of f_s frames per second, and the total outgoing bit budget for a GOP of frame length L_G is $\frac{R_{CH}L_G}{f_s}$. The RCPC code rates are chosen from a candidate set **R** of punctured code rates $\{R_1, R_2, R_3, ..., R_K\}$. We aim to minimize the sum of weighted slice loss distortion over the AWGN channel. In the sum, the CMSE distortion $D(j)$ contributed by the loss of the j-th slice is weighted by the probability of losing that slice, which depends on the slice size

Table 4 Percentage slice misclassification by degree (1°, 2°, and 3°) for each priority for Akiyo, Foreman, and Bus

Sequence	1°($P1/P2/P3/P4$)	2°($P1/P2/P3/P4$)	3°($P1/P2/P3/P4$)
Akiyo(L)	3.6/18.4/7.8/4.2	0.2/2.4/0.2/0.0	0.0/0.28/0.0/0.0
Akiyo(H)	2.7/7.8/15.8/5.6	0.1/2.5/0.2/1.0	0.2/0.0/0.0/0.4
Foreman(L)	4.9/9.3/10.0/4.9	0.4/0.3/1.0/0.7	0.0/0.28/0.0/0.0
Foreman(H)	5.2/10.2/9.5/11.0	0.4/0.2/1.0/0.5	0.0/0.0/0.0/0.0
Bus(L)	4.9/9.8/9.0/3.9	0.6/0.2/0.5/1.7	0.0/0.0/0.0/0.2
Bus(H)	4.9/10.1/8.4/3.5	0.5/0.4/0.6/1.1	0.0/0.0/0.0/0.0

$S(j)$ in bits and on the bit error probability p_b after channel decoding. Here p_b depends on the channel signal to noise ratio (SNR) and on the RCPC code rate r_i selected for the slice priority i. The optimization problem is formulated as [34]:

$$\min_{\mathbf{r}} \left\{ \sum_{i=1}^{4}\sum_{j=1}^{n_i} \left[1 - (1 - p_b(SNR, r_i))^{\left(\frac{S(j)}{r_i}\right)} \right] D(j) \right\}$$

subject to

$$(1) \quad \sum_{i=1}^{4}\sum_{j=1}^{n_i} \frac{S(j)}{r_i} \leq \left(\frac{R_{CH}L_G}{f_s} \right)$$

$$(2) \quad r_{i-1} \leq r_i \quad \text{for} \quad i = 2, 3, 4$$

$$(3)$$

Here n_i is the number of slices of priority i. The formulation only considers slice loss distortion and ignores compression distortion. Constraint (1) is the channel bit rate constraint and constraint (2) ensures that higher priority slices have code rates which are at least as strong as the code rates allocated to the lower priority slices. The optimization problem is solved using the Branch and Bound (BnB) algorithm with interval arithmetic analysis [18] to yield the optimal UEP code rates.

The mother code of the RCPC code has rate 1/4 with memory M = 4 and puncturing period P = 8. Log-likelihood ratio (LLR) was used in the Viterbi decoder. The RCPC rates were {(8/9), (8/10), (8/12), (8/14), (8/16), (8/18), (8/20), (8/22), (8/24), (8/26), (8/28), (8/30), (8/32)}. The channel bit rate used for videos encoded at 1Mbps (2Mbps) is 2Mbps (4Mbps). Figures 5, 6, 7 and 8 show the average PSNR and video quality metric (VQM) score computed over 100 realizations of each AWGN channel SNR for Akiyo, Foreman, Bus and Table Tennis. A VQM value of 0 (1) represents the best (worst) video quality. The UEP performance of Predicted_CMSE_GOP scheme closely follows that of the Measured_CMSE_GOP. This demonstrates that our predicted CMSE based priority assignment scheme performs well. The other two schemes shown in these figures will be discussed in the next section.
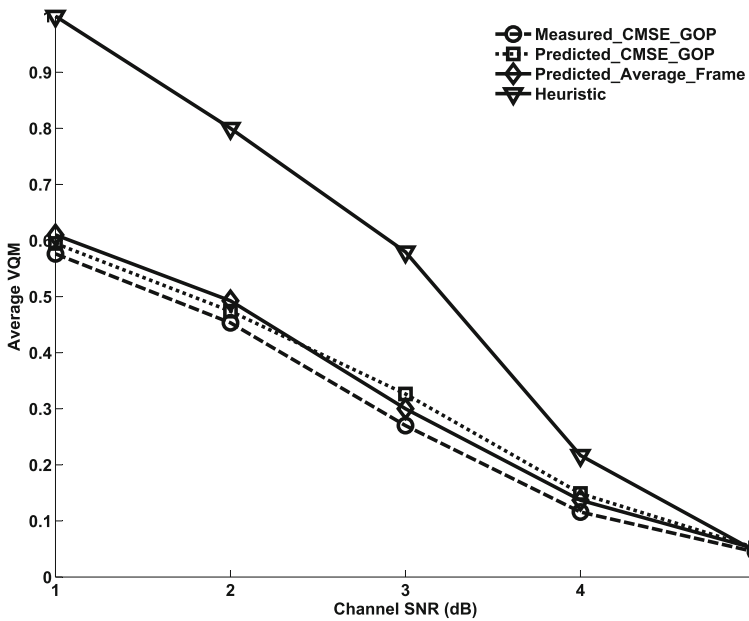
6 Frame-level slice priority assignment schemes

The GOP-based slice priority assignment scheme Predicted_CMSE_GOP requires the predicted CMSE values of all the slices of a GOP before it can assign priority to them. This introduces the delay of at least one GOP time, which may not be acceptable in applications with stringent delay constraints, such as video conferencing. In this section, we present a frame-level slice priority assignment scheme, denoted Predicted_Average_Frame, which determines the priority of the slices of a frame based on their predicted CMSE, the frame type and location within the GOP. This enables "on-the-fly" slice prioritization, without requiring the slices of future frames in the GOP.

Using all the videos in our training set (shown in Table 1) encoded at different bit rates, we use the Measured_CMSE_GOP scheme to compute the average percentage of slices in each priority class contributed by each frame within a GOP for a specific GOP length and configuration. Our GOP structure (IDR B P ... B) has a length of 20 frames and consists of one IDR frame, nine P frames and 10 B frames. Since the B frames are not used as reference frames, we have combined them together to get one set of ratios for the priorities. Table 5 shows the details of the prioritization ratios at a frame level. For example, the ratios for an IDR frame are (53 % , 28 % , 14 % , 5 %). This means that 53 % (5 %) of the slices with the highest CMSE in every IDR frame are assigned to P1 (P4). These ratios will vary when different encoding configurations are used.
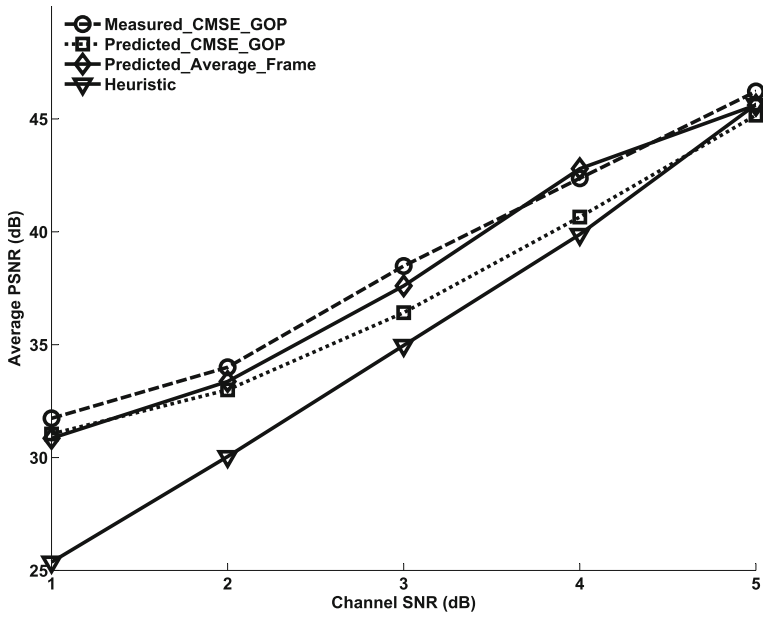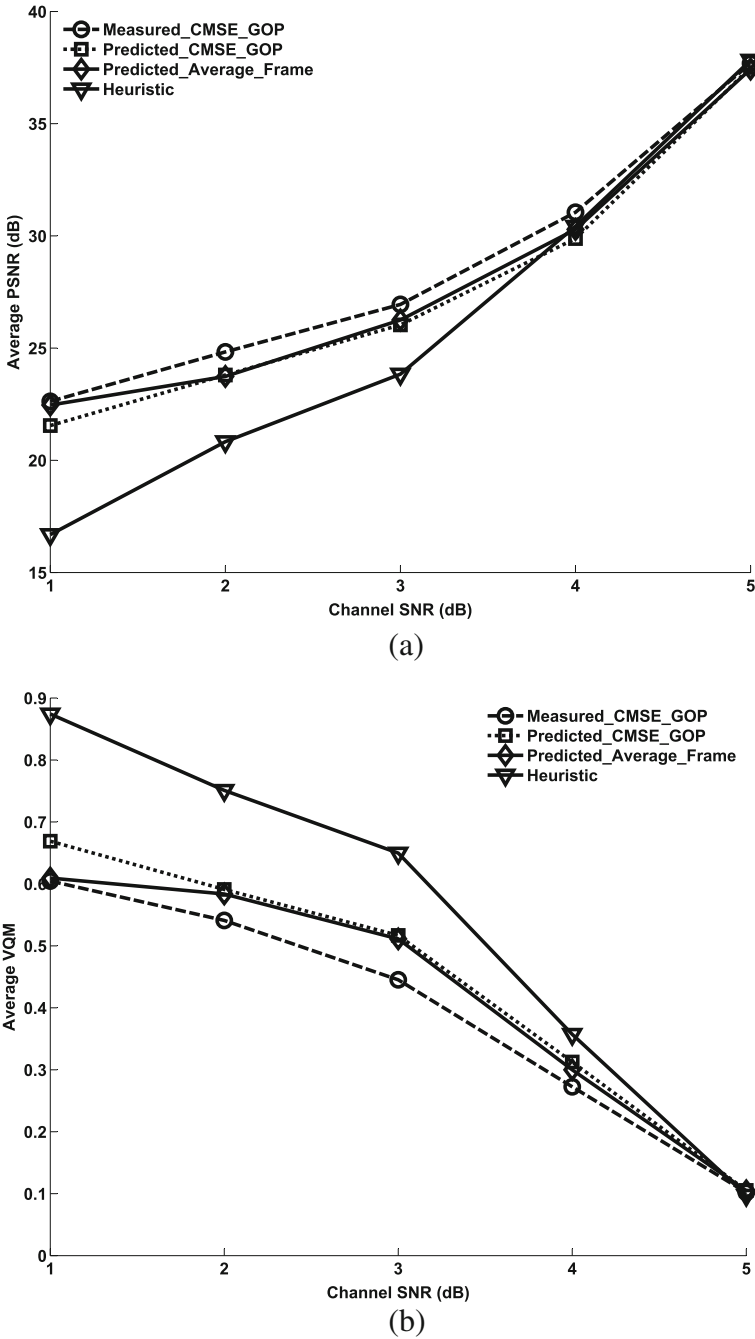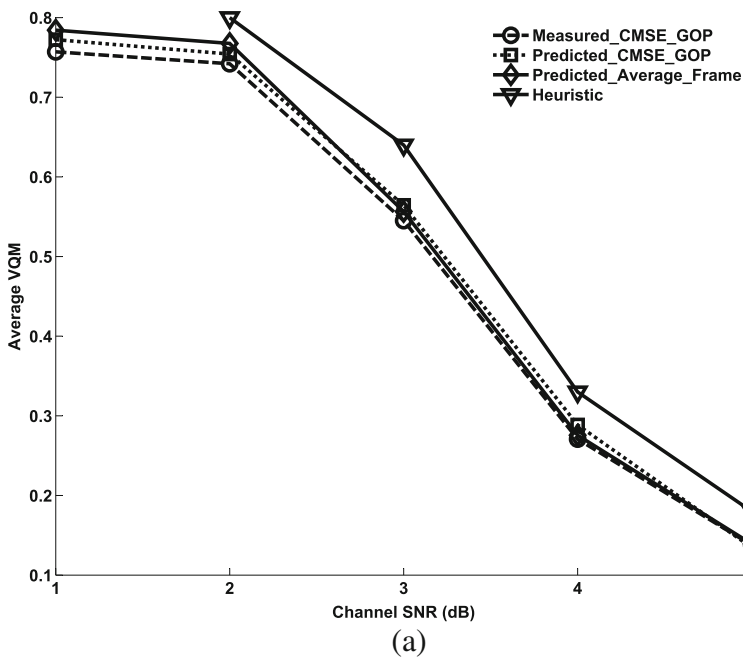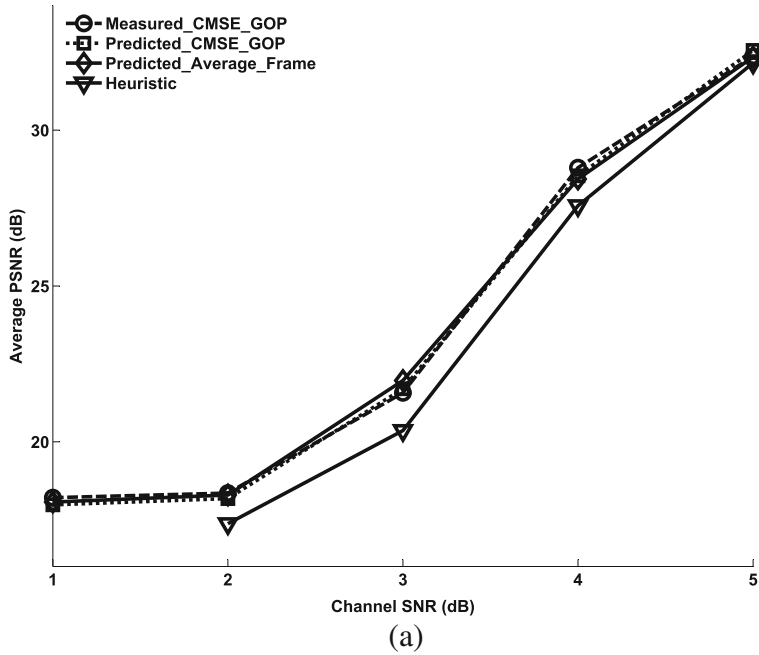
(a)

(b)

Fig. 5 (**a**) Average PSNR and (**b**) Average VQM performance of GOP and frame-level slice priority schemes over a 2Mbps AWGN channel for Akiyo encoded at 1Mbps

(a)

(b)

Fig. 6 (**a**) Average PSNR and (**b**) Average VQM performance of GOP and frame-level slice priority schemes over a 2Mbps AWGN channel for Foreman encoded at 1Mbps

Fig. 7 (**a**) Average PSNR and (**b**) Average VQM performance of GOP and frame-level slice priority schemes over a 2Mbps AWGN channel for Bus encoded at 1Mbps
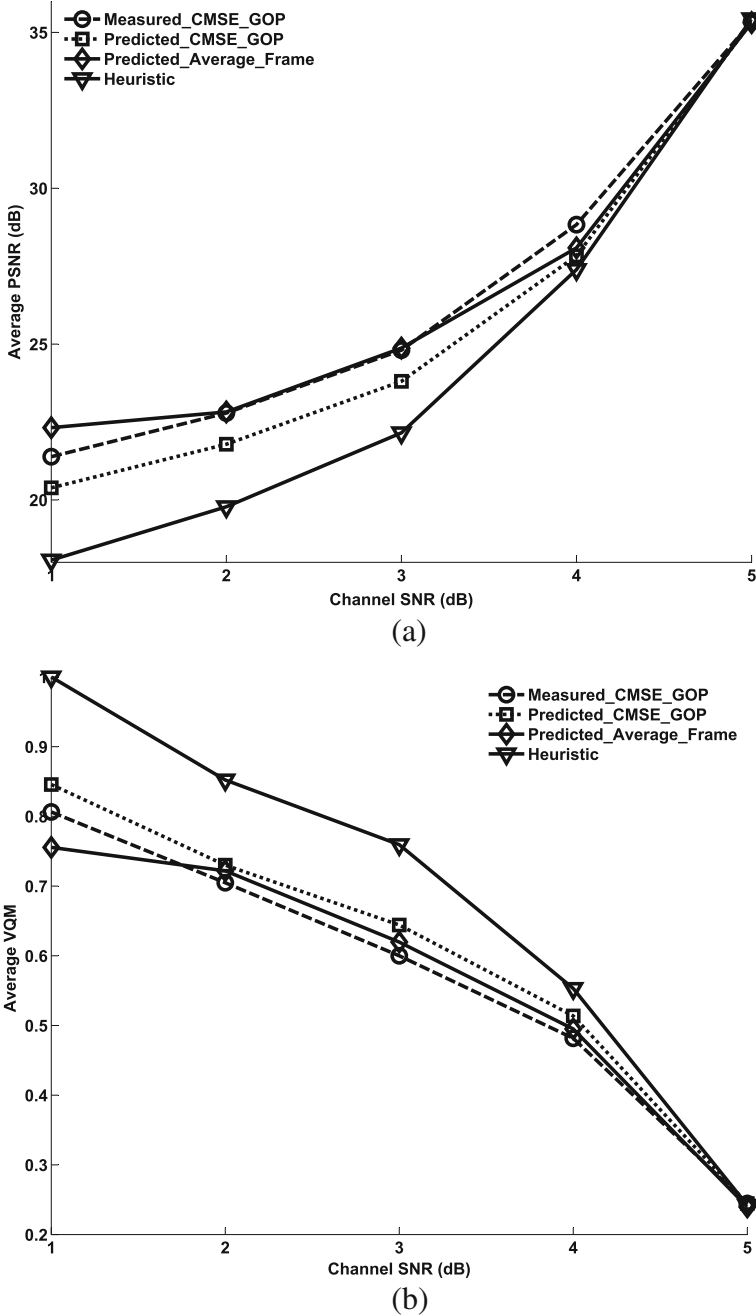
Fig. 8 (a) Average PSNR and (b) Average VQM performance of GOP and frame-level slice priority schemes over a 4Mbps AWGN channel for Table Tennis encoded at 2Mbps

Table 5 Percentage slices contributed to different priorities by each frame type and location within the gop

Frame Type	%Slices (P1/P2/P3/P4)
IDR	(53/28/14/5)
P Frame 1	(36/34/20/10)
P Frame 2	(34/37/19/10)
P Frame 3	(30/38/20/12)
P Frame 4	(25/35/27/13)
P Frame 5	(19/35/29/17)
P Frame 6	(14/31/33/22)
P Frame 7	(10/26/35/29)
P Frame 8	(5/20/36/39)
P Frame 9	(2/9/18/71)
B	(0/10/19/71)

We also examine the performance of a simple frame type and location based priority assignment scheme, called Heuristic, where all the IDR slices are assigned to the highest priority $P1$, all the slices of the first half of P frames of each GOP to $P2$, the P frame slices from the latter part of each GOP to $P3$, and all the B frame slices to the lowest priority $P4$, without considering their CMSE values. Note that several papers in the literature use this simple scheme [13, 15, 16, 44, 53]. We use a UEP scheme to study the performance of these frame-level slice prioritization schemes over AWGN channels for four test videos (slow motion Akiyo, medium motion Foreman, and high motion Bus) encoded at 1Mbps and a larger resolution video Table Tennis encoded at 2Mbps. Figures 5–8 show the average PSNR and VQM performance of the GOP-level Measured_CMSE_GOP and Predicted_CMSE_GOP schemes, the frame-level Predicted_Average_Frame scheme, and the Heuristic scheme. As discussed in the previous section, the performance of Predicted_CMSE_GOP scheme closely follows that of the Measured_CMSE_GOP. The frame-level priority assignment scheme, i.e., Predicted_Average_Frame achieves a performance which is very close to the GOP-based slice priority schemes while the Heuristic scheme performs significantly worse. Other videos tested produced similar results.

7 Application of slice prioritization to slice discard

In this section, we present an application of our slice prioritization schemes for slice discard in a network affected by congestion. We consider a simple topology, where two video users VS_1 and VS_2 are transmitting video data at $R1$ Kbps and $R2$ Kbps, respectively, to the destination D through an intermediate router. When there is no congestion, the outgoing rate R of the intermediate router equals $(R1 + R2)$ Kbps. We study the slice discard performance of three slice prioritization schemes Measured_CMSE_GOP, Predicted_CMSE_GOP, and Predicted_Average_Frame. In these schemes, the lowest priority slices are dropped from different frames of each GOP. We also study the performance of a Drop-Tail (DT) scheme where the slices are discarded from the tail end of each GOP.

We use two video sequences Foreman (VS_1) and Bus (VS_2) encoded at 1Mbps each. The outgoing bit rate of the intermediate router is 2Mbps. The error free PSNR values of the two sequences are 38.9dB and 33.4dB, respectively. Table 6 shows the video PSNR and VQM results obtained by discarding slices intelligently over the traditional DT scheme

Table 6 Performance of the slice discard schemes

| Slice Loss (%) | PSNR (dB) | | | | VQM | | | |
	10 %		30 %		10 %		30 %	
Slice Discard Scheme	VS_1	VS_2	VS_1	VS_2	VS_1	VS_2	VS_1	VS_2
Measured_CMSE_GOP	34.8	28.7	31.4	25.3	0.257	0.324	0.334	0.561
Predicted_CMSE_GOP	33.8	28.4	30.8	25.1	0.289	0.325	0.343	0.571
Predicted_Average_Frame	33.5	28.0	29.8	24.7	0.290	0.382	0.357	0.591
Droptail (DT)	31.2	26.0	27.8	22.9	0.335	0.501	0.381	0.750

when the outgoing rate R drops from 2Mbps to 1.8Mbps (10 % loss) and 1.4Mbps (30 % loss). We distribute the slice losses equally between the two users. Our slice prioritization schemes achieve at least 2dB PSNR and corresponding VQM gain over the DT scheme, as dropping consecutive slices of frame(s) from the tail of each GOP degrades the video quality significantly.

8 Conclusions

We presented a low-complexity and low-delay GLM to predict the CMSE of a slice loss using a combination of video factors that can be extracted while the current frame is being encoded, without using any information from the future frames of the GOP. We demonstrated the accuracy of the prediction model using cross-validation. The slices of each GOP were classified into four priorities using the predicted CMSE values, and their performance was close to that of the measured CMSE over AWGN channels. We extended our GOP based slice prioritization scheme to achieve frame-level prioritization for low-delay applications. We also analyzed the priority misclassifications and showed that second degree and third degree priority misclassifications were minimal. We also demonstrated an application of our slice prioritization scheme towards slice discard in a congested network.

References

1. Akaike H (1974) A new look at the statistical model identification. IEEE Trans Autom Control 19(6):716–723
2. Anderson TW, Darling DA (1954) A test of goodness of fit. J Am Stat Assoc 49(268):765–769
3. Babich F, D'orlando M, Vatta F (2008) Video quality estimation in wireless IP networks: algorithms and applications. ACM Trans Multimedia Comp Commun A 4(1):1–18
4. Baccaglini E, Tillo T, Olmo G (2008) Slice sorting for unequal loss protection of video streams. IEEE Signal Process Lett 15:581–584
5. Bandyopadhyay SK, Wu Z, Pandit P, Boyce JM (2006) An error concealment scheme for entire frame losses for H.264/AVC. In: IEEE International Sarnoff Symposium on Circuits and Systems, pp 1–4
6. Breiman L (2001) Random Forests. Mach Learn 45(1)
7. Breiman L, Friedman J, Stone CJ, Olshen RA (1984) Classification and Regression Trees. Chapman and Hall

8. Burnham KP, Anderson DR (1998) Model Selection and multimodel inference: A practical information-theoritical approach. Springer, Verlag
9. Chakareski J, Apostolopoulos J, Wee S, Tan WT, Girod B (2005) Rate-distortion hint tracks for adaptive video streaming. IEEE Trans Circ Syst Video Technol 5(10):1257–1269
10. Chakareski J, Frossard P (2005) Rate-distortion optimized bandwidth adaptation for distributed media delivery. In: IEEE ICME, pp 763–766
11. Chakareski J, Frossard P (2006) Rate-distortion optimized distributed packet scheduling of multiple video streams over shared communication resources. IEEE Trans Multimedia 8(2):207–218
12. Chartrand R, Yin W (2008) Iteratively reweighted algorithms for compressive sensing. In: IEEE ICASP, pp 3869–3872
13. Chen WT, Lin TC, Chen JC (2008) Dynamic packet selection for H.264 video streaming over IEEE 802.11e WLANs. In: IEEE WCNC, pp 3133–3138
14. Chikkerur S, Sundaram V, Reisslein M, Karam LJ (2011) Objective video quality assessment methods: A classification, review, and performance comparison. IEEE Trans Broadcast 57(2):165-182
15. Feamster N, Balakrishnan H (2002) Packet loss recovery for streaming video. In: Proc. Int. Packet Video Workshop, pp 1–11. Pittsburg, PA
16. Fiandrotti A, Gallucci D, Masala E, Magli E (2008) Traffic prioritization of H.264/SVC video over 802.11e ad hoc wireless networks. In: IEEE Computer Communications and Networks, pp 1–5. St. Thomas, US Virgin Islands
17. Hemami SS, Reibman AR (2010) No-reference image and video quality estimation: applications and human-motivated design. Elsevier J Signal Process Image Commun 25:469-481
18. Ichida K, Fujii Y (1979) An interval arithmetic method for global optimization. Springer-Verlag J Comput 23:85–97
19. Kanumuri S, Cosman P, Reibman AR, Vaishampayan VA (2006) Modeling packet-loss visibility in MPEG-2 video. IEEE Trans Multimedia 8(2):341-355
20. Kambhatla KKR, Kumar S, Cosman PC (2012) H.264/AVC video packet aggregation and unequal error protection for noisy channels. In: ICIP, pp 1649–1652
21. Kambhatla KKR, Kumar S, Paluri S, Cosman PC (2012) Wireless H.264 video quality enhancement through optimal prioritized packet fragmentation. IEEE Trans Multimedia 14(5):1480–1495
22. Kumar S, Xu L, Mandal MK, Panchanathan S (2006) Error resiliency schemes in H.264/AVC standard. Elsevier J Vis Commun Image Represent, Spec issue Emerg H.264/AVC Video Coding Stand 17(2):183–185
23. Li F, Liu G (2009) Compressed-domain-based transmission distortion modeling for precoded H.264/AVC video. IEEE Trans Circ Syst Video Technol 19(12):1908-1914
24. Li Z, Chakareski J, Niu X, Zhang Y, Gu W (2009) Modeling and analysis of distortion caused by markov-model burst packet losses in video transmission. IEEE Trans Circ Sys Video Technol 19(7):917–931
25. Liang YJ, Apostolopoulos JG, Girod B (2008) Analysis of packet loss for compressed video: Effect of burst losses and correlation between error frames. IEEE Trans Circ Syst Video Technol 18(7):861-874
26. Lin TL, Kanumuri S, Zhi Y, Poole D, Cosman P, Reibman AR (2010) A versatile model for packet loss visibility and its application to packet prioritization. IEEE Trans Image Process 19(3):722-735
27. Liu H, Xu H, Zhao S (2012) Consistent-degradation macroblock grouping for parallel video streams over DiffServ networks. Elsevier J Comput Commun 35:151-158
28. Lottermann C, Steinbach E (2014) Modeling the bit rate of H.264/AVC video encoding as a function of quantization parameter, frame rate and GOP characteristics. In: IEEE ICMEW, pp 1–6
29. Martin JCD, Quaglia D (2001) Distortion-based packet marking for MPEG video transmission over DiffServ networks. In: ICME, pp 399–402
30. Masala E, Martin JCD (2003) Analysis-by-synthesis distortion computation for rate-distortion optimized multimedia streaming. In: IEEE Int Conf Multimedia and Expo 3:345–348
31. McCullagh P, Nelder J (1983) Generalized Linear Models Chapman and Hall
32. Miguel V, Cabrera J, Jaureguizar F, Garcia N (2011) High-definition video distribution in 802.11g home wireless networks. In: IEEE ICCE, pp 211–212
33. Mullet GM (1976) Why regression coefficient have the wrong sign. J Qual Technol 8(3)
34. Paluri S, Kambhatla KKR, Kumar S, Bailey B, Cosman PC (2012) Predicting slice loss distortion in H.264/AVC video for low complexity data prioritization. In: ICIP, pp 689–692
35. Perez P, Garcia N (2011) Video prioritization for unequal error protection. In: IEEE ICCE, pp 315–316
36. Reibman AR, Vaishampayan VA, Sermadevi Y (2004) Quality monitoring of video over a packet network. IEEE Trans Multimedia 6(2):327-334
37. Reibman AR, Poole D (2007) Characterizing packet-loss impairments in compressed video. In IEEE ICIP 5:77–80
38. Schier M, Welzl M (2012) Optimizing selective ARQ for H.264 live streaming: a novel method for predicting loss-impact in real time. IEEE Trans Multimedia 14(2):415-430

39. Schmidt C, Rose K (2007) First-order distortion estimation for efficient video streaming at moderate to high packet loss rates. In: Int. Packet Video Workshop, Lausanne, Switzerland, pp 318–325
40. Seshadrinathan K, Soundararajan R, Bovik AC, Cormack LK (2010) Study of subjective and objective quality assessment of video. IEEE Trans Broadcast 19(6):1427-1441
41. Shen Y, Cosman PC, Milstein LB (2006) Video coding with fixed-length packetization for a tandem channel. IEEE Trans Image Process 15(2):273–288
42. Srinivasan SK, Vahabzadeh-Hagh J, Reisslein M (2010) The effects of priority levels and buffering on the statistical multiplexing of single-layer H.264/AVC and SVC encoded video streams. IEEE Trans Boradcast 56(3):281-287
43. Stockhammer T, Hannuksela MM, Wiegand T (2003) H.264/AVC in wireless environments. IEEE Trans Circ Syst Video Technol 13(7):657–673
44. Talari A, Rahnavard N (2009) Unequal error protection rateless coding for efficient MPEG video transmission. In: IEEE Military Communications Conference, pp 1–7
45. The R Project for Statistical Computing.http://www.r-project.org/
46. Venables W, Ripley B (2010) Modern Applied Statistics with S (Statistics and Computing). Springer
47. Vito FD, Farinetti L, Martin JCD (2002) Perceptual classification of MPEG video for differentiated-services communications. In: IEEE ICME, vol 1, pp 141–144
48. Wang Y, Wu Z, Boyce JM (2006) Modeling of transmission-loss-induced distortion in decoded video. IEEE Trans Circ Syst Video Technol 16(6):716-732
49. Wiegand T, Sullivan GJ, Bjntegaard G, Luthra A (2003) Overview of the H.264/AVC video coding standard. IEEE Trans. Circuits Syst. IEEE Trans Circ Syst Video Technol 13(7):560-576
50. Wu Z, Boyce JM (2006) An error concealment scheme for entire frame losses based on H.264/AVC. In: IEEE ISCAS, pp 4463–4466
51. Zhang F, Steinbach E, Zhang P (2014) MDVQM: A novel multidimensional no-reference video quality metric for video transcoding. Elsevier, J Vis Commun Image R 25:542-554
52. Zhang R, Regunathan SL, Rose K (2000) Video coding with optimal inter/intra-mode switching for packet loss resilience. IEEE J Sel Areas Commun 18(6):966-976
53. Zhang W, Zheng Q, Lian Y (2009) Tree-aware selective frame discard for P2P IPTV system on set-top boxes. IEEE Trans Conum Electron 55(4):982-1987
54. Zhang X, Peng XH, Wu D, Porter T, Haywood R (2009) A hierarchical unequal packet loss protection scheme for robust H.264/AVC transmission. In: IEEE CCNC, pp 1–5

Seethal Paluri received her B.S. in Electrical and Computer Engineering from University of the West Indies, Trinidad and Tobago in 2005 and her M.S. degree in Electrical Engineering from University of Southern California, Los Angeles in 2008. She is Ph.D. candidate in Computational Sciences with major in Mathematics and minor in Electrical Engineering, jointly awarded by San Diego State University and Claremont Graduate University. Her research interests are in the areas of error resilient image and video compression, error concealment techniques, and region of interest (ROI) based coding. She also works on developing statistical modeling techniques for video quality prediction and prioritization for reliable real-time Quality of Service (QoS) aware streaming over wireless networks. She is also a member of Scholars without Borders.

Kashyap K. R. Kambhatla received his B.Tech. degree in Electronics and Communication Engineering from Jawaharlal Nehru Technological University, Hyderabad, India in 2004 and the M.S. degree in Electrical and Computer Engineering from Clarkson University, Potsdam, NY, in 2006. He received his Ph.D. with honors in Electrical and Computer Engineering jointly awarded by the Jacobs School of Engineering, University of California San Diego and the College of Engineering, San Diego State University. During the summer of 2006, he worked on video telephony in 3G EV-DO Rev. A and B cellular networks as a research intern in the Next-Gen Wireless Technology Group at Sprint Advanced Technology Labs, Burlingame, CA. His research interest is in the area of real-time cross-layer quality-ofservice (QoS) aware video streaming over ad-hoc wireless communication networks through packet size and physical layer FEC code rate adaptation. He also works on network optimization problems and developing efficient medium access control protocols for reliable video transmission. He is also a member of Phi Kappa Phi and Scholars without Borders.

Barbara A. Bailey obtained her B.S in Mathematics and Physical Education from Springfield College, Massachusetts in 1981, and her M.S in Mathematics from University of Illinois, Champaign-Urbana, and her Ph.D. in Biomathematics from North Carolina State University, Raleigh. She served as a Postdoctoral Research Fellow at National Center for Atmospheric Research NCAR, Boulder CO during 1996-1998. She joined the University of Illinois at Urbana-Champaign as an Assistant Professor of Statistics during 1998-2005. She was a Visiting Assistant Professor with the Department of Mathematics at the University of Colorado at Denver and Health Sciences Center. She is currently an Associate Professor in the Department of Mathematics and Statistics at San Diego State University. Her research interests are in the area of temporal and spatial modeling of ecological and environmental processes using nonlinear and multivariate statistical analysis.

Pamela C. Cosman (S'88-M'93-SM'00-F'08) obtained her B.S. with Honor in Electrical Engineering from the California Institute of Technology in 1987, and her M.S. and Ph.D. in Electrical Engineering from Stanford University in 1989 and 1993, respectively. She was an NSF Postdoctoral Fellow and a Visiting Professor at the University of Minnesota during 1993-1995. In 1995, she joined the faculty of the department of Electrical and Computer Engineering at the University of California, San Diego, where she is currently a Professor and Vice Chair. Her research interests are in the areas of image and video compression and processing, and wireless communications. Dr. Cosman is the recipient of a Career Award from the National Science Foundation, a Powell Faculty Fellowship, and a Globecom 2008 Best Paper Award. She was a guest editor of the June 2000 special issue of the IEEE Journal on Selected Areas in Communications on "Error-resilient image and video coding," and was the Technical Program Chair of the 1998 Information Theory Workshop in San Diego. She was an associate editor of the IEEE Communications Letters (1998- 2001) and IEEE Signal Processing Letters (2001-2005). She was the Editor-in-Chief (2006-2009) as well as a Senior Editor (2003-2005, 2010-present) of the IEEE Journal on Selected Areas in Communications. She is a member of Tau Beta Pi and Sigma Xi.

John D. Matyjas received his Ph.D. in electrical engineering from State University of New York at Buffalo in 2004. Currently, he is serving as the Connectivity & Dissemination Core Technical Competency Lead at the Air Force Research Laboratory (AFRL) in Rome, NY. His research interests include dynamic multiple-access communications and networking, spectrum mutability, statistical signal processing and optimization, and neural networks. He serves on the IEEE Transactions on Wireless Communications Editorial Advisory Board. He is also the recipient of the 2012 IEEE R1 Technology Innovation Award, 2012 AFRL Harry Davis Award for "Excellence in Basic Research," and the 2010 IEEE Int'l Communications Conf. Best Paper Award. He is an IEEE Senior Member, chair of the IEEE Mohawk Valley Signal Processing Society, and member of Tau Beta Pi and Eta Kappa Nu.

Sunil Kumar obtained his B.E. in Electronics Engineering from S.V. National Institute of Technology in 1988, and his M.E. and Ph.D. in Electrical and Electronics Engineering from Birla Institute of Technology and Science, Pilani, India, in 1992 and 1997, respectively. From 1997-2002, he was a postdoctoral fellow and adjunct faculty in Electrical Engineering Systems department at the University of Southern California, Los Angeles, CA. From 2002 to 2006, he was an Assistant Professor in Electrical and Computer Engineering at Clarkson University, Potsdam, NY. Since 2006, he is an Associate Professor and Thomas G. Pine Faculty Fellow in the Electrical and Computer Engineering Department at the San Diego State University, San Diego, CA. His research interests include QoS-aware cross layer wireless network protocols and robust multimedia compression techniques, including H.264. He has published more than 100 research articles in refereed journals and conferences, including two book/book chapters. His research has been supported by NSF, U.S. Air Force Research Lab, Dept. of Energy, California Energy Commission, and industry. He was a guest editor of the April 2006 special issue of Elsevier Journal of Visual Communication and Image Representation on "Emerging H.264/AVC Video Coding Standard". He has organized workshops and technical sessions in several international conferences.