Transmission of video telephony images over wireless channels ☆

Hang Liu and Magda El Zarki

Video Processing and Telecommunications Laboratory, Department of Electrical Engineering, University of Pennsylvania, Philadelphia, PA 19104, USA

Abstract. In this paper, the effects of digital transmission errors on H.263 codecs are analyzed and the transmission of H.263 coded video over a TDMA radio link is investigated. The impact of channel coding and interleaving on video transmission quality is simulated for different channel conditions. Fading on radio channels causes significant transmission errors and H.263 coded bit streams are very vulnerable to errors. Powerful Forward Error Correction (FEC) codes are therefore necessary to protect the data so that it can be successfully transmitted at acceptable signal power levels. FEC, however, imposes a high bandwidth overhead. In order to make best use of the available channel bandwidth and to alleviate the overall impact of errors on the video sequence, a two-layer data partitioning and unequal error protection scheme based on H.263 is also studied. The scheme can tolerate more transmission errors and leads to more graceful degradation in quality when the channel SNR decreases. In lossy environments, it can improve the video transmission quality at no extra bandwidth cost.

1. Introduction

The advances in low bitrate video coding technology have led to the possibility of delivering video services to users through band limited wireless networks. Both ITU-T/SG15 and ISO-MPEG4 are working to set standards for very low bit rate video coding. Recently, ITU-T/SG15 finished the first draft recommendation of H.263 which targets the transmission of video telephony through the Public Switched Telephone Network (PSTN) at data rates less than 64 kbit/s [1]. The experts' group is starting to adapt H.263 for wireless applications because the low bit rate makes it best suitable for band limited wireless networks. A couple of proposals are currently being evaluated [2].

H.263 coded bit streams are very vulnerable to errors and require high channel reliability. Radio channels on the other hand are error prone. Fading on radio channels causes significant transmission errors. Therefore, powerful Forward Error Correction (FEC) codes are necessary to protect the data so that it can be successfully transmitted at acceptable signal power levels. However, the extra bandwidth for FEC overhead is also critical in wireless networks because the bandwidth in the wireless domain is much more limited than in wireline networks. The relationship between error control and video transmission quality must therefore be investigated in order to get an optimal trade-off.

Layered (or classified) source coding and unequal error protection have received a lot of attention in recent years [2–5]. They are two techniques that one can implement to alleviate the impact of errors on the video sequence and make best use of the channel bandwidth. There are two main issues related to layered source coding and unequal error protection. One is how to partition the data into different priority layers. The other is how to choose the appropriate error protection schemes for the different priority layers. In order to obtain optimal transmission quality, these two issues have to be jointly considered in conjunction with the channel characteristics.

In this paper, the effects of digital transmission errors on H.263 codecs are analyzed, the transport of H.263 video over a TDMA radio link is systematically investigated, and the impact of FEC and interleaving on video quality is evaluated. Numerical results for the performance of the video transmission under different channel coding and interleaving strategies are presented for various channel conditions. In order to make best use of the available channel bandwidth and to alleviate the overall impact of errors on the video sequence, a two-layer source coding and unequal error protection scheme based on H.263 is also studied. The scheme can tolerate more transmission errors and leads to more graceful degradation in quality when the channel SNR decreases. In lossy environments, it improves video transmission quality at no extra bandwidth cost.

This paper is organized as follows: Section 2 outlines the H.263 coding structure and analyzes the effect of transmission errors on the H.263 video. Based on this analysis, a two-layer data partitioning and unequal error protection scheme is presented. In section 3, we briefly describe the FEC and channel model used in this work. In section 4, the transmission of single priority layer and two priority layer H.263 video with different FEC, interleaving and data partitioning strategies has been simu-

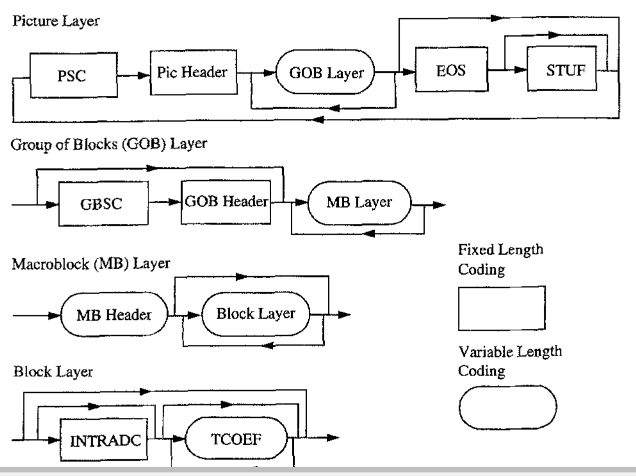
2. H.263 video coder and effect of transmission errors

2.1. Outline of H.263 video coding structure

H.263 coded data is arranged in a hierarchical structure with four syntax layers as shown in Fig. 1 [1]. From top to bottom they are: Picture, Group of Blocks (GOB), Macroblock (MB) and Block. Data for each picture is composed of a picture header followed by data for GOBs, eventually followed by an optional end-of-sequence code (EOS) and stuffing bits (STUF). The picture header contains a picture start code (PSC), temporal reference (TR), type information (PTYPE), quantizer information (PQUANT), continuous presence multipoint (CPM), picture logical channel indicator (PLCI) if CPM mode is indicated by external means, extra insertion information (PEI), and spare information (PSPARE). Temporal reference for B-frames (TRB) and quantization information for B-pictures (DBQUANT) are also included in the picture headers if the optional PB-frames mode is used.

Each picture is divided into GOBs. Data for each GOB consists of a GOB header followed by data for macroblocks. Each GOB comprises one macroblock row for sub-QCIF, QCIF, and CIF format pictures, and two macroblock rows for 4CIF and four macroblock rows for 16CIF. For the first GOB, the GOB header is empty and therefore only macroblock data is present. For all other GOBs, it is up to the encoder whether or not the GOB header will be left empty or not. If the GOB header is not empty, it will include GOB start code (GBSC), group number (GN), GOB logical channel indicator (GLCI) if CPM mode is indicated by external means, GOB frame ID (GFID), and quantizer information (GQUANT).

A macroblock contains four 8 by 8 blocks of luminance data (Y) and the two corresponding 8 by 8 blocks of chrominance data (one of each of the blue chrominance Cb and red chrominance Cr). Data for each macroblock consists of a macroblock header followed by data for the blocks. The Macroblock header includes coded macroblock indication (COD), macroblock type & coded block pattern for chrominance (MCBPC), macroblock mode for B-blocks (MODB), coded block pattern for B-blocks (CBPB), coded block pattern for luminance (CBPY), quantizer information (DQUANT), motion vector data (MVD), motion vector data for optional advanced prediction mode (MVD2-4), and motion vector data for B-macroblocks (MVDB). Some fields may not be present in the MB header depending on the picture header and the other MB header fields. A block comprises an 8-row by 8-column matrix of luminance or chrominance data samples (Y, Cb or Cr). The DC coefficient for INTRA block (INTRADC) is present for every block of the intra coded macroblocks. Transform coefficients (TCOFFs) represent the other DCT coefficients coded by run-length and variable length coding (VLC) that are present if indicated by MCBPC, CBPB, or CBPY.



2.2. Effect of transmission errors and data partitioning

The impact of a single bit in error will depend on which bit is hit. A single bit error in the DCT coefficients will at least damage one block as the DCT coefficients of each block are coded using run-length and variable length coding. If loss of synchronization occurs, all the subsequent blocks in the GOB may be destroyed. As motion vectors are differentially encoded for the inter-coded macroblocks of the same GOB, then a bit error in a motion vector may result in the corruption of this macroblock and the following predicted macroblocks. In addition, because VLC is also employed in the macroblock headers, synchronization will probably be lost when a bit error occurs. Therefore, the worst case of a bit error in the macroblock headers is the loss of the complete GOB. The start codes in the picture and GOB headers provide the synchronization in the spatial domain and stop error propagation. The effect of transmission errors is confined to a GOB alone, i.e., one or more bit errors in a GOB does not affect other GOBs. Only if a bit error occurs in the control information symbols of the picture headers, may it seriously impair the total frame.

In the temporal domain, predictive coded pictures (P-pictures) are coded using motion compensated prediction from a past intra or predictive coded picture and are generally used as a reference for further prediction. The error will propagate among the P-pictures until the next intra coded picture (I-picture) which is coded using information only from itself, thereby providing error resilience in the temporal domain. Furthermore, as the higher frequencies are visually less important than the lower ones, a bit error occurring in the higher frequency DCT coefficients will have very little impact on video links; both random and bursty errors generated by noise and fading exist when signals are transmitted. Powerful FEC is required in order to combat the errors. However, FEC adds a lot of overhead to the system which at worst could render the delivery of acceptable quality video impossible because the required data rate exceeds the available channel capacity. A natural approach is to rearrange the coded video information such that the important information is better protected and more likely to be received correctly. The unimportant information is less protected in order to reduce bandwidth overhead.

Based upon the above observations, a simple approach is to partition the H.263 coded data into two priority layers (classes) [5]. The high priority layer (partition 0 or base layer) consists of the important data, which contains the control information, motion vectors and maybe lower order DCT coefficients, while the low priority layer (partition 1 or enhancement layer) contains the higher frequency DCT coefficients. The GOB start codes and GOB numbers may or may not be redundantly copied in the low priority layer, depending on the ratio of additional overhead to significance of synchronization and error recovery in the low priority layer. In general, it is better to copy them in the low priority layer if a lot of data is partitioned into this layer. A priority breakpoint in the picture headers indicates what elements are to be included in the high priority layer. The remainder of the bitstream is to be placed in the low priority layer. This is similar to the MPEG-2 data partitioning syntax [6] except that one more bit is used in the picture header to indicate whether or not the GOB start codes and GOB numbers are copied in the low priority layer. Fig. 2 shows an example of how the decoder switches between the partitions when a bit stream with two partitions is decoded. The controller can change the priority breakpoint so that the I-pictures have more data in the high priority layer than the P-pictures.

The two priority layers from the video source coder employ different FEC codes, reflecting the importance of the information. After channel coding, a sequential multiplexing scheme can be used to multiplex the encoded data from the different priority layers [2]. The two layers of data are serially interleaved with each other in the multiplexed stream. Each layer is preceded by a layer header in order to allow demultiplexing at the receiver side. For each frame, the data of the high priority layer with the prefixed layer header is transmitted first, followed by the low priority layer header and data. Interleaving can then be used on the multiplexed stream to randomize the bursty errors. Finally, the coded data is placed in the transmitter buffer for transmission.

If block FEC codes are employed and the priority layer length from a coded frame is not exactly a multiple of the FEC block code length (i.e., when the coded data from one priority layer of a frame is divided into blocks, the last block does not have enough bits), the last block can include some bits from the same priority layer of the next transmitted frame. In our simulation, block codes with 127 bits long are employed. The priority layer header is of fixed length and includes three fields. The first 1 bit field indicates the layer type. The second field is 7 bits long and indicates how many bits in the last block come from the next frame. The third field is 8 bits long and indicates the length of the data field following the header and belonging to this priority layer in units of 127 bits. The layer length is variable depending on the coded frames. Generally, 8 bits are enough to represent the layer length of encoded QCIF format video sequences. In case the layer length is longer than what the length field can represent, the field is extended by another 8 bits. Then the value of the first 8 bit length field is $2^8 - 1$ and the value of the last 8 bit length field is the difference of the layer length and $2^8 - 1$. If convolution codes are used, the layer header only needs two fields, one is for the layer type and the other for the layer length in bits because the data for the layer can exactly come from one frame [2].

Note that the picture start codes are not necessary after channel coding because the layer header information has indicated where a picture will start. Therefore, the picture start code will be removed before channel coding at the transmitter and added after channel decoding at the receiver. The layer headers are well protected for transmission so that the errors in the layer headers are negligible. The bit stream is corrupted during the transmission over the wireless channel. At the receiver, the corrupted bit stream is deinterleaved, and the two layers are reconstituted by the demultiplexer, then channel decoding and source decoding are performed. Good error tolerance can be achieved if the high priority data is well protected. This scheme has the advantage of minimum complexity and a bit rate efficiency very close to the single layer encoder. Furthermore, the high priority layer can carry enough information to produce an accep-

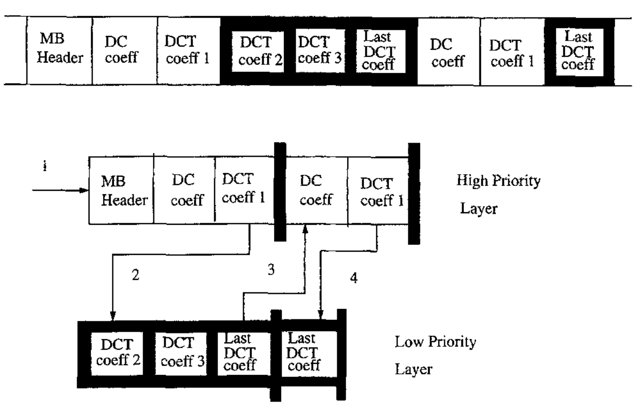


Fig. 2. A segment from a bit stream with two partitions. The priority breakpoint is set so that one (last, run, level) coefficient event is

3. Channel error protection

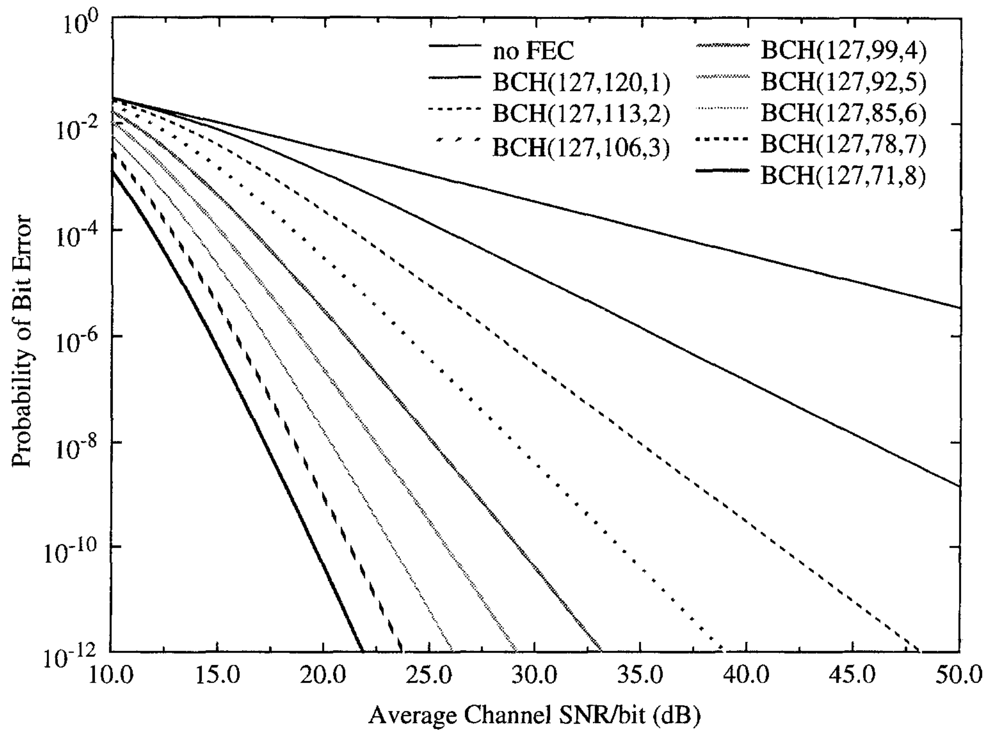
3.1. Forward error correction codes

Real-time video services require high reliability, limited time delays, and reasonably high transmission rates. The selection of FEC codes needs to take into consideration of several factors in order to meet the video transmission requirements: (1) The capability of the FEC: the objective is to improve the end-to-end bit error rate as much as possible. It should be noted however that the capability of the FEC codes strongly depends on the channel characteristics and error patterns. (2) The overhead or code rate: the FEC codes should add as little as possible overhead and maximize the code rate. (3) The block size for each code: most codes become more robust with larger blocks because more redundancy is available for a given code rate. However, the increase in block size causes additional delay. (4) The complexity: it should be simple so that the design/implementation cost can be minimized. Interleaving spreads the bursty errors due to Rayleigh fading into random errors required by most FEC codes. It should be noted that a single bit error could have the same impact on the reconstructed pictures as if all the bits of the GOB are in error because of error propagation in a GOB in the compressed video bit stream. The spreading of errors, via interleaving, in a compressed video bit stream, may damage more GOBs thereby having a negative effect on the overall performance unless the used channel coding is strong enough to correct those erroneous bits. The same observation can be made for MPEG coded video [7]. Furthermore, interleaving results in delay. Some of the above criteria in the selection of FEC codes and interleaving degrees are contradictory. The overriding issue is that the design of an effective error control scheme should consider all of the above factors to achieve the best engineering trade-off.

BCH codes provide a good trade-off in terms of error correction capability versus complexity. They can effectively correct random errors and have also been successfully applied in bursty error environments when combined with interleaving. We use BCH codes in our simulations. Fig. 3 depicts the BER performance versus channel SNR over a Rayleigh fading channel for BCH(127, 120, 1), BCH(127, 113, 2), BCH(127, 106, 3), BCH(127, 99, 4), BCH(127, 92, 5), BCH(127, 85, 6), BCH(127, 78, 7) and BCH(127, 71, 8) codes. We assume here that the modulation scheme is $\pi/4$-QPSK and interleaving is used to randomize the error bursts. The associated overheads are 5.5%, 11%, 16.5%, 22%, 27.5%, 33%, 38.5%, and 44%, respectively.

3.2. Channel model

The channel and transmission model used in the simulation is based on the Personal Access Communications Services (PACS) system [8,9], which is one of the industry standard proposals for emerging Personal Communications Services (PCS) in the United States. PACS uses TDMA and FDD, with 8 slots per carrier and 2.5 ms frame duration. $\pi/4$-QPSK is chosen as the



modulation format because of its high spectrum and power efficiency. The channel bandwith is 300 kHz and the radio link rate is 384 kb/s with 32 kb/s basic data rate for each user. One time slot carries 120 bits, including 80 bits of user information. A base station can support multiple handsets simultaneously with a transmitter and a receiver. Furthermore, it is able to allocate several slots to a single call to provide higher data rates.

A fading simulator based on the above channel and transmission model is used to generate the error patterns of the channel [10]. We assume that the channel coherence bandwidth is much larger than the signal bandwidth and delay spread is not a serious problem. Otherwise, anti-intersymbol interference (ISI) measures such as adaptive equalizers should be used to alleviate ISI. The carrier frequency is 1.9 GHz which is in the frequency band of PCS.

4. Simulation results

4.1. Transmission of standard single layer H.263 video

Fig. 4 shows the block diagram of the wireless video transmission system under investigation. Source coding is performed first, then various FEC and interleaving error control schemes are used. To evaluate the effect of channel errors, error patterns generated by the fading simulator are added modulo-2 to the binary bit stream representing the coded output from the interleaver. The coded bit stream with errors is then reconstructed into a video sequence.

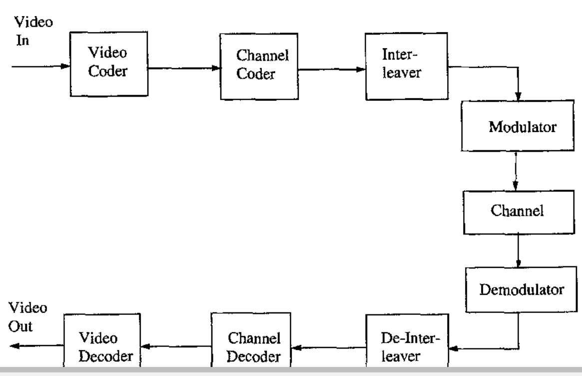
Simulation was carried out on several QCIF video sequences. The results of the "Mother and Daughter" sequence are reported in this paper, which contains typical video telephony-like images. The original YUV 4 : 1 : 1 video sequence is encoded with 15 frames per second, and an I frame is used every second (i.e. every 15 frames). Video encoding and decoding are performed with the modified Telenor R&D H.263 software and the optional PB-frames mode, unrestricted motion vector mode and advanced prediction mode are employed.

Peak Signal-to-Noise Ratio (PSNR) as a function of the average channel SNR is used as an objective measure of video quality for a given error control scheme. The overall bit rate including the video and FEC overhead is always 32 kb/s for all simulation scenarios. 200 encoded frames with different FEC and interleaving error protection schemes are transmitted at various average channel SNR conditions. In order to reduce the sensitivity of the encoded video to the temporal and spacial location of an error, each transmission was run 20 times using a different starting time. The average value over all the runs for each transmission is presented. For the two layer coded sequences, the sum of the video data in the high and low priority layers combined with the FEC used for each layer and the overhead due to data partitioning determines the overall data rate.

Fig. 5 depicts the average PSNR as a function of the average channel SNR using BCH codes with 127 bits block length and different error correction capabilities. The standard single layer H.263 coded bit streams are used and interleaving (INV) is performed over 20 127-bit-blocks. The vehicle speed is 2.5 mi/h. For each FEC case, the video quality rapidly degrades as the average channel SNR decreases below a threshold. This is because there is a dramatic drop in video quality once errors occur in the headers and motion vectors. The threshold of the curve occurs at a lower channel SNR for more powerful FEC codes. However, the PSNR is lower for stronger FEC codes at high channel SNR when there is no error, because higher overhead required by stronger FEC codes reduces the video source rate.

Fig. 6 shows the impact of interleaving on video quality. BCH(127, 113, 2) is used as error protection with different interleaving degrees and the vehicle speed is 2.5 mi/h. When the degree of interleaving increases, the error bursts are better randomized so that FEC codes, which are ideally suited to correct uncorrelated errors, can handle them better. On the other hand, as discussed in section 3, the spreading of errors in a compressed video bit stream may damage more GOBs if the channel coding can not correct these errors. This is why higher degrees of interleaving result in better PSNR performance when the average channel SNR is high and error rate is low. However, no interleaving is the best when the average channel SNR is very low and error rate is very high.

Interleaving also causes additional delay. As an example, we assume that the processing delay of the transmitter and the receiver is 50 ms which is the total time required to perform video encoding, channel encoding, modulation, demodulation, channel decoding, and video decoding, etc., and each protected PB frame is 2.54 kbit including FEC overhead. The transmission delay for each PB frame is 79 ms for a 32 kb/s channel. If interleaving is performed over a PB frame (20 127-bit-long blocks), the receiver can start processing the data only after all the interleaved bits are received. This means that



DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming inter-face) to developers that want to integrate case filings into their apps.

LAW FIRMS

Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS

Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS

Sync your system to PACER to automate legal marketing.

WHAT WILL YOU BUILD? | **sales@docketalarm.com** | **1-866-77-FASTCASE**

fastcase®
Smarter legal research.