System and model for real-time wireless video transmission

Dmitry Aleksandrov, Vladimir Bashun, Anton Sergeev, Andrey Turlikov
St. Petersburg State University of Airspace Instrumentation

Abstract

In this paper the problem of real time transmission of compressed video over a variable UWB channel is investigated. The main goal is to create joint compression and transmission scheme (jpeg progressive + ARQ protocol) that will provide the required level of PSNR and probability of lost rows in reconstructed frames. For this purpose an ARQ protocol for multi stream video delivery is developed. Also a special model is presented to find the optimal parameters (packets sizes, number of slices) of the joint coding scheme minimizing the number of dropped rows and maximizing PSNR of reconstructed frame. The suggested solution was developed for Cable Replacement program and can be used with existing UWB PHY/MAC. Nevertheless the developed methods of wireless video transmission are general and can work with other compression algorithms (J2k) and wireless systems (WiFi, WiMAX).

I. INTRODUCTION

Cable replacement program using wireless technologies, particularly Ultra-wide band (UWB), allows different devices to interoperate - without cables - as soon as they are in proximity. UWB [1] can be used to enable high-speed (up to 480 Mbps), high quality video retransmission inside a wireless personal area network (WPAN). The variety of devices within the entertainment cluster can be wirelessly connected: DVDs, HDTVs, STBs, PVRs, PDAs, laptobs, MP3 players and stereos, digital camcorders and digital cameras, and other CE devices found throughout the home. For example, UWB could connect a wall-mounted plasma display or HDTV to an STB or DVD player, without cables [8].

Transmitted video data should be efficiently compressed because of large sizes of multimedia data (1,5 Gbps for HDTV 1080i), constrained bandwidth, and fast change of channel conditions. Variable throughput, fluctuations of wireless channel and delay limits should be also taken into account.

The research shows [5] that lossless video coding algorithms (JPEG-LS, JPEG2000 lossless) do not provide necessary compression rate so only lossy coding schemes are considered below. In this paper a lightened tiny version of JPEG algorithm was used. But the described protocol/model can work with any progressive lossy coder, for example JPEG2000 [7]

The investigated joint compression and transmission scheme for real time video delivery over the wireless channel should satisfy the following requirements:

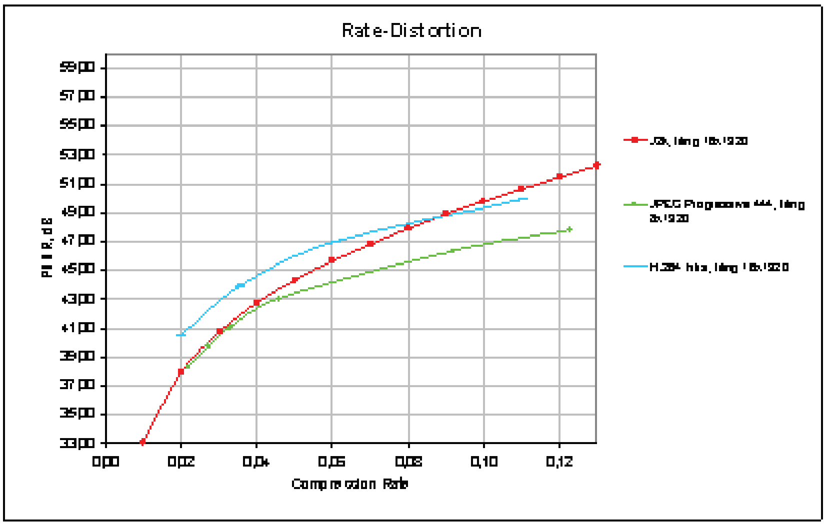

Fig. 1. Rate-Distortion curves for J2k, H.264 and JPEG Intra compression algorithms in tiling mode

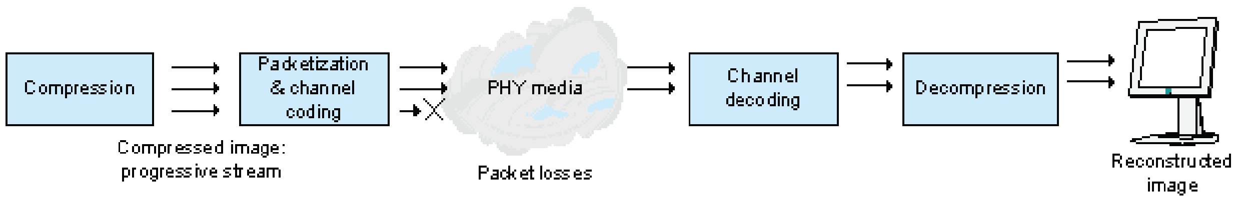

Fig. 2. Scheme of compression and transmission system for real-time video delivery

- Input: HDTV 1080i/60 8bit, 30 FPS
- Image quality: No visual difference between source and lossy compressed picture that stands for PSNR $>$ 37 dB
- Delay min for use in latency critical applications (gaming, conferences, video)
- Target Compression Ratio (CR): 10 - 13 times plus possibility of multilayered video streams with priorities
- Complexity and delay constraints. In this paper we consider the situation with a minimal buffer for storage of one block of encoded data (row of domains, tile etc.). Strong memory constraints (100 KBytes) and rigid delay limits do not allow to apply motion compensation/estimation methods or any other complex inter/intra coding algorithms.
- Work with existing UWB chips. Uncoded stream from PHY layer is not available

Here input frame is divided into non-overlapping successive tiles (8 rows of pixels, blocks of size 8x1920) that are compressed and transmitted independently. This requires only 50 Kbytes of memory on coder and decoder.

Brief structure of joint coding and transmission scheme is shown in Figure 2. JPEG (progressive mode) coding algorithm [6] is selected for frames compression due to very small complexity level [5] and good compression efficiency (see Fig. 1) and is not described here. It's also necessary to note that the task of video compression for transmission over the channel with constrained bandwidth is strongly connected with Rate-Distortion Optimization problem, but through scope constraint it's not examined here. So the main goal is to develop ARQ protocol for multi stream video delivery.

A similar work was done in [4]. It also considered a transmission of video stream with delay constraints over the erroneous wireless channels. However, a transmission of progressive stream was not fully exploited. Adaptive selection of quantizers was investigated to fit the time-variant channel for two situations only: either there is enough time to send whole block of data, or the whole block will be lost and therefore progressive coding and multi stream transmission was not used.

II. MAIN PART

A. Transmission protocol

The main goal of video transmission protocol is to transmit block of data with delay that does not exceed some fixed threshold, and combat errors that occur during transmission. Obvious way to deal with packets losses is to use error correction techniques. The basic error correcting methods that can be used is forward error correction (FEC) and automatic repeat request (ARQ). UWB supports both methods but in this paper, we consider usage of existing UWB hardware for transmission. This means, that physical and MAC layer of UWB can not be modified. In order to combat errors, we are able to use existing forward error correction techniques on UWB PHY layer, and UWB MAC [1] for erroneous packets retransmission. UWB MAC supports several types of acknowledgment policies (no acknowledgment, immediate acknowledgment, or acknowledgment for group of packets)

Packet loss as a result of errors and delays can be reduced, but it can not be eliminated completely. Increase of FEC rate or number of retransmissions is limited by necessity to transmit a certain number of compressed information (e.g. row of domains or frame) within a certain time interval. For real-time video increase of transmission delay is similar to packet loss (i.e. packets that arrive too late are considered to be lost). It means that packet can not be re-transmitted forever until it successfully received, and some packets are always lost. Moreover, increase of FEC and number of retransmissions leads to protocol overhead growth. This results in decreasing of throughput even in absence of errors.

The protocol considers existence of two streams with different priorities. In this case it is natural to put data of different priorities into different packets. High-priority data is put into the first packet, and low-priority data is placed into the second packet. First packet is send with immediate acknowledgment policy, and is resent until it would be received successfully. Only after that low priority data is sent.

B. Efficiency function

Given the set of quantizers (that determine average compression rate) and PHY transmission rate is fixed, it is necessary to select protocol parameters (packet sizes, number of retransmissions) in order to maximize the reconstructed image quality.

Traditional metric for estimation of end-to-end distortion image distortion is $PSNR$ [3],[4]. However, average$PSNR$ shouldn't be used as the only metric, because the human visual system perceives video in ways that are not easily reproduced in quantifiable metrics like PSNR. For real-time video over wireless transmission system there is common situation when there is no data at decoder (as a result of packet loss or high delays). We call such situation outage. If number of outage blocks is not very big, average $PSNR$ of video sequence will be reduced insignificantly. At the same time, frames with outage rows are notable by human visual system. Additional constraint on outage probability (P_{out}) is required. Outage probability allows estimating the percent of data blocks that are corrupted. The following criteria are used to optimize protocol parameters for wireless transmission:

- Maximize PSNR of the transmitted and reconstructed frame.
- Outage probability (probability of lost row) - the portion of transmitted rows that were lost due to errors during the transmission, should not exceed some threshold

It is clear that main optimization parameter is size of high-priority packet. It determines available number of retransmissions, and thus influence on the outage probability.

C. Optimization problem formulation

The main goal is to maximize average reconstructed image quality (PSNR). At the same time, the outage probability should not exceed certain limit. The optimization problem can be formulated in the following way:

$$\begin{cases} avPSNR_{movie} \to \max \\ P_{out} \leq P_{out}^{*} \end{cases},$$

$$avPSNR_{movie} = 10\log_{10}\frac{(2^n-1)^2}{MSE_{movie}},$$

$$MSE_{movie} = \frac{1}{N_{rows}}\sum_{i=1}^{N_{rows}} rowMSE_i,$$

$$rowMSE_i = \frac{1}{N_{pixels}}\sum_{j=1}^{N_{pixels}} e(\bar{x})^2$$

where $avPSNR_{movie}$ - average movie PSNR, $rowMSE_i$ - mean square error for row of domains, N_{rows}- number of domain rows in movie,N_{pixels}- number of pixels in row of domains, P_{out}^{*} - outage probability constraint

D. System model

In order to find optimal system parameters, we introduce several assumptions: **A1.** Progressive video codec is used, i.e. if decoder receives only part of data, it can reconstruct image of lower quality. Transmission protocol (described above) is fixed

A2. Bit error rate is fixed, and errors are uniformly distributed. It means that PER (Packet Error Rate) can be calculated for given BER (Bit Error Rate) and packet size. This assumption may be not quite realistic. Yet, UWB provide coding techniques (such as interleaving, band hopping, etc.) that make this assumption close to reality. Channel measurements with UWB devices also demonstrate this.

A3. For every block of data, we have distortion as a function of a number of compressed bytes (Rate-Distortion function). A rough approximation assumes linear dependence between size of row of domains and PSNR. It is assumed that there exists a maximum value of PSNR that can not be improved with further growth of source coding rate. Mean square error for outage block of data is taken equal to64^2.

A4. Distribution of sizes of compressed domain rows can be approximated theoretically by truncated normal distribution with known mean value $M$and variance σ^2 and constraints LB, RB (estimated from empirical data). Since size of compressed row of domains is always integer, probability that size will be equal to l is $\Pr\{L = l\} \hat{=} p(l) = F_1(l) - F_1(l-1)$ (where $F_1(l)$ is cumulative distribution function)

E. Theoretical estimation

Exploiting the above mentioned assumptions, we can find the theoretical estimation for optimal size of high-priority packet. The results of theoretical estimation can then be compared to results of simulation with real video sequence. We denote$MSE(l)$ a random variable that is equal to $MSE$of reconstructed image on decoder

$$MSE(l) = \begin{cases} MSE_{\min} \text{ all data was received without errors} \\ MSE(l) \ l \text{ bytes of data was received without errors} \\ MSE_{\max} \text{ no data} \end{cases}$$

where$MSE_{\min}$ is an average MSE for compressed row of domains (estimated from empirical data). $MSE_{\max} = 64^2$, $MSE(l) = \frac{255^2}{10^{\frac{PSNR(l)}{10}}}$, where $PSNR(l)$ can be found exploiting rate-distortion model (assumption A.3).

If our assumptions are correct, mean square error $MSE_{movie}$of real video sequence should approximate average of random variable $E(MSE(l))$ with growth of number of frames in movie ($N_{rows} \to \infty$). We are going to find an expected distortion. Considering assumption A.2, we can calculate probability of error in packet of size $l$bytes in the following way: $PER(l) = 1 - (1 - BER)^{l \cdot 8}$, where $BER$is the bit error rate.

Main parameters of transmission scheme are:

- Time interval for transmission $T_{total}$and size of compressed data l
- Maximum number of transmissions of high-priority packet$\max_tr_num(L1)$, depends on selected transmission rate, time interval for transmission T_{total} and size $L1$ of high-priority packet. Size of packet payload of UWB is variable, and may lay in range (0,4095) bytes.
- $avail\_tr(l, L1)$, the maximum number of transmissions of first packet, provided that there is still enough time to send remaining data. It depends on selected transmission rate, time interval for transmission T_{total} and sizes $L1, l$ of first packet and all data accordingly.

We want to estimate the outage probability and expected distortion$E(MSE(l))$. If we fix size of high-priority packet L1, several situations are possible (see table I).

The expected distortion can be calculated as $E(MSE(l)) = \sum_{i=1}^{6} P_i \cdot MSE_i$Outage probability is equal to$P_{out} = P_2 + P_6$, and optimization problem can be reformulated as follows:

$$\begin{cases} E(MSE(l)) \to \min \\ P_{out} \le P_{out}^{*} \end{cases}$$

As can be seen, in order to select optimal size for first packet it is enough to know parameters (mean, variance) of distribution of sizes of compressed domain rows and parameters for rate-distortion model.

TABLE I

A. $LB < l < L1$, in this case all data can be put into one packet.

1.	Packet received successfully (all data is received)	$P_1 = \sum_{l=LB}^{L1} \left(1 - PER(l)^{\max\_tr\_num(l)}\right) \cdot p(l)$	$MSE_1 = MSE_{\min}$
2.	Packet was not received (no data at decoder)	$P_2 = \sum_{l=LB}^{L1} PER(l)^{\max\_tr\_num(l)} \cdot p(l)$	$MSE_2 = MSE_{\max}$

B. $L1 < l < RB$, data is split into several packets. Note, that if first (high-priority) packet will be resent several times (as a result of packet loss), there may be not enough time to send remaining (low-priority) data. Let us denote

$$P_{1st\_packet} = (1 - PER(L1)) \cdot \sum_{k=0}^{avail\_tr(l,L1)-1} PER(L1)^k$$

$$P_{1st\_packet\_only} = 1 - P_{1st\_packet} - PER(L1)^{\max\_tr\_num(L1)}$$

where P_{1st_packet} is the probability that first packet was received successfully, and there is enough time to send remaining data,

$P_{1st\_packet\_only}$ is the probability that first packet was received successfully, and there is not enough time to send remaining data

3.	Both packets are received successfully (all data is received)	$P_3 = \sum_{l=L1}^{RB} P_{1st\_packet} \cdot (1 - PER(l - L1)) \cdot p(l)$	$MSE_3 = MSE_{\min}$
4.	High priority packet received successfully, low-priority packet was not received	$P_4 = \sum_{l=L1}^{RB} P_{1st\_packet} \cdot PER(l - L1) \cdot p(l)$	$MSE_4 = MSE(L1)$
5.	High priority packet received successfully, and there was no time to send low-priority data	$P_5 = \sum_{l=L1}^{RB} P_{1st\_packet\_only} \cdot p(l)$	$MSE_5 = MSE(L1)$
6.	Packet was not received (no data at decoder)	$P_6 = \sum_{l=L1}^{RB} PER(L1)^{\max\_tr\_num(L1)} \cdot p(l)$	$MSE_6 = MSE_{\max}$

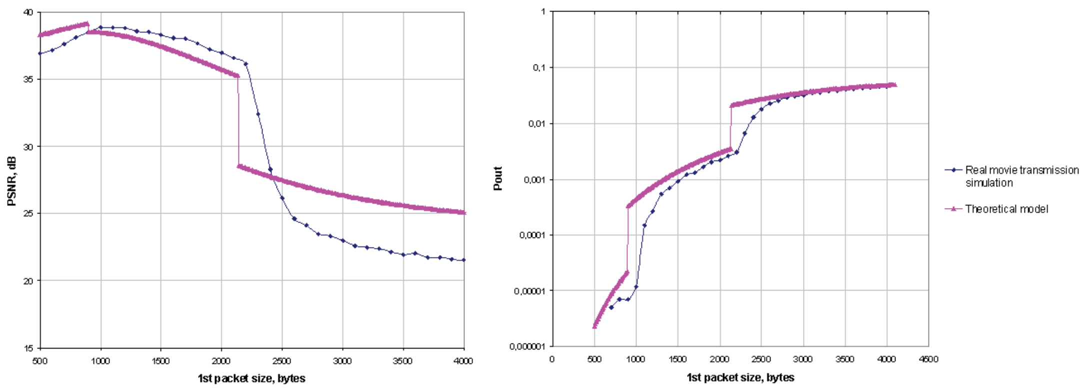

Fig. 3. Dependence of average PSNR and P_{out} on size of 1st packet (L1) for BER = 10e-5

F. Results

Experiment results and theoretical estimations are shown in the Figures 3,4.

One can see that outage probability and average PSNR have step-like form that can be easily explained. Since the overall delay for each tile must be the constant, the transmission window (time, available for transmission of row of domains) is also fixed and depends on frame rate and encoded tile size only. It means that increase of the packet size leads to decrease of number of available retransmissions (since we have strict time limitations): 2?1?0. Steps (sharp drop of PSNR and jump of P_{out}) occur when number of available retransmissions is decreased by one. Otherwise if number of available retransmissions doesn't change when packet size is growing smoothly then P_{out} is also increasing smoothly little by little. One can also see that theoretical model estimation provides rather good approximation of the real compression and transmission system. It means that simulation model can be used for analysis and selection of optimal parameters (if average (M) and deviation (σ) values of compressed tiles distribution, RD function and $MSE_{\max}$ are determined correctly). Firstly the required acceptable level of outage probability P_{out} ("step"

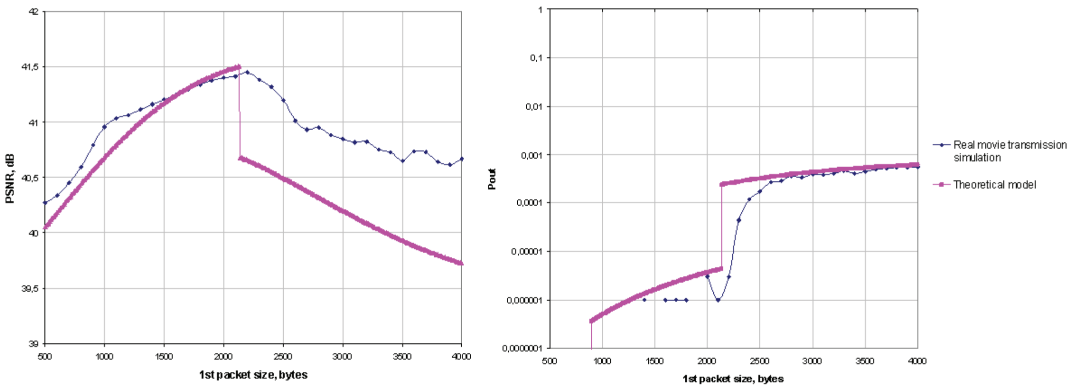

Fig. 4. Dependence of average PSNR and P_{out}on size of 1st packet (L1) for BER = 10e-6

that gives the range of first packet sizes) should be selected and than a packet size on this ”step” that maximizes mean PSNR can be chosen. For example taking a Tx speed = 480 Mbs and $BER = 10^{-6}$, we can select high-priority packet size equal to 886 bytes that gives a pair [outage probability, PSNR] = $[3.51 * 10^{-7}$, 40.5dB]

ACKNOWLEDGMENT

This research activity was sponsored by a special grant of Intel CTG research council

REFERENCES

[1] Standard ECMA-368, “High Rate Ultra Wideband PHY and MAC Standard.”
[2] Stankovic et al, “Live video streaming over packet networks and wireless channels,” *13th Packet Video Workshop, Nantes*, 2003.
[3] F. Zhai, Y. Eisenberg, T. N. Pappas, R. Berry, A.K. Katsaggelos, “Rate-distortion optimized hybrid error control for real-time packetized video transmission,” *IEEE Transactions on Image Processing*, Processing 15(1): 40-53, 2006.
[4] A. Ortega, “Optimization techniques for adaptive quantization of image and video under delay constraints,” *Ph.D. Thesis, Dept. of Electrical Engineering, Columbia University*, Jun. 1994.
[5] Diego Santa Cruz et al, “OISO/IEC JTC1/SC29/WG1 N1816 Coding of Still Pictures,” *An analytical study of JPEG 2000 functionalities*, July 2000.
[6] William B. Pennebaker, Joan L. Mitchel, “JPEG still image data compression standard,” *NY*, 1993.
[7] “JPEG 2000 Standard Part 1,” *Final Committee Draft v.1.0*, 16 March 2000.
[8] “Ultra-Wideband (UWB Technology) Enabling high-speed wireless personal area networks. White Paper,” *300982-002 US, Intel Corporation*, 2005, http://www.intel.com/technology/comms/uwb/download/ultra-wideband.pdf.