Available online at www.sciencedirect.com
SCIENCE DIRECT®
Signal Processing: *Image Communication* 19 (2004) 67–79
SIGNAL PROCESSING: IMAGE COMMUNICATION
www.elsevier.com/locate/image

Unequal error protection for MPEG-2 video transmission over wireless channels

Chung-Lin Huang*, Sling Liang

Electrical Engineering Department, National Tsing-Hua University, Hsin-Chu 30043, Taiwan

Received 17 June 2002; received in revised form 12 February 2003; accepted 11 August 2003

Abstract

This paper proposes an unequal error protection (UEP) method for MPEG-2 video transmission. Since the source and channel coders are normally concatenated, if the channel is noisy, more bits are allocated to channel coding and fewer to source coding. The situation is reversed when the channel conditions are more benign. Most of the joint source channel coding (JSCC) methods assume that the video source is subband coded, the bit error sensitivity of the source code can be modeled, and the bit allocations for different subband channels will be calculated. The UEP applied to different subbands is the rate compatible punctured convolution channel coder. However, the MPEG-2 coding is not a subband coding, the bit error sensitivity function for the coded video can no longer be applied. Here, we develop a different method to find the rate-distortion functions for JSCC of the MPEG-2 video. In the experiments, we show that the end-to-end distortion of our UEP method is smaller than the equal error protection method for the same total bit-rate.

Keywords: Joint source channel coding (JSCC); Unequal error protection (UEP); Rate compatible punctured convolution (RCPC) coding; Equal error protection (EEP)

1. Introduction

Wireless image and video transmission has become an essential component of wireless systems. It is the main system bottleneck because it requires more bandwidth than transmission of the other information sources. Conventional communication systems, which consider separate source and channel coding, are not suitable for transmitting video in wireless channel because they failed to consider the highly time-varying channel characteristics, the severe influence of channel loss on the encoded source, and the unequal importance of transmitted bits.

A common approach for building joint source channel coding (JSCC) is to cascade the existing source coding and channel coding. An important question is how to distribute the source bits and the channel bits between source coder and channel coder so that the resulting end-to-end distortion is minimized. Subband coding has been applied for source coder [17,19], which facilitates the rate-distortion function modeling, the bit allocation based on the so-called bit error sensitivity, and the UEP channel coding. The bit allocation methods [4,6,12] have been proposed for JSCC under the conditions that the video is subband coded and the

*Corresponding author. Fax: +886-3571-5971.
E-mail address: clhuang@ee.nthu.edu.tw (C.-L. Huang).

0923-5965/$ - see front matter
doi:10.1016/j.image.2003.08.018

subband distortion function (in terms of the number of bits of the source code and the channel code) can be computed. However, the bit-rate allocation for encoding MPEG-2 video has not been mentioned because it is difficult to model the propagation property of the variable length coding and the bit error sensitivity. This paper proposes a new method to find the rate-distortion functions and how to apply the UEP coding for transmitting the MPEG-2 video over the wireless networks.

For image and video transmission applications, the bit-rate is always a major concern for transmitting the encoded video in the band-limited and noisy channel. Hanzo and Streit [10] contrived a range of programmable constant-rate videophone codecs, which can adjust their coding rate in order to accommodate their stream in a conventional speech channel. Bit allocation is a classic problem for source coding, where a given number of bits are distributed among a finite set of quantizers to minimize a distortion measure. It also has been studied by Tao et al. [20] and applied to MPEG video encoding. Similar to bit allocation in source coding, the JSCC developed to optimally split the redundant information between source and channel coded data. Kung et al. [14] developed an adaptive JSCC method that the encoder may adjust quantization parameter and channel code rate based on the current bit error rate (BER). They proposed on-line estimation of R-D model and quantization model.

Hagenauer [8,9] mentioned that a priori information about the source helps the Viterbi decoder estimating the correct path for rate compatible punctured convolution (RCPC) encoded data. Bush et al. [1] demonstrated that, for certain situations, it pays to leave redundancy in the source instead of trying to get rid of it by compression and than reinserting it with the channel coding. However, the question of optimal partitioning the parity information between source and channel coders was not addressed. Kozintsev et al. [13] addressed how to efficiently represent an image or video source into compressed and uncompressed subsets, each suited to the appropriate mode of transmission. Ruf and Modestino [18] proposed a fixed-length JSCC scheme for image transmission over additive white Gaussian (AWGN) channels, in which different channel codes are applied to different bits according to their respective importance on the reconstructed image. Cai and Chen [4] also proposed a fixed-length JSCC scheme for generalized Gaussian sources and an all-pass filtering source reshaping [5].

Studies have been conducted on the analysis of rate-distortion curve for MPEG source codecs [16]. A more general rate-distortion analysis for transform coding has been developed by He and Mitra [11]. Frossard and Verscheure [7] separated the total distortion into source distortion and channel distortion. They proposed a distortion model relating the channel distortion to the packet loss ration and average burst length, which is almost independent of source rate. However, when the MPEG source is transmitted in noisy channels, the end-to-end distortion consists of the channel error, spatial and temporal error propagation, and the error concealment process. Bystrom et al. [2] extended their earlier work to independently combined source channel rate allocation [3] and employed models of universal distortion-rate characteristics, which are functions of BER per frame and can be jointly encoded with the channel codes.

The goal of JSCC is to distribute the total fixed bit-rates to source coder and channel coder so that the end-to-end distortion is minimized. This paper describes how the quantization error in the source coding stage and the channel noise in transmitting stage contributes to the overall distortion as an explicit function of the parameters such as source bit-rate and specified channel code rates for operation over AWGN channel. Based on the different importance of source components in reconstructed image, we may develop a parametric rate-distortion-based bit allocation scheme for both the source and channel coders (see Fig. 1).

For MPEG-2 video, the quantization error contributes to the source distortion, and a channel transmission error that contributes to the channel distortion. Channel distortion is a function of BER. With RCPC, we can easily select different code rates without changing the structure of the channel codec. We extend the concept of the universal distortion-rate characteristics [3]

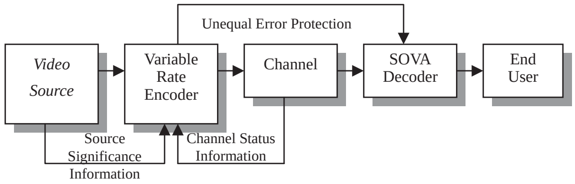


Fig. 1. JSCC system block.

together with the RCPC channel codec to optimally allocate source and channel coding bit-rate for MPEG-2 video which is more complicated than H.263. We have developed an optimization JSCC scheme under a certain channel state and total bit budget. Different from [3], our system may also adjust the JSCC coding scheme based on varying channel state and total bit budget. Our method is also different from [7] in that the FEC packet (instead of RCPC) is applied to the video packet. In [7], they did not consider that the overall distortion is sensitive to the influence of channel and occurs on different video sources such as I frame, P frame, or B frame.

This paper addresses that the channel distortion functions for different video sequences are similar. The difference can be simplified as a scaling factor. In the experiments, we will show that the end-to-end distortion of our UEP method is smaller than the equal error protection (EEP) method for the same total bit-rate.

2. System overview

The distortion in MPEG-2 video source coding is caused mainly by quantization process and other truncation operations. Even without channel errors, the quantization error will induce the propagation distortion of the reconstructed video. The MPEG-2 encoder consists of block-DCT, quantization, zig-zag scan, run-length and Huffman coding. The MPEG-2 bit streams are sensitive to channel error due to its variable length characteristics. When there are transmission errors, it is difficult for the decoder to locate and isolate the error bits. Therefore, the error propagation is unavoidable during decoding process. We can only assume the rest of bit stream after the error bit is lost until next synchronization flag, usually at the next slice header.

If there are channel errors, the MPEG-2 decoder may lose synchronization in the decoding stage. The BER of the received bit stream may be very trivial, but it may cause severe destruction of the image contents due to the error propagation of the source decoding. Though the strategies of error control are not discussed in the scope of the MPEG-2 standard, the decoding algorithm should be resilient to propagation of errors and facilitate concealment of errors. The decoded pictures via error concealment have significant increment in PSNR.

If the errors are detected through the arrival of invalid bit stream, the decoder will discard the rest of macroblocks until the next slice is detected. These skipped macroblocks are undefined and they are assigned some fixed values depending on decoders. To compensate these corrupted macroblocks, the blocks at the same position in the previous frame is copied to that corresponding macroblock. The visibility of error is then rather low unless they occur in I-pictures, in which case the substitutes in previous frame are no longer good enough and they may cause the errors propagate through all GOPs. To solve this problem, the corrupted macroblocks of I-pictures may be spatially interpolated by the neighboring macroblocks or by temporal error concealment.

To minimize the impact of the transmission error, an appropriate choice of channel error correcting/detecting codes is necessary. Using RCPC, we need only one convolutional encoder to generate flexible codes and employ the soft decision Viterbi decoder [9,21]. RCPC codes are families of channel codes that are obtained by puncturing the output of a "mother" convolution code. Puncturing is the process of removing, or deleting, bits from the output sequences in a predefined manner so that fewer bits are transmitted than the original code leading to a higher code rate. The idea of puncturing was extended to include the concept of rate compatibility [8]. Rate compatibility requires that a higher-rate code is a subset of a lower-rate code, or the lower-protection code is embedded into the higher-protection code. This is accomplished by puncturing a mother

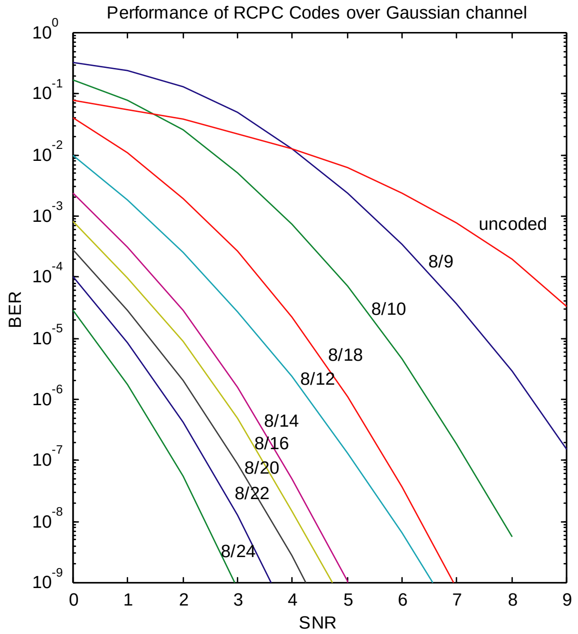


Fig. 2. Performance of RCPC code family.

code of rate $1/n$ to achieve higher rate (less protection).

No constructive method is known for determining a good mother code and the puncturing table $a(l)$ for an RCPC code family. Therefore, a computer full search has been performed under some restrictions [15]. In this paper, we choose the code (1101, 1111, 1011), with memory $M = 3$, period $P = 8$, and $l = 1, \ldots, 16$, as our RCPC mother code. We generate a RCPC code with code rate $P/(P + l)$, $l = 1, \ldots, 16$. The performance of selected RCPC codes on a Gaussian channel with soft decision under different channel states are simulated and results are given in Fig. 2. As shown in Fig. 2, lower code rate makes lower bit error probabilities, which means better protection for combating the channel errors. There is an interesting region that the uncoded bits outperform some channel coded bits at lower SNR.

3. Joint source and channel coding

In the conventional communication system, the source coder removes the redundancy from the source so that the source data can be compressed, whereas the channel coder inserts the controllable redundancy, the channel bits, for detection and correction of the possible channel errors. According to Shannon's information theory, source and channel coding can be treated separately. However, for real video transmission, even though the source encoded bit stream is almost statistically independent, the individual bits normally differ in their relative importance or sensitivity to the channel errors. For wireless and satellite communication, the channel is inherently unreliable and channel errors are unavoidable. Under these situations, the separate design of source and channel coding fails to provide the best solution.

In our systems, we split the encoded MPEG-2 bit stream into three layers due to their unequal error sensitivity. The most importance layer is I-pictures in bit stream, then P-pictures, and the least important layer is B-pictures. We define the rate-distortion function as

$$D_{\text{Total}} = \sum_{i=0}^{K-1} \frac{S_i}{S}[E_i + A_i P_i], \tag{1}$$

where K *indicates* the number of layers, ($K = 3$), S denotes the total bits in source data, S_i means the number of bits in ith layer, E_i shows the average source distortion induced by ith layer, A_i is the bit error sensitivity of ith layer, and P_i denotes the equivalent channel bit error rate after channel coding for the ith layer. A_iP_i stands for channel distortion. The overall bit-rate (bits/s) can be written as

$$R_{\text{Total}} = \sum_{i=0}^{K-1} \frac{S_i}{S}[R_{\text{S}}^i + R_{\text{C}}^i], \tag{2}$$

where R_{Total} indicates the total bit-rate, R_{S}^i means the source bit-rate for ith layer, R_{C}^i is the channel redundancy bits for ith layer.

If we can find the A_i for each layer, then we can easily estimate total distortion of the reconstructed video. Unfortunately, the MPEG-2 encoded bit stream is complex and the bits in the same layer may have different error sensitivity. In addition, the errors in I- and P-pictures will also propagate to the following pictures. It is difficult to evaluate the effect of propagation errors in different pictures. Therefore, we need to model the universal multi-dimensional rate-distortion function that can be applied to characterize the source coding errors, channel errors, and error propagation, and then we can describe the bit error sensitivity through the relationship of the rate-distortion function and the channel BER.

To measure the distortion, we use the mean square error (MSE) to measure the end-to-end video distortion. The distortion of the N-frame sequence is defined as

$$D_{\text{seq,Total}} = \frac{1}{M}\sum_{m=1}^{M}\frac{1}{N}\sum_{n=1}^{N}\sum_{h=1}^{H}\sum_{w=1}^{W}(x_n(h,w) - \tilde{x}_n(h,w))^2, \quad (3)$$

where H is the height of pictures, W indicates the width of pictures, N denotes the frame number in sequence, and M means the number of times the simulation. $x_n(h,w)$ is the intensity of the pixel located at (h,w) of the nth original picture, and $\tilde{x}_n(h,w)$ is the intensity of the pixel located at (h,w) of the nth reconstructed image frame. The parameter M is assumed to be a large number to assure statistical sufficiency for modeling the resulting distortion. The encoded video sequence is a color CIF (352×288) "Stefan" sequence, with $N = 40$, and $M = 1000$. Here, we denote the logarithms of the inverse BER (P_i) as

$$\tilde{P}_i = \log_{10}\left(\frac{1}{P_i}\right), \quad i \in \{\text{I}, \text{P}, \text{B}\}. \quad (4)$$

To simulate the transmission of video sequence, the video sequence is encoded with a specified bit-rate, and then split into three layers for separate transmission. There are many combinations of different layer protection via equivalent channel BER. In our example, we choose $\tilde{P}_i = 3, \ldots, 9$, identifying the channel BER ranges from 10^{-3} to 10^{-9}. The simulated result is shown in Fig. 3.

In Fig. 3, if we fixed P_P, the distortion curve is monotonically decreasing with the increasing inverse BER P_B, and we find the same tendency when we fixed P_B. We may find that the influences of the P_P and P_B on the rate-distortion curve are different. The 1D rate-distortion curve of P_P (fixed P_I and P_B) has higher slope than the 1D rate-distortion curve of P_B (fixed P_I and P_P), which

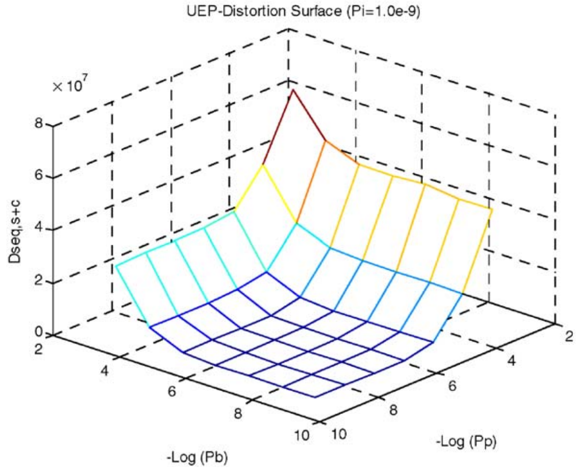


Fig. 3. Distortion–BER relation.

means that under the same protection, the error sensitivity of P-pictures is much higher than B-pictures. At very low BER, there is a distortion floor, which is independent of different P_I, P_P or P_B. We find that it is the source distortion and it will vary with different source bit-rate. Therefore, we may assume that the source distortion is unrelated with the channel distortion, and rewrite the rate-distortion function as follows:

$$D_{\text{seq,S+C}} = \sum_{i=0}^{K-1}(D_\text{S}^{(i)} + D_\text{C}^{(i)}) = D_\text{S}(R_\text{S}) + D_\text{C}(\tilde{P}_\text{I}, \tilde{P}_\text{P}, \tilde{P}_\text{B}), \quad (5)$$

where $D_\text{S}(R_\text{S})$ is the source distortion and $D_\text{C}(\tilde{P}_\text{I}, \tilde{P}_\text{P}, \tilde{P}_\text{B})$ indicates the channel distortion, and

$$R_\text{Total} = R_\text{S} + \sum_{i=0}^{K-1}\frac{S_i}{S}CR_i, \quad (6)$$

where $K = 3$, $i = 0$, 1 or 2 means I-, P-, or B-pictures, respectively, R_Total indicates the total available bit-rate, R_S is the source coding bit-rate, S denotes the total bits in source data, S_i is the number of bits in ith layer, and CR_i is the equivalent channel code rate for BER P_i. Given total bit-rate R_Total and channel status (SNR), by using the universal rate-distortion functions, we may find R_S and R_C that create the minimal end-to-end distortion.

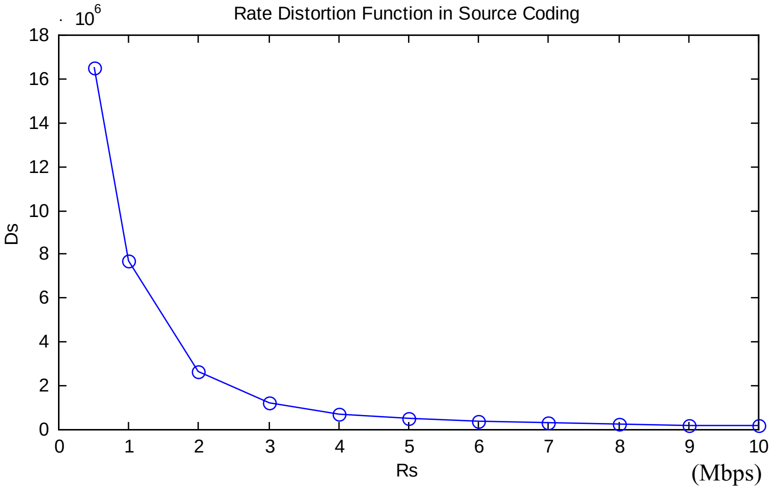


Fig. 4. Rate-distortion function of source coding for the sequence "Stefan".

To find the advantages of using the rate-distortion function for the JSCC, we simulate the entire source encoding and decoding and find different distortions D_S under different source bit-rates. The rate-distortion function for MPEG-2 video source is shown in Fig. 4. To simulate the channel distortion, we find that D_C is relatively complex because there are three variables to be considered. Similar to Fig. 3, we only simulate the integer point of the inverse bit error probabilities, $-\log P$, from 3 to 9 for I-, P-, and B-pictures. The channel distortion is defined as the MSE between the two video source sequences before and after channel decoding, respectively. To simulate the video sequence transmission, we require a set of RCPC codes of which the different code rates will generate different channel bit error probabilities for certain specific channel status (in terms of SNR).

Here, we select the range of channel BER from 10^{-3} to 10^{-9} that is normally used for simulating the video transmission over the AWGN channel. The code rate selection under different SNR is list in Table 1. For example, if the known channel state is SNR = 3 dB, there are seven possible combinations of code rates that can be used in channel coding, the corresponding channel bit error probabilities are $10^{-3.58}$, $10^{-4.56}$, $10^{-5.81}$, $10^{-6.31}$, $10^{-7.06}$, $10^{-7.9}$, and $10^{-9.08}$.

Table 1
Code rate selection for different SNR based on available bit-error rate

Rate	SNR (dB)								
	0	1	2	3	4	5	6	7	8
							3.47	4.42	5.53
8/9							5.34	6.72	8.26
8/10					3.12	4.15	7.43	9.06	
8/12				3.58	4.66	5.95	8.19		
8/14			3.60	4.56	5.63	6.87			
8/16		3.51	4.55	5.81	7.30	8.97			
8/18	3.09	4.01	5.06	6.31	7.82				
8/20	3.56	4.54	5.70	7.06	8.56				
8/22	4.00	5.07	6.37	7.90					
8/24	4.55	5.75	7.26	9.08					

Once we have the D_S and D_C, we can combine them to obtain the universal rate-distortion function $D_{seq,S+C}$. For a given total bit-rate R_{Total}, we may search all possible combinations of (R_S, P_I, P_P, P_B) to find the best one that minimizes the end-to-end distortions. The flow diagram of the joint source and channel bit allocation is described in Fig. 5. The optimal solutions for different UEP of the source with minimum end-to-end distortion $D_{seq,S+C}$ for different R_{Total} with channel status SNR = 3 dB are listed in Table 2.

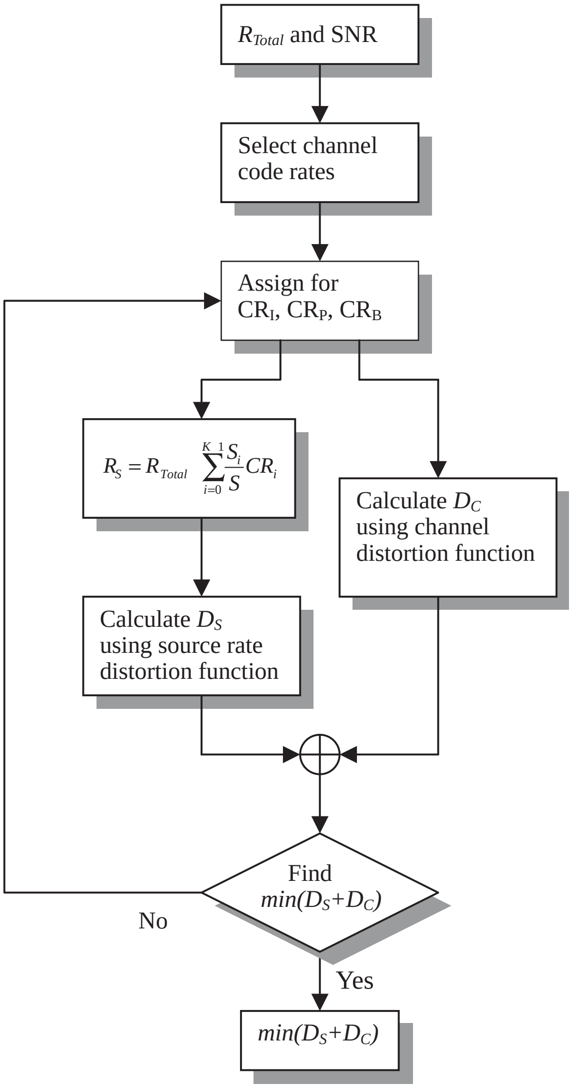


Fig. 5. Flowchart of bit-rate distribution.

Table 2
Experimental results of optimal bit-rate allocation with SNR = 3

R_{Total} (Mbps)	R_S (Mbps)	$\tilde{P}_I$	$\tilde{P}_P$	$\tilde{P}_B$
2.0	0.9	7.90	6.31	5.81
3.0	1.4	6.31	6.31	5.81
4.0	1.8	7.9	6.31	5.81
5.0	2.2	7.06	7.06	5.81
6.0	2.5	7.06	7.06	6.31
7.0	2.9	7.06	7.06	6.31
8.0	3.3	7.60	7.06	6.31
9.0	3.7	7.90	7.06	6.31
10.0	4.1	7.90	7.06	6.31
11.0	4.6	7.06	7.06	6.31
12.0	4.7	7.90	7.06	7.06
13.0	5.1	7.90	7.06	7.06
14.0	5.5	7.90	7.06	7.06
15.0	5.8	7.90	7.06	7.06
16.0	6.0	7.90	7.06	7.06

Table 3
Comparison of UEP and EEP

R_{Total} (Mbps)	R_S (Mbps)	$\tilde{P}_I$	$\tilde{P}_P$	$\tilde{P}_B$	D_{seq}
4.0	1.8	7.90	6.31	5.81	3267123.5
4.0	1.4	7.90	7.90	7.90	4972335.8
4.0	1.7	6.31	6.31	6.31	3662436.6
4.0	2.0	5.81	5.81	5.81	3555650.7

Table 4
Comparison of UEP and EEP

R_{Total} (Mbps)	R_S (Mbps)	$\tilde{P}_I$	$\tilde{P}_P$	$\tilde{P}_B$	D_{seq}
10.0	4.1	7.90	7.06	6.31	716350.5
10.0	3.6	7.90	7.90	7.90	878059.6
10.0	3.9	7.06	7.06	7.06	772455.4
10.0	4.4	6.31	6.31	6.31	771110.8

Selecting two entries from Table 2 ($R_{Total} = 4.0$ and $R_{Total} = 10.0$), we compare the end-to-end distortion of the optimal bit-distribution for unequal error protection (UEP) with that of the even distribution for EEP in Tables 3 and 4. For different channel status (different SNR), we may find the other set of best bit-rate assignment for (R_S, P_I, P_P, P_B) that also create the minimum end-to-end distortion under certain fixed total bit-rate R_{Total}.

The frame to frame comparison of the PSNR for the UEP and EEP of the MPEG-2 encoded video "Stefan" is shown in Fig. 6, and the visual comparison is shown in Fig. 7.

4. Modeling the rate-distortion function for JSCC

We have proved that the JSCC bit-rate distribution with UEP provides the optimal solution via full search of vector (R_S, P_I, P_P, P_B) with a given R_{Total}. However, the rate-distortion functions of D_S and D_C usually are not on-line obtainable, so that they cannot be used in distortion minimization operation. In this section, we will introduce how to develop a rate-distortion model and employ it to reduce the computational complexity for real-time application.

As we mentioned before, the overall rate-distortion function can be written as

$$D_{seq,S+C} = D_S(R_S) + D_C(\tilde{P}_I, \tilde{P}_P, \tilde{P}_B) \quad (7)$$

of which the source distortion is the function of R_S only. From the source rate-distortion function in Fig. 4, we find that the source coding distortion is monotonically decreasing with increasing bit-rate; therefore, it is reasonable that the source rate-distortion function can be approximated by

$$\tilde{D}_S(R_S) = \frac{\alpha_S}{(R_S)^{\beta_S}}, \quad (8)$$

where α_S and β_S are two unknown parameters to be determined. A solution to the unknown parameters is to minimize the following curve fitting formula:

$$\sum_{n=1}^{N}[D_{S,n}(R_{S,n}) - \tilde{D}_S(R_S)], \quad (9)$$

where N is the number of selected training data, and $D_{S,n}(R_{S,n})$ are the simulation results of the distortion for N different rates. Using the non-linear parameter estimation method, we can find the parameters α_S and β_S easily. The example of the estimated source rate distortion is shown in Fig. 8(a) by dotted line and the video source sequence is "Stefan". We find that using only three training points is sufficient for modeling the source rate-distortion function. The reasons for such a simple process are that the implementation is simple and the loosely fitted regions of the rate-distortion curve are not used in normal cases.

The small mismatch in low-source bit-rate ($R_S < 1$ Mbps) can be neglected because not all regions will be used in the bit allocation stage. From different types of MPEG-2 video and channel status, we find that the most commonly used source bit-rate is between 1.0 and 6.0 Mbps.

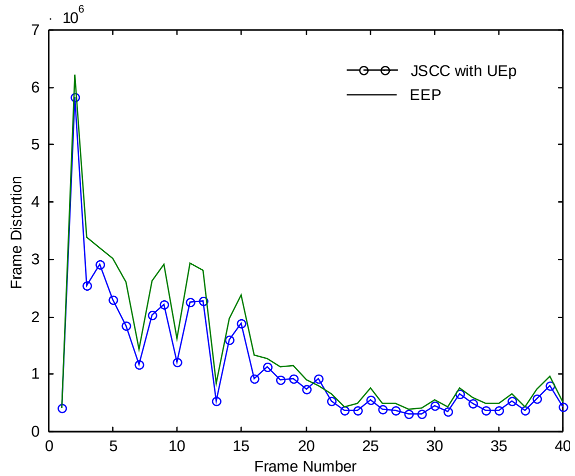
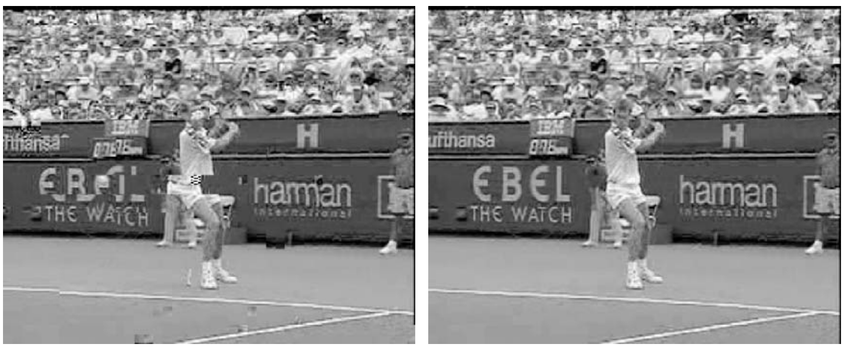


Fig. 6. Frame to distortion for EEP and UEP of the MPEG-2 encoded video "Stefan": (a) EEP $D_{seq,} = 0.878 \times 10^6$, and (b) UEP $D_{seq} = 0.716 \times 10^6$.



(a) Equal error protection $D_{seq,}$=0.878 · 10^6. (b) Unequal error protection D_{seq}=0.716 · 10^6.

Fig. 7. Rate allocation results with the overall bit-rate of $R_{Total} = 10$ Mbps.

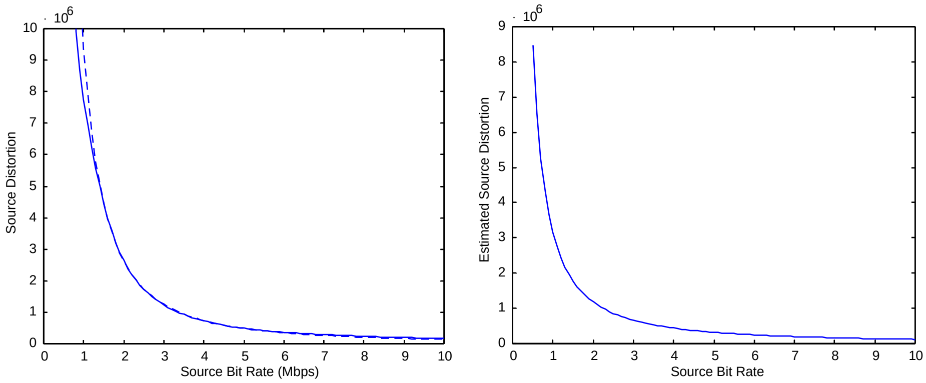


Fig. 8. The source rate distortion D_S for sequences (a) "Stefan", and (b) "Coast Guard".

The estimated source distortion function of "Coast Guard" is also shown in Fig. 8(b) with only three training points $R_S = 2.0, 4.0, 6.0$ Mbps.

To estimate the channel distortion for JSCC, we illustrate some examples of the simulated channel distortion–BER relation in Fig. 9 for sequence "Stefan". We can find that the channel distortion is also monotonically increasing with increasing channel bit error probabilities. Furthermore, I-pictures induce much more distortion than P-pictures under the same bit error probabilities, and the similar phenomenon occurs between the P- and B-pictures. These figures describe how the error propagates between pictures in the same GOP. In other words, most MPEG-2 encoded bit streams undergo similar error propagation effect induced by the channel bit errors. Therefore, the channel distortion models in Fig. 9 can be modified for estimating the channel rate distortion of other MPEG-2 encoded sequence.

Since the essence of MPEG-2 video coding is to encode the residue images of motion compensation, the channel error will create the error propagation of the source image which is related with the structure of GOPs. The propagation error is much larger than the original motion compensation error. Therefore, the channel distortion surfaces of different video sequences should be similar if they are encoded by the same GOP structure. We use the video sequence "Coast Guard" to verify our assumption. After the simulation of the channel distortion for the sequence "Coast Guard", we find that the similar channel distortion models can be generated, and the only difference is the scale parameter $\tilde{\alpha}_i$ which is about 0.68. It is understandable because the channel distortions caused by P_P or P_B are basically the errors in the residue images, which are similar for different video sequence with the same GOP structure. However, the channel distortions caused by different P_I do not change linear proportion. Considering this similarity, and the decomposition of the channel distortion function D_C from 3D function into a set of 2D functions, we may rewrite the channel distortion function as

$$D_C(\tilde{P}_I, \tilde{P}_P, \tilde{P}_B) = \sum_{i=1}^{O} \gamma_i \cdot \tilde{D}_C^i(\tilde{P}_P, \tilde{P}_B)\big|_{P_I = P_i}, \qquad (10)$$

where O is the set of some specified bit error probabilities, in our example, $O = 6$ and

$$\tilde{P}_i = \log\left(\frac{1}{P_i}\right), \quad P_i \in \{10^{-4}, 10^{-5}, \ldots, 10^{-9}\}. \qquad (11)$$

The points outside these six surfaces can be interpolated. For two different encoded MPEG-2 video sequences, the corresponding channel distortion functions (with the same P_I), $D_C(\tilde{P}_P, \tilde{P}_B)$, are different by a scale γ_i. To generate the channel distortion model for any video sequence, we only need to scale the pretrained channel distortion model (i.e., Fig. 9).

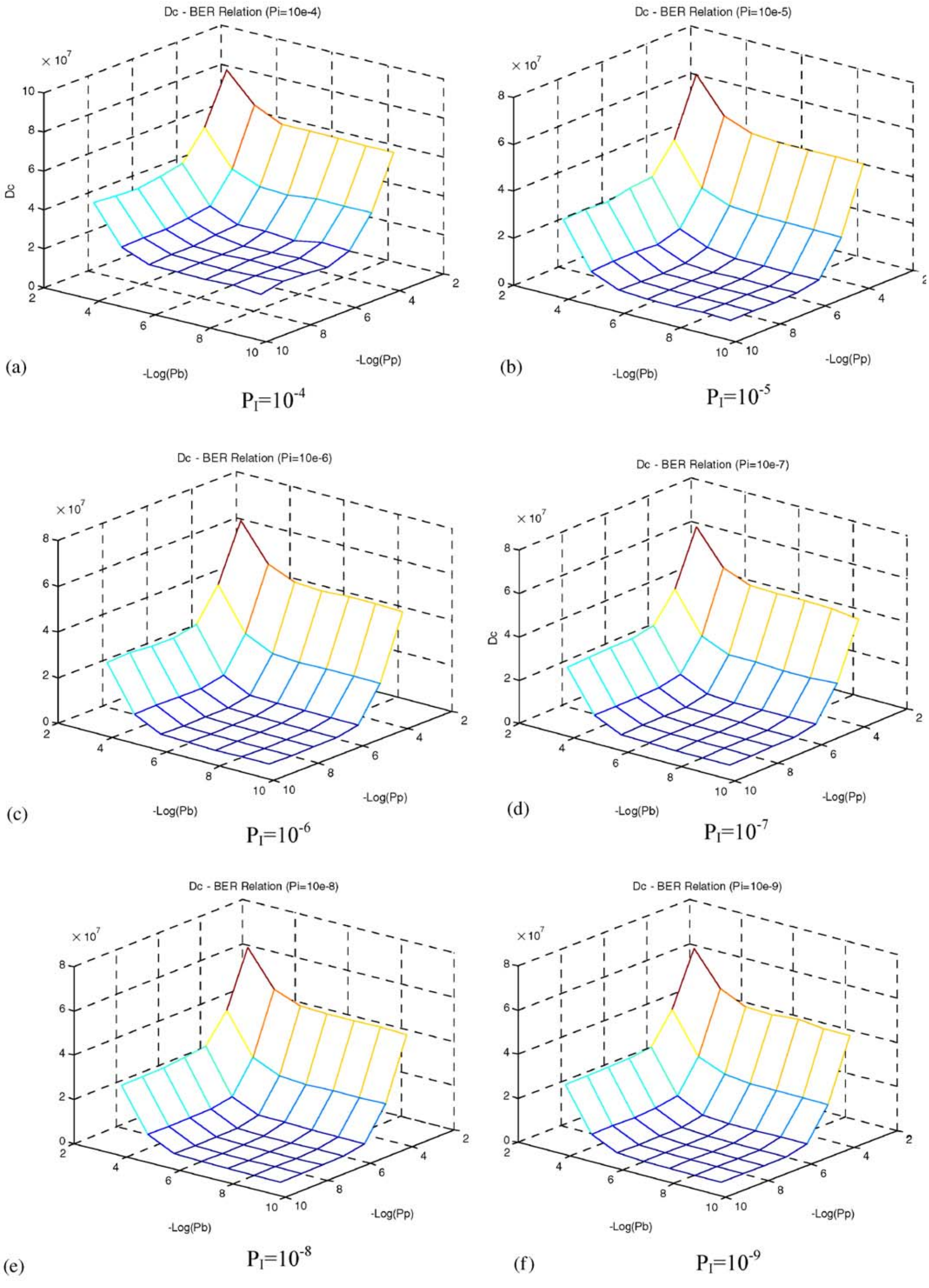


Fig. 9. Examples of channel distortion for sequence ''Stefan''. (a) $P_\mathrm{I} = 10^{-4}$, (b) $P_\mathrm{I} = 10^{-5}$, (c) $P_\mathrm{I} = 10^{-6}$, (d) $P_\mathrm{I} = 10^{-7}$, (e) $P_\mathrm{I} = 10^{-8}$, and (f) $P_\mathrm{I} = 10^{-9}$

5. Simulations

Here, we may apply the proposed scheme for transmitting video sequence "Coast Guard". The optimal bit-rate distribution can be evaluated easily. The estimated source distortion for sequence "Coast Guard" is shown in Fig. 8(b) for only three sets training data at $R_S = 2.0$, 4.0, 6.0 Mbps. Given the six channel-distortion surfaces in Fig. 9 as the primitive function of D_C and the trained scale parameter $\gamma_i = 0.68$, we can obtain another set of channel rate-distortion surfaces for modeling the new input video sequence as long as the two MPEG-2 videos have the same GOP structure. The scale parameter γ_i can be easily obtained by simulating one data point (with same BER P_I, P_P, P_B) in each channel distortion surface, i.e., the parameter γ_i is equivalent to the ratio of the channel distortions of the two encoded videos. Based on the source distortion model and channel distortion model, the end-to-end distortion under each combination of vector (R_S, P_I, P_P, P_B) with a given R_{Total} and channel status (SNR) is calculated. In real transmission, given different channel conditions, we may select the correct entry from the table and the transmitted video will have minimum end-to-end distortion. Nine different optimal bit allocations for video "Coast Guard" are listed in Table 5. To verify the advantage of bit allocation, we compare the overall distortion of the unequal bit allocation (UEP) with the three equal bit distributions (EEP) in Table 6.

We may also apply our JSCC scheme for the other video sequences. Before the channel coding of the MPEG-2 encoded video sequence, we need to generate the source rate-distortion function. To accomplish this, we need only to calculate three source distortions at three different source rates. To obtain the channel rate-distortion function, we need to model the six channel distortion functions at six different channel loss rates for I frame (P_I). For instance, take the video sequence "Foreman" as another example, the estimated source distortion model is shown in Fig. 10 which is generated by the three training distortion data at $R_S = 2.0$, 4.0, 6.0 Mbps. After the six simulated points on six channel distortion surface of $P_I = 10^{-4}$–10^{-9}, the average of the scale parameter γ_i is about 0.71.

Table 5
Bit allocation for sequence "Coast Guard" with SNR = 3

R_{Total} (Mbps)	R_s (Mbps)	$\tilde{P}_I$	$\tilde{P}_P$	$\tilde{P}_B$
2.0	0.9	7.90	6.31	5.81
3.0	1.3	7.06	7.06	5.81
4.0	1.8	7.90	6.31	5.81
5.0	2.0	7.90	7.06	6.31
6.0	2.5	7.90	7.06	6.31
7.0	2.9	7.90	7.06	6.31
8.0	3.3	7.06	7.06	6.31
9.0	3.7	7.90	7.06	6.31
10.0	3.9	7.90	7.06	7.06

Table 6
Testing for bit allocation in Table 5

R_{Total} (Mbps)	R_S (Mbps)	P_I	P_P	P_B	D_{seq}
6.0	2.5	7.90	7.06	6.31	6684513.7
6.0	2.1	7.90	7.90	7.90	8290945.1
6.0	2.4	7.06	7.06	7.06	7090214.2
6.0	2.6	6.31	6.31	6.31	6728699.1

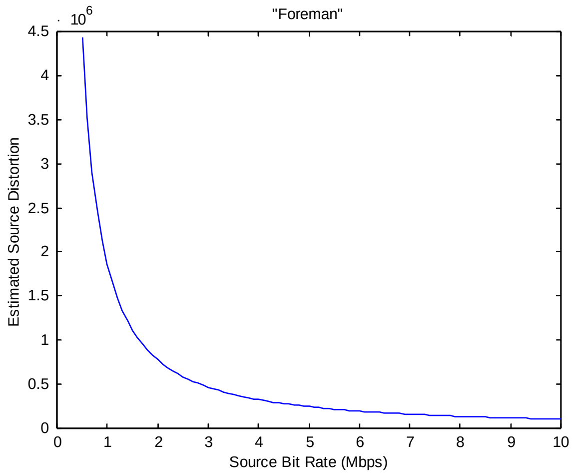


Fig. 10. Estimated source distortion for sequence "Foreman".

Table 7
Bit allocation for sequence "Foreman" with SNR = 3

R_{Total} (bps)	R_s (bps)	$\tilde{P}_I$	$\tilde{P}_P$	$\tilde{P}_B$
2.0	0.9	7.90	6.31	5.81
3.0	1.3	7.06	7.06	5.81
4.0	1.6	9.08	7.06	6.31
5.0	2.0	9.08	7.06	6.31
6.0	2.5	7.06	7.06	6.31
7.0	2.9	7.06	7.06	6.31
8.0	3.1	7.90	7.06	7.06
9.0	3.5	7.90	7.06	7.06
10.0	3.8	9.08	7.06	7.06

Using the interpolated source and channel rate-distortion functions, the optimal bit allocations for sequence "Foreman" can be obtained, which are shown in Table 7.

6. Conclusions

We have described a JSCC scheme for MPEG-2 video transmission. The bit-rate allocation algorithm and the concept of UEP are employed. This scheme provides an optimal bit-rate distribution via full search of all possible combinations of the BERs of the three picture types. To employ the UEP and reduce the complexity of channel coder, we use the RCPC codes. Employing the rate-distortion models can significantly reduce the computation of searching for an optimal solution from all possible combinations of (R_S, P_I, P_P, P_B). In our system, there are different bit error sensitivities in blocks, slices and different pictures of the same type. We believe that more layers can contribute to a better bit-rate allocation. We have found an effective way to find the rate-distortion function for transmitting different MPEG-2 encoded video sequences.

References

[1] G. Bush, F. Burket, J. Hagenauer, B. Kukla, To compress or not to compress? Proceedings of the GLOBECOM, 1996, pp. 198–203.

[2] M. Bystrom, J.W. Modestino, Combined source-channel coding schemes for video transmission over an additive white Gaussian noise channel, IEEE Journal on Selected Areas in Communication 18 (6) (June 2000) pp. 880–890.

[3] M. Bystrom, T. Stockhammer, Modeling of operational distortion-rate characteristics for joint source-channel coding of video, Proceedings of the IEEE ICIP, 2000.

[4] J. Cai, C.W. Chen, Robust joint source-channel coding for image transmission over wireless channel, IEEE Transactions on CAS for VT, September 2000, pp. 962–966.

[5] Q. Chen, T.R. Fisher, Image coding using robust quantization for noisy digital transmission, IEEE Trans. Image Process. 7 (April 1998) 496–505.

[6] G. Cheung, A. Zakhor, Joint source/channel coding for scalable video using model of rate distortion functions, IEEE Trans. IP 9 (3) (2000) 340–354.

[7] P. Frossard, O. Verscheure, Joint source/FEC rate selection for optimal MPEG-2 video delivery, IEEE ICASSP, 2000.

[8] J. Hagenauer, Rate-compatible punctured convolutional codes (RCPC codes) and their application, IEEE Trans. Commun. Vol. 36 (4) (April 1988) 389–400.

[9] J. Hagenauer, Source-controlled channel decoding, IEEE Trans. Commun. Vol. 43 (September 1995) 2449–2457.

[10] L. Hanzo, J. Streit, Adaptive low-rate wireless videophone schemes, IEEE Transactions on CAS for VT, Vol. 5 (4), 1995, pp. 305–318.

[11] Z. He, S.K. Mitra, A unified rate-distortion analysis framework for transform coding, IEEE Transactions on CAS for VT, Vol. 11 (12), December 2001.

[12] L.P. Kondi, A.K. Katsaggelos, Joint source-channel coding for scalable video over DS-CDMA multipath fading channels, Proceedings of the IEEE ICIP, 2001.

[13] I. Kozintsev, K. Ramchandran, A Hybrid compressed-uncompressed framework for wireless image transmission, Proc. IEEE ICIP 2 (October 1997) 77–80.

[14] W.Y. Kung, C.S. Kim, R. Ku, C.C.J. Kuo, Adaptive joint source/channel coding for robust video transmission, Proceedings of the ISCAS 2002, Phoenix, May 2002.

[15] L.H.C. Lee, New rate-compatible punctured convolutional codes for Viterbi decoding, IEEE Trans. Commun. 41 (12) (December 1994) pp. 3073–3079.

[16] L.J. Lin, A. Ortega, Bit-rate control using piecewise approximated rate-distortion characterisrtic, IEEE Transactions on CAS for VT, Vol. 8, 1998, pp. 446–459.

[17] K. Ramchandran, M. Vetterli, Multiresolution joint source-channel coding, in: V. Poor, G.W. Wornell (Eds.), Wireless Communications, Signal Processing Perspectives, Prentice-Hall, Englewood Cliffs, NJ, 1998, pp. 282–329.

[18] M.J. Ruf, J.W. Modestino, Operational rate-distortion performance for joint source and channel coding of images, IEEE Trans. IP 8 (3) (1999) 305–320.

[19] K. Sayood, H.H. Out, N. Demir, Joint source/channel coding for variable length codes, IEEE Trans. Commun. 48 (May 2000) 787–794.

[20] B. Tao, B.W. Dickinson, H.A. Peterson, Adaptive model-driven bit allocation for MPEG video coding, IEEE Transactions on CAS for VT, Vol. 10, 2000, pp. 147–157.

[21] A.J. Viterbi, J.K. Omura, Principles of Digital Communication and Coding, McGraw-Hill, New York, 1979.