Multimedia Systems manuscript No.
(will be inserted by the editor)

György Dán · **Viktória Fodor** · **Gunnar Karlsson**

Robust Source-Channel Coding for Real-time Multimedia

Abstract Multimedia applications operating in today's Internet have to employ some form of error resilience to cope with losses. For interactive applications with strict delay constraints the latency introduced by these schemes has to be low as well. Furthermore the parameters of the applied scheme have to be set based on measurements in a possibly rapidly changing environment. In this paper we propose a robust method, called min-max-α, for optimal source-channel code rate allocation to deal with time-varying packet channels and channel state estimation errors. We evaluate its performance when used with forward error correction (FEC) and multiple description coding (MDC) in both stationary and non-stationary environments. We show that on a stationary channel robust rate allocation is suboptimal in terms of mean distortion, but it achieves a lower variance, while on a non-stationary channel it prevents severe degradation of the quality. We apply the proposed rate allocation method to motion compensated video and show that it performs better on a non-stationary channel than minimization of the mean distortion proposed earlier.

Keywords Source-channel coding · MDC · FEC · Rate-distortion · Real-time multimedia

1 Introduction

Real-time audio and video communication is becoming more and more important in the Internet. The quality provided to these applications however is still unpredictable due to packet losses, end-to-end delay and delay jitter. Though much research and standardization has been done in recent years, there is no support yet from the network side for QoS in form of resource reservation and call admission control. Thus applications that require low packet loss and delay jitter have to employ some end-to-end mechanisms that compensate the disturbances introduced by the network and that are robust to variations and estimation errors in the quality of the transmission both on short and long timescales. The problem of delay jitter is often solved on the receiver side via adaptive playout algorithms [31,32]. Packet losses can be compensated for on the sender side and the receiver side. On the receiver side receiver-based error concealment algorithms can be used like insertion or interpolation [26,39]. On the sender side redundant information can be added to the data flow or error resilient source coding can be used. Whenever a packet is lost, the redundant information can be used to reconstruct the lost information. Forward error correction (FEC), and the recently re-discovered multiple description coding (MDC) have been proposed for this purpose [5,37]. Error resilience features are included in recent video coding standards, like the H.264 [41].

The ratio of redundancy has to be chosen in the sender in a way that maximizes the perceived quality at the receiver. There are several problems that have to be solved during this optimization, called optimal rate allocation. Since the receiver cannot calculate the distortion of the audio-visual information, it is the sender that has to estimate the effects of losses on the perceived quality at the receiver. The sender can attempt to do this by maintaining an estimate of the channel state, i.e., the state of the transmission path from the sender to the receiver, and an estimate model of the perceived quality as a function of the redundancy rate and the channel state. Most of the work in the literature assumes a stationary channel with known parameters and aims at minimizing the mean distortion [2,13,24,29,34,35,44]. A few other works include feedback in the optimization problem, but consider the channel state estimate to be accurate and use it to minimize the mean distortion [1,5,33].

In this paper we propose a method, called min-max-α, for optimal rate allocation to deal with time-varying channels and channel state estimation errors. Min-max-α aims to minimize the maximum distortion for a set of possible future channel states.

We compare the performance of the proposed method to that of the common approach in the literature, i.e., minimizing the mean distortion for the measured packet loss probability, under various network scenarios. First, we illus-

ACCESS Linnaeus Center, School of Electrical Engineering, KTH, Royal Institute of Technology, Stockholm, Sweden. E-mail: {gyuri,vfodor,gk}@ee.kth.se .

trate the behavior of our method when FEC or MDC is used for error resilience. We show that although the min-max-α method is suboptimal in terms of mean distortion on a stationary channel, it achieves significantly lower variance. On non-stationary channels the proposed method prevents severe degradation of the quality which would otherwise occur due to the misestimation of the channel parameters. Finally we show how the min-max-α method can improve the error-resilience of H.264 coded video in a non-stationary environment.

The joint selection of source coding parameters and transmission power for video transmission over wireless channels was considered in [8]. The authors used the term min-max to denote the minimization of the mean distortion given a bound on the maximum transmission power and the minimization of the transmission power given a maximum mean distortion. Despite of the similar name used for the approach, the problem presented in that paper is different from ours. In [30] the authors considered the problem of media broadcasting in a wireless cell and minimized the maximum disappointment of the users in the cell, that is, the difference between their perceived performance and the best possible performance. Despite the similar name used for these approaches, the addressed problems and the applied analytical tools are different from ours. The authors in [9] minimized the linear combination of the mean and the variance of the distortion of motion compensated video for given packet loss probabilities. Our approach is conceptually different and offers two advantages. It is computationally less intensive and it is not restricted to video coding: it is applicable to a wide variety of error resilience techniques as we show it in the paper.

The rest of the paper is organized as follows. In Section 2 we present the rate allocation methods for joint source-channel coding. In Section 3 we describe the considered transmission channel and the method used to estimate its state. In Section 4 we show how the proposed min-max-α method can be used to set the redundancy rate for FEC and MDC. In Section 4.2 we analyze the behavior of the rate-allocation methods on a stationary channel under various scenarios. In Section 4.3 we consider the case of non-stationary channels. In Section 5 we show how the rate-allocation methods compare when applied to H.264 coded video. We conclude our work in Section 6.

2 Rate allocation problem

We consider the scenario when delay-sensitive multimedia traffic, such as real-time video, is transmitted from a sender to a receiver through a network. Packets can be lost on the transmission path due to congestion, on unreliable wireless links, or can be discarded at the receiver due to late arrival. To cope with losses the sender adds redundant information to the data flow. We assume that the available transmission rate at the sender is given, thus redundancy can be increased by decreasing the source coding rate only.

We consider a receiver described with distortion rate function $D = D(R_a, p_\omega, p_{\omega|\omega}, \beta)$, where R_a is the available code rate, p_ω is the stationary loss probability, $p_{\omega|\omega}$ is the conditional loss probability (the probability that a packet is lost, given that the preceding packet was lost) and β is the ratio of redundancy. The higher the ratio of redundancy β, the higher the distortion will be in the absence of losses. In the presence of losses, however, a higher value of β gives increased resilience to errors. The distortion rate function depends on the source and on the means of error resilience used.

The goal of a rate allocation method is to adjust the value of β at the sender in order to minimize the effect of losses introduced by the network. The most widely used approach to solve this problem is to minimize the mean distortion for the stationary loss probability, i.e., to find the ratio of redundancy β^* that minimizes

$$D(R_a, \hat{p}_\omega, \hat{p}_{\omega|\omega}, \beta), \qquad (1)$$

where $\hat{p}_\omega$ is the estimated value of the loss probability and $\hat{p}_{\omega|\omega}$ is the estimated value of the conditional loss probability.

There are several shortcomings of this approach however [17]. This approach does not take into account the sensitivity of the human perception to short term variations of the quality. Psychoacoustic models and models of the human visual system prefer constant quality, as human observers can adapt on the long term to reasonably low, but constant quality. Even on a stationary channel with loss probability p_ω the number of lost packets j over a finite period of time, say, in a block of n packets, is a random variable (e.g. on a channel described by the Bernoulli loss model and mean loss probability p_ω, the number of lost packets in a block of n packets follows a binomial distribution with parameters (n, p_ω)). Hence, the estimated loss probability $\hat{p}_\omega$ is a random variable as well. If $D(R_a, p_\omega, p_{\omega|\omega}, \beta)$ is not a linear function of p_ω and $p_{\omega|\omega}$, then

$$D(R_a, p_\omega, p_{\omega|\omega}, \beta) \neq E_{\{\hat{p}_\omega, \hat{p}_{\omega|\omega}\}}[D(R_a, \hat{p}_\omega, \hat{p}_{\omega|\omega}, \beta)],$$

even though $E[\hat{p}_\omega] = p_\omega$ and $E[\hat{p}_{\omega|\omega}] = p_{\omega|\omega}$ if the estimator is unbiased. The fluctuations in the loss probability over short intervals can degrade the perceived quality.

2.1 Weighted mean distortion

To incorporate the effects of short term fluctuations of the channel state, we define the weighted mean distortion, a new evaluation criterion instead of the mean distortion defined by (1).

Let us consider a channel described with the model Ψ and the set of channel parameters $\mathcal{P}$. Let us denote by $E[s]$ the mean number of packets received between two loss events, and by $P(j, E[s]+1)$ the distribution of the number of lost packets in a block of $E[s]+1$ packets calculated based on the model Ψ and the parameters $\mathcal{P}$. Let us denote the density function of $j/(E[s]+1)$ by $\pi(p)$ and its distribution function by $\Pi(p)$. Furthermore we denote the γ percent confidence interval of p by Ω_γ, i.e., $\Omega_\gamma = \{p : \gamma/2 \leq 100\Pi(p) \leq$

$100-\gamma/2)$. Using these definitions we define the weighted mean distortion as

$$D=\int_{p\in\Omega_\gamma} D(R_a,p,p_{\omega|\omega},\beta)\pi(p)dp. \quad (2)$$

This definition of the mean distortion introduces the temporal behavior of the channel into the evaluation.

To capture the variation of the distortion we define the standard deviation of the distortion in a similar way.

$$\sigma=\sqrt{\int_{p\in\Omega_\gamma} E[\tilde{D}(R_a,p,p_{\omega|\omega},\beta)^2]\pi(p)dp-D^2}. \quad (3)$$

Finally we define the coefficient of variation of the distortion as $CoV=\sigma/D$. Given these evaluation criteria in the following subsection we introduce three rate allocation methods.

2.2 Rate allocation methods

In order for the application to be able to perform the rate allocation, it has to assume a channel model $\hat{\Psi}$ and has to estimate the channel parameters $\hat{\mathcal{P}}$. Let us denote by $\hat{s}$ the estimated mean number of packets received between two loss events, by $\hat{p}_\omega=1/(\hat{s}+1)$ the estimated loss probability, and by $\hat{P}(j,\hat{s}+1)$ the estimated distribution of the number of lost packets in a block of $\hat{s}+1$ packets calculated based on the model $\hat{\Psi}$ and $\hat{\mathcal{P}}$. Let us denote the density function of $j/(\hat{s}+1)$ by $\hat{\pi}(p)$ and its distribution function by $\hat{\Pi}(p)$. Furthermore we denote the estimated α percent confidence interval of p by $\hat{\Omega}_\alpha$, i.e., $\hat{\Omega}_\alpha=\{p:\alpha/2\le 100\hat{\Pi}(p)\le 100-\alpha/2)$.

Crisp rate allocation

The rate allocation methods used in earlier works [14,22, 6,34] consider a stationary channel with known loss probability. In this case the density function of the number of lost packets in a block becomes $\hat{\pi}(p)=\delta(p-\hat{p}_\omega)$, where δ is the Dirac delta. Thus the optimal ratio of redundancy is given by

$$\beta_c^*=\arg\min_\beta D(R_a,\hat{p}_\omega,\hat{p}_{\omega|\omega},\beta), \quad (4)$$

and can be found by evaluating the the mean distortion for all possible redundancy values. We will refer to this method later as the crisp rate allocation (CRA) method. To evaluate the performance of CRA the corresponding weighted mean distortion will be calculated using (2) and denoted by D_{CRA}.

Weighted optimal rate allocation

The rate allocation method that achieves the minimal mean distortion in the sense of subsection 2.1 is the one that finds the optimal value of β which minimizes (2),

$$\beta_w^*=\arg\min_\beta\int_{p\in\hat{\Omega}_\alpha} D(R_a,p,\hat{p}_{\omega|\omega},\beta)\hat{\pi}(p)dp. \quad (5)$$

The corresponding weighted mean distortion can be calculated using (2) and we denote it by D_{WOA}. We will refer to this method later as the weighted optimal rate allocation (WOA) method. Note, that the WOA method is computationally expensive, since the distortion has to be calculated for all possible future channel states.

The CRA method is a special case of the WOA method: by taking the density function $\hat{\pi}(p)=\delta(p-\hat{p}_\omega)$ in (5) we get (4).

Min-max-α rate allocation

The rationale behind the proposed min-max-α rate allocation is to minimize the maximum of the distortion over the channel states that have a certain probability of occurrence. It will perform worse than optimal whenever the channel state is close to the expected behavior, is robust however to fluctuation and to sudden changes of the channel state. In certain cases, like for example in the case of motion compensated video, the effects of decreased mean distortion below a certain value are not noticed by human observers [38, 40], while an improvement in the worst case performance remains noticeable. The optimal value of the ratio of redundancy β in the min-max-α sense is

$$\beta_\alpha^*=\arg\min_\beta\max_{p\in\hat{\Omega}_\alpha} D(R_a,p,\hat{p}_{\omega|\omega},\beta) \quad (6)$$

We denote the corresponding weighted mean distortion by $D_{MMA-\alpha}$, which can be calculated using (2). We will refer to this method later as the min-max-α rate allocation (MMA-α) method. Calculation of the optimal ratio of redundancy in the min-max-α sense is not more computation intensive than that of the CRA method. If the distortion rate function $D(R,p_\omega,p_{\omega|\omega},\beta)$ is monotonically increasing in p_ω for any value of β and $p_{\omega|\omega}$, then it is enough to find the optimal value of β for $max_p\{p\in\hat{\Omega}_\alpha\}$. Practical distortion-rate functions have this property, as increasing the loss probability does, ceteris paribus, increase the distortion. If the mean and the median of the distribution $\hat{\pi}(p)$ are equal, then the CRA method is a special case of the MMA-α rate allocation method, with $\alpha=100$.

3 Description of the transmission channel

3.1 Channel model

We assume that the transmission channel, i.e., the packet loss process on the transmission path from the sender to the receiver, can be described with a two state Markovian model, often referred to as the Gilbert model [16]. The Gilbert model is widely used for the design and evaluation of error control solutions for channels with correlated losses due to its simplicity and analytical tractability [14,21]. The Gilbert model's capability to model the packet loss process has been investigated both via measurements [21,42,45] and analytically [7,43]. Around half of the measured traces of Internet end-to-end packet losses reported in [21,42,45] could be modeled sufficiently well with a two state Gilbert model, while analytical results show that the Gilbert model can capture the loss process of a single multiplexer if the level of statistical multiplexing is high.

The Gilbert model is a two state time-discrete Markovian model, where state 0 corresponds to the good state of the channel, i.e., the packet is received, and state 1 to the bad state of the channel, i.e., the packet is lost. Let us denote the transition probability from state 0 to state 1 by p and from state 1 to state 0 by q. Given the values of p_ω and $p_{\omega|\omega}$ the parameters of the Gilbert model can be set as $q = 1 - p_{\omega|\omega}$ and $p = qp_\omega/(1-p_\omega)$. The Gilbert model can then be used to calculate the probability of j losses in a block of n packets as shown in [10].

The number of packets between two loss events are i.i.d. random variables with probability mass function

$$P(s=i) = \begin{cases} 1-q & i=0 \\ q(1-p)^{i-1}p & i>0 \end{cases}, \quad (7)$$

so that the expected value of the number of packets between two loss events is $E[s] = q/p$, and its variance

$$Var[s] = \frac{q(2-p)-q^2}{p^2}. \quad (8)$$

In the paper we use the Gilbert model as Ψ, and the set of parameters $\mathcal{P}$ is $\{p_\omega, p_{\omega|\omega}\}$.

3.2 Estimation of the channel parameters

Adaptive applications that want to optimize their performance for the actual network conditions have to assume a channel model $\hat{\Psi}$ and have to estimate the state of the transmission channel $\hat{\mathcal{P}}$. If the assumed channel model is the Bernoulli model, then the channel parameter to be estimated is $\hat{p}_\omega$. If the assumed channel model is the Gilbert model, then the channel parameters to be estimated are $\hat{p}_\omega$ and $\hat{p}_{\omega|\omega}$. In particular, for the above rate allocation methods the application needs an estimate of the loss probability and the conditional loss probability. In the following we discuss the estimation of these parameters.

3.2.1 Estimation of p_ω

On a stationary channel the estimate $\hat{p}_\omega$ can get arbitrarily close to the stationary loss probability p_ω by calculating the average over a sufficiently long interval. Estimation methods to be used on a non-stationary channel however have to follow the changing state of the channel, and thus averaging has to be done over a short interval, which in turn results in a higher variance of the estimate. The methods proposed in the literature to solve the problem of loss probability estimation can be classified into two groups. Methods belonging to the first group count the number of lost packets in a loss window, e.g. the dynamic loss window method [12]. Methods belonging to the other group measure the packet loss intervals, i.e., the number of received packets between two loss events, and calculate the estimated packet loss interval based on a number of past values. The exponentially weighted moving average and the adaptive loss interval methods (ALI) [12,45] belong to this group. The authors in [45] compared the performance of several estimators using measured traces and concluded that none of them performs significantly better than the others.

In this analysis we will use the ALI method due to its simplicity, even though any other on-line estimation method could be used instead. The estimation of the average packet loss interval is done using the following formula

$$\hat{s} = \frac{\sum_{i=1}^{n} w_i s_i}{\sum_{i=1}^{n} w_i}, \quad (9)$$

where s_i is the i^{th} most recent measured packet loss interval and w_i are the corresponding weighting coefficients defined as

$$w_i = \begin{cases} 1 & 1 \le i \le n/2 \\ \frac{i-n/2}{n/2+1} & n/2 < i \le n \end{cases}, \quad (10)$$

and $\hat{p}_\omega$ is defined as

$$\hat{p}_\omega = \frac{1}{\hat{s}+1}. \quad (11)$$

As we are not aware of any analytical evaluation of the ALI method, we present a brief evaluation here. Assuming that the s_i are i.i.d random variables with mean $1/p_\omega - 1$ and variance σ^2, the variance of $\hat{s}$ is

$$Var[\hat{s}] = Var[\frac{\sum_{i=1}^{n} w_i s_i}{\sum_{i=1}^{n} w_i}] = \frac{\sigma^2 \sum_{i=1}^{n} w_i^2}{(\sum_{i=1}^{n} w_i)^2} = \frac{\sigma^2 8(4n+7)}{27n(n+2)}. \quad (12)$$

Comparing this to the variance of the minimum variance unbiased estimator (i.e., the unbiased estimator with the least possible variance), which is σ^2/n, the ratio of the variances is

$$\frac{Var[\hat{s}]}{\sigma^2/n} = \frac{8(4n+7)}{27(n+2)}, \quad (13)$$

which is bounded from above by $39/29$ and goes to $32/27$ as n goes to infinity, thus the variance of the ALI method is close to the minimal.

If the s_i are correlated, and the covariance $cov(s_i,s_j) = E[s_i s_j] - E[s_i]E[s_j]$ is a function of $|i-j|$ only (e.g. the process is second order stationary), i.e., $\rho_{i,j} = cov(s_i,s_j)/\sigma^2 = \rho_{|i-j|}$, then the variance of the estimator is

$$Var[\hat{s}] = Var\left[\frac{\sum_{i=1}^{n} w_i s_i}{\sum_{i=1}^{n} w_i}\right] \quad (14)$$

$$= \frac{\sigma^2(\sum_{i=1}^{n} w_i^2 + 2\sum_{i=2}^{n}\sum_{j=1}^{i-1} w_i w_j \rho_{i-j})}{(\sum_{i=1}^{n} w_i)^2}, \quad (15)$$

which is always greater than the one given in (12). The estimated value of the packet loss interval $\hat{s}$ is approximately normally distributed with mean and variance given above due to the central limit theorem. These analytical results are in accordance with simulation results shown in [11].

3.2.2 Estimation of $p_{\omega|\omega}$

The estimation of the conditional loss probability does not have much coverage in the literature, presumably because of the difficulties regarding the collection of a sufficiently large sample size on a non-stationary channel. For this reason instead of attempting to estimate the conditional loss probability we investigate in Section 4.2 how different assumptions on the conditional loss probability influence the efficiency of the rate allocation methods.

3.2.3 Estimation of Ω_α

Given $\hat{p}_\omega$ (i.e., $\hat{s}$), $\hat{p}_{\omega|\omega}$ and the estimated channel model $\hat{\Psi}$, the application can derive the confidence interval Ω_α used in the rate allocation methods shown in Subsection 2.2. E.g., if $\hat{\Psi}$ is the Gilbert model, then $\hat{\mathcal{P}}$ is $\{\hat{p}_\omega, \hat{p}_{\omega|\omega}\}$, and the probability of j losses in a block of n packets can be calculated as shown in [10].

4 Min-max-α for FEC and MDC

In this section we illustrate the use of min-max-α via two simple examples. We apply it to set the redundancy rate for media-dependent forward error correction (FEC) and to set the side and central distortions for multiple description coding (MDC). First we give the distortion-rate functions $D(R_a, p_\omega, p_{\omega|\omega}, \beta)$ for FEC and MDC. These functions are then used to determine the optimal redundancy and the minimum achievable distortion with the different rate allocation methods as given by (4)-(6). We use two metrics for evaluation, the mean distortion and the variance of the distortion.

4.1 Distortion-rate bounds for FEC and MDC

In order to keep the examples simple, we consider a memoryless Gaussian source with unit variance and use the squared distortion measure, which is the most common distortion measure. The reasons for the choice of the Gaussian source are twofold. Gaussian mixture models using the weighted sum of Gaussian densities have lately found application in speech coding [20,36]. Furthermore, it is known that DCT transform coefficients in motion compensated pictures have a generalized Gaussian distribution [19]. The distortion-rate function for a Gaussian source with unit variance and squared distortion measure in the absence of losses (i.e., due to quantization) is given as

$$D(R) = 2^{-2R}, \tag{16}$$

where R is the code rate and $D(R)$ is the distortion [18].

To calculate the variance of the distortion we have to calculate its second moment $E[\tilde{D}^2]$, since $V[\tilde{D}] = E[\tilde{D}^2] - D^2$, where D is the mean distortion. The coefficient of variation is then the square root of the ratio of the variance and the mean $CoV[\tilde{D}] = \sqrt{V[\tilde{D}]/D^2}$. In the following we give a lower bound on the CoV of the distortion for a Gaussian source with zero mean and the squared distortion measure.

Proposition 1 *Given a Gaussian random variable with 0 mean and variance σ^2 with pdf p, a squared difference distortion measure $d(x,y) = L(x-y) = (x-y)^2$, then*

$$CoV[\tilde{D}] \geq \sqrt{\frac{4e\Gamma\left(\frac{3}{4}\right)^4}{\pi^2} - 1}.$$

The proof of the proposition can be found in the Appendix.

The bound is independent of the code rate R, and can easily be checked that it holds for $R = 0$, in which case the second moment of the distortion is given by the fourth central moment of the Gaussian r.v., i.e., $E[\tilde{D}^2] = 3\sigma^4$, and thus the variance of the distortion is $V[\tilde{D}] = 2\sigma^4$, and $CoV[\tilde{D}] = \sqrt{2}$. Throughout the paper we will use this value as the CoV of the distortion. We introduce the notation $\varsigma = CoV[\tilde{D}]^2 + 1 = 3$, so that given the mean distortion for code rate R, we estimate the second moment of the distortion as

$$E[\tilde{D}^2] = \varsigma D(R)^2. \tag{17}$$

In the presence of information loss on the transmission channel, the distortion depends not only on the code and redundancy rate but also on the error control scheme applied. In the following we discuss the distortion-rate characteristics of FEC and MDC for total available rate R_a.

4.1.1 Distortion-rate bounds for FEC

In the following analysis we consider media-dependent FEC, proposed by the IETF and implemented in Internet audio tools like Rat [28] and Freephone [15], and in the H.264 video coding standard. The idea behind media-dependent FEC is to add a redundant copy of the original packet to one of the subsequent packets. The redundant packet is heavily compressed, so that quality reconstructed from the redundant packet is low, but still better than when there is nothing to play out. Proposed ways to improve the performance of the scheme are to increase the offset between the original packet and the redundant one [23] and to send multiple redundant copies in subsequent packets [5]. The performance of this FEC scheme has been evaluated via simulations in [27] and analytically in [3,4,6]. The results show that if the ratio of the traffic implementing FEC is small, streams can benefit from using FEC.

In the case of media-dependent FEC the first and the subsequent $\nu - 1$ (redundant) descriptions are encoded independent from each other, and thus, if any of them is received, its distortion is given by (16). We denote the rate allocated to the primary encoding with R_1^{FEC} and the rate allocated to the k^{th} (redundant) copy with R_k^{FEC}. The redundancy ratio introduced by the FEC is then $\beta = \sum_{k=2}^{\nu} R_k^{FEC}/R_1^{FEC}$, where $R_1^{FEC} = R_a/(1+\beta)$. We denote the distortion of the original encoding with D_1^{FEC}, while the distortion of the k^{th} redundant, low quality encoding with D_k^{FEC}. In the case, when both the primary and some redundant encodings are received, the redundant encodings can not be used to reduce the distortion of the original encoding. In our analysis we consider the case $\nu = 2$, as this is most commonly used.

The mean distortion of the FEC scheme with two descriptions can be calculated as the weighted sum of the distortions of the cases when both descriptions are received, only one of them is received or none of them is received

$$D^{FEC}(R_a, p_\omega, p_{\omega|\omega}, \beta) = (p_{\alpha\alpha}(n) + p_{\alpha\omega}(n))D_1^{FEC} + p_{\omega\alpha}(n)D_2^{FEC} + p_{\omega\omega}(n), \quad (18)$$

where $p_{\alpha\alpha}(n)$, $p_{\alpha\omega}(n)$, $p_{\omega\alpha}(n)$, $p_{\omega\omega}(n)$ are the probabilities of the joint loss or reception of two packets n packets apart, and can be calculated based on p_ω and $p_{\omega|\omega}$. The second moment of the distortion can be calculated similarly

$$E[\tilde{D}^{FEC}(R_a, p_\omega, p_{\omega|\omega}, \beta)^2] = \{(p_{\alpha\alpha}(n) + p_{\alpha\omega}(n))(D_1^{FEC})^2 + p_{\omega\alpha}(n)(D_2^{FEC})^2 + p_{\omega\omega}(n)\}\varsigma. \quad (19)$$

4.1.2 Distortion-rate bounds for MDC

MDC addresses the problem of joint source and channel coding. Originally designed for the transmission of multiple descriptions of a single source over independent channels, it has been rediscovered recently for use in packet switched networks [37]. In the case of MDC, several coded descriptions of the same source are sent over different channels. If only one of the descriptions is received, it is used for reconstruction with a certain accuracy. If more than one descriptions are received, then the information from the other descriptions can be used to enhance the accuracy (in contrast to FEC, where the redundant copy can not be used to enhance quality). In a packet switched network, instead of using separate channels, one can put the different encodings into different packets and send them in subsequent packets, similarly to the case of FEC. In the general case, the amount of information sent over the separate channels (packets) can be different; however in single-path packet networks, which offer identical treatment to all packets, it can be shown that balanced MDC, i.e., the one sending the same amount of information in all packets, is optimal [18].

In the case of MDC we consider the (balanced) two-channel case. We denote the rate allocated to individual descriptions by $R_1^{MDC} = R_a/2$. The distortion when both descriptions are received, called the central distortion, is denoted by D_0^{MDC} and the distortion if only one of the descriptions is received, called the side distortion, is denoted by D_1^{MDC}. The distortion rate bounds for the 2-channel MDC are [25]

$$D_1^{MDC} \geq 2^{-2R_a/2} \quad (20)$$

$$D_0^{MDC}(R_a, D_1^{MDC}) \geq 2^{-2R_a}\gamma(R_a, D_1^{MDC}), \quad (21)$$

where $\gamma(R_a, D_1^{MDC}) = 1$ if $2D_1^{MDC} > 1 + D_0^{MDC}$ and

$$\gamma(R_a, D_1^{MDC}) = \frac{1}{1 - \{(1 - D_1^{MDC}) - \sqrt{(D_1^{MDC})^2 - 2^{-2R_a}}\}^2}$$

otherwise. Equation (21) shows that if the side distortion is not large then the central distortion is higher than the distortion rate minimum. We can interpret $2^{-2R_a/2}/D_1^{MDC}$ as the ratio of redundancy β, the higher its value, the more robust MDC will be to errors. For a primer on rate distortion theory and multiple description coding see [18].

The mean distortion and the second moment of the distortion of MDC can be calculated similarly to that of FEC

$$D^{MDC}(R_a, p_\omega, p_{\omega|\omega}, \beta) = p_{\alpha\alpha}(n)D_0^{MDC}(D_1^{MDC}) + (p_{\alpha\omega}(n) + p_{\omega\alpha}(n))D_1^{MDC} + p_{\omega\omega}(n), \quad (22)$$

$$E[\tilde{D}^{MDC}(R_a, p_\omega, p_{\omega|\omega}, \beta)^2] = \{p_{\alpha\alpha}(n)D_0^{MDC}(D_1^{R_a,MDC})^2 + (p_{\alpha\omega}(n) + p_{\omega\alpha}(n))(D_1^{MDC})^2 + p_{\omega\omega}(n)\}\varsigma. \quad (23)$$

4.2 Performance evaluation considering stationary channel

In this subsection we discuss how the rate allocation methods shown in Section 2 behave in a stationary environment used in combination with FEC and MDC.

We evaluate the performance of the CRA and the MMA-α methods for α=1,5,10,25 by comparing the mean and the standard deviation of the weighted distortions at the optimal level of redundancy. We show distortion values relative to the ones achievable with the WOA method. WOA would minimize the weighted distortion in all stationary cases, but it is not feasible for real-time applications due to its complexity.

Both for FEC and MDC we consider the case of two descriptions, and the spacing between the packets $n = 1$. For brevity we will show figures for $D(R_a) = 10^{-4}$ only. The choice of a fixed rate does not limit the validity of our results, similar results can be obtained for other values of R_a. If the application uses some form of rate control, and hence R_a varies over time as a function of the packet loss probability, then the optimal rate of redundancy has to be chosen with respect to the actual available rate. Often we use FEC as an example, the results with MDC are similar, as it is shown in some cases.

We consider the transmission of a sequence of packets through a lossy transmission path. Applications set the redundancy rate based on the estimated loss parameters and the estimated channel model. We calculate using Ψ and $\mathcal{P}$ the corresponding weighted mean distortion (2) and the standard deviation of the distortion (3).

We consider three cases. First we assume that the application has perfect channel information. Then, we assume, that the application has a perfect estimate on the loss probability but not on the conditional loss probability. In both of these cases $\hat{p}_\omega = p_\omega$. Finally, we consider the case when the application uses the ALI method to estimate the loss probability.

4.2.1 Perfect channel information

First we consider the case when the application has a perfect estimate of the channel state, i.e., the loss probability $\hat{p}_\omega = p_\omega$ and the conditional loss probability $\hat{p}_{\omega|\omega} = p_{\omega|\omega}$. The channel model $\hat{\Psi}$ used by the application is the Gilbert model, thus $\hat{\Psi} = \Psi$. Figure 1 shows the mean distortions achieved using the MMA-1 and the CRA methods divided by that of the WOA method. The figure shows that in terms of mean distortion the CRA method is optimal ($D^{FEC}_{CRA}/D^{FEC}_{WOA} \approx 1$). The MMA-1 method performs worse than the CRA method, especially if both the loss probability and the conditional loss probability are low. Figure 2 shows the the standard deviations of the distortions achieved using the MMA-1 and the CRA methods divided by that of the WOA method. As can be seen, the MMA-1 method has a significantly lower standard deviation than the CRA method, thus MMA-1 effectively decreases the fluctuations of the perceived quality.

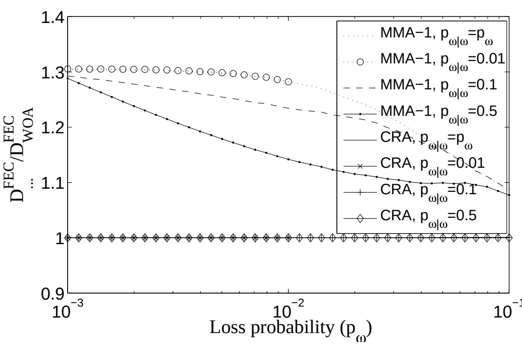

Fig. 1 Mean distortions vs. p_ω for the MMA-1 and the CRA methods compared to the WOA method.

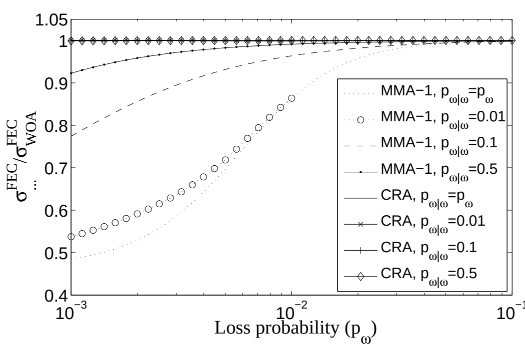

Fig. 2 Standard deviations vs. p_ω for the MMA-1 and the CRA methods compared to the WOA method.

4.2.2 Partial channel information

In this subsection we consider the case when the application has a perfect estimate of the loss probability, $\hat{p}_\omega = p_\omega$, but has no information about the conditional loss probability. First we consider the case when the channel model $\hat{\Psi}$ used by the application is the Bernoulli model, i.e., the application assumes independent losses. The Bernoulli-model is a lower order model than the actual channel model Ψ, the Gilbert-model. Fig. 3 shows the distortions obtained by the MMA-1 and the CRA methods using the Bernoulli channel model divided by the distortion obtained by the WOA method with perfect channel state information (i.e., aware of $p_{\omega|\omega}$, denoted by WOA-i in the figures). The figure shows significant difference compared to Fig. 1 when the loss process is highly correlated, for modest levels of correlation, i.e., below a conditional loss probability of 0.1 there is practically no difference.

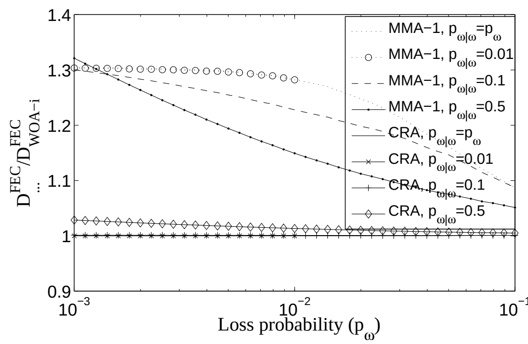

Fig. 3 Mean distortions vs. p_ω and $p_{\omega|\omega}$ for the MMA-1 and CRA methods with partial channel information compared to the WOA method with perfect channel information.

To see how the relationship of $\hat{p}_{\omega|\omega}$ and $p_{\omega|\omega}$ influences the results, we now consider the case when the channel model $\hat{\Psi}$ used by the application is the Gilbert model, and the value $\hat{p}_{\omega|\omega}$ is chosen by the application between p_ω and 0.9. The loss probability is $p_\omega = 10^{-3}$. Fig. 4 shows the ratio of the mean distortion for the CRA method choosing $\hat{p}_{\omega|\omega}$ and that of the CRA method with perfect channel information (denoted by CRA-i), as well as the ratio of the mean distortion for the MMA-1 method choosing $\hat{p}_{\omega|\omega}$ and that of the MMA-1 method with perfect channel information (denoted by MMA-1-i). The figure shows that overestimating $p_{\omega|\omega}$ results in poorer performance than underestimating it. Comparing the curves corresponding to the two rate allocation methods we see that the CRA method is more sensitive to the misestimation of $p_{\omega|\omega}$ than the MMA-1 method. Fig. 5 shows the ratios of the standard deviations for the CRA and MMA-1 methods for the same scenario. The conclusions are similar, the CRA method is more sensitive to the misestimation of $p_{\omega|\omega}$ than the MMA-1 method, and overestimating $p_{\omega|\omega}$ results in higher variance than underestimating it.

4.2.3 Measured channel information

In this subsection we consider the case when the application uses the ALI method to estimate the loss probability, and

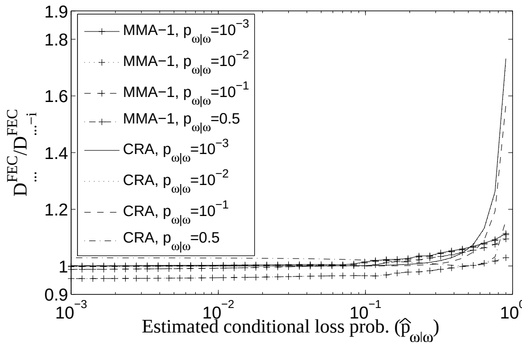

Fig. 4 Mean distortions vs. $\hat{p}_{\omega|\omega}$ for the MMA-1 and CRA methods with partial channel information compared to the case with perfect channel information.

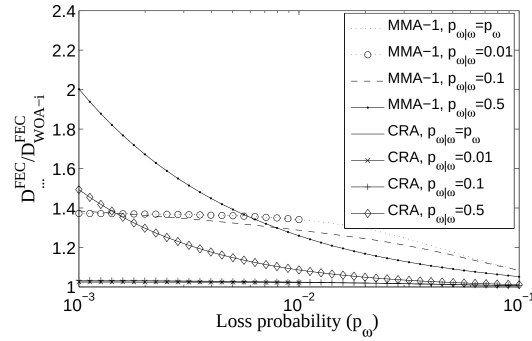

Fig. 6 Mean distortions vs. p_ω for the MMA-1 and the CRA methods compared to the WOA method with perfect channel information.

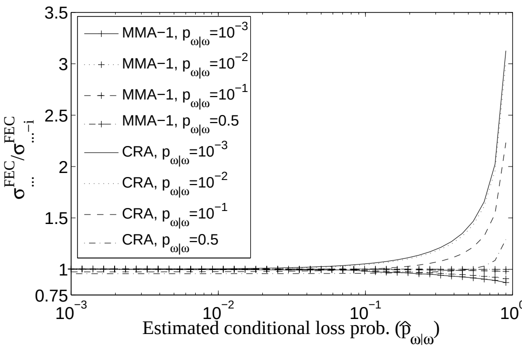

Fig. 5 Standard deviations vs. $\hat{p}_{\omega|\omega}$ for the MMA-1 and the CRA methods with partial channel information compared to the case with perfect channel information.

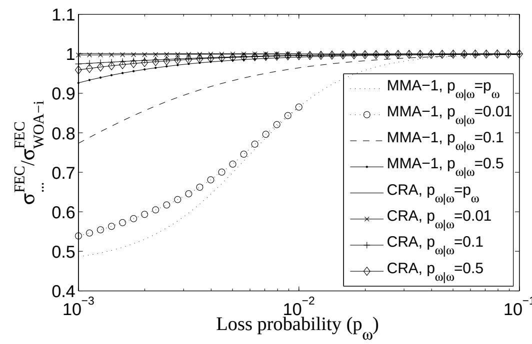

Fig. 7 Standard deviations vs. p_ω for the MMA-1 and the CRA methods compared to the WOA method with perfect channel information.

hence the estimate $\hat{p_\omega} = 1/(1+\hat{s})$ is a random variable. We calculate the 99 percent confidence interval of $\hat{p}_\omega$ based on Ψ (the Gilbert model), $\mathcal{P}$ (p_ω and $p_{\omega|\omega}$) and (12). Then we calculate the optimal redundancy rate using the CRA, the WMA and the MMA-α methods corresponding to the $\hat{p}_\omega$ values in the confidence interval. Based on $\mathcal{P}$, (2) and (3) we calculate the distortions corresponding to the optimal redundancy rate, and calculate the weighted average of these distortions, where the weights are the probabilities of the occurrence of $\hat{p}_\omega$. We assume that in lack of a reliable estimate of $\hat{p}_{\omega|\omega}$ the application assumes independent losses. Based on the observations of the previous subsection this behavior is the most robust for FEC and MDC, while it will not influence the results significantly.

Fig. 6 shows the mean distortion achieved using the MMA-1 and the CRA methods based on the estimated channel parameters divided by the mean distortion achieved using the WOA method based on the correct channel parameters. As an effect of the estimation of p_ω the mean distortions increase especially for correlated losses. Fig. 7 shows the standard deviation achieved using the MMA-1 and the CRA methods based on the estimated channel parameters divided by the standard deviation achieved using the WOA method based on the correct channel parameters.

Next we compare results obtained with the MMA-α method for different values of α. Fig. 8 shows the mean distortions of the MMA-α methods and the CRA method divided by that of the WOA method with perfect channel information for the case of independent losses. Fig. 9 shows the standard deviations for the same scenario. The figures show the trade-off between the mean distortion and the standard deviation, and that α has to be selected according to the application's requirements on these values.

To compare the performance of FEC and MDC we show the ratio of their mean distortions using the MMA-1 and the CRA method in Fig. 10. The curves corresponding to the same conditional loss probability but different rate allocation method show similar charateristics. The results, MDC is always better in terms of mean distortion, coincide with those obtained for the long term average in [6]. In Fig. 11 we show the ratios of the standard deviations of FEC and MDC for the CRA and the MMA-1 methods. The figure shows that MDC gives a lower variance than FEC independent of the rate allocation method used. Comparing the CRA and MMA-1 methods shows that while the ratio of the mean distortions is slightly higher with the MMA-1 method, the ratio of the standard deviations is slightly lower.

Finally we evaluate how the mean distortion evolves in the case of misestimation of p_ω. Fig. 12 shows the ratio of

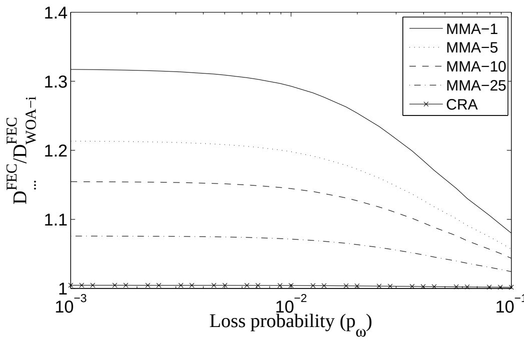

Fig. 8 Mean distortions vs. p_ω for the MMA-α and CRA methods compared to the WOA method with perfect channel information.

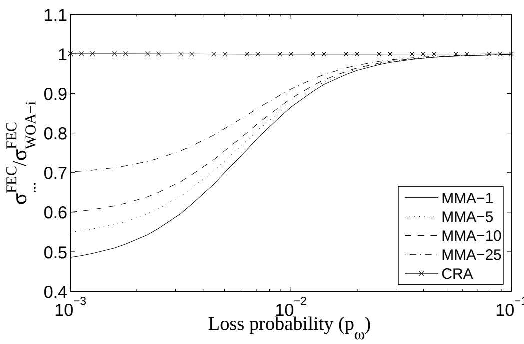

Fig. 9 Standard deviations vs. p_ω for the MMA-α and CRA methods compared to the WOA method with perfect channel information.

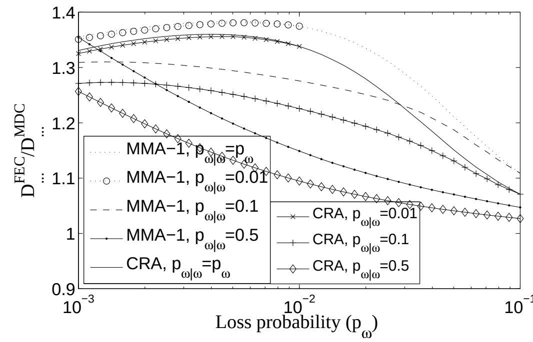

Fig. 10 Ratio of the mean distortions vs. p_ω of FEC and MDC using the the MMA-1 and the CRA method.

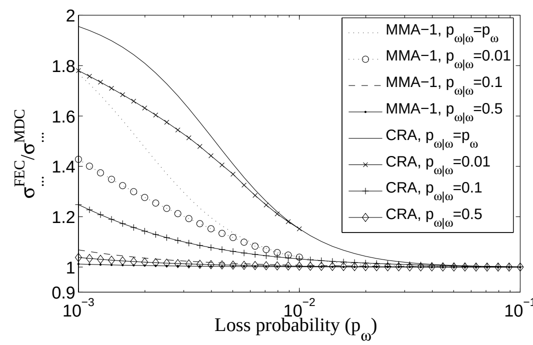

Fig. 11 Ratio of the standard deviations vs. p_ω of FEC and MDC using the MMA-1 and the CRA methods.

the mean distortions for FEC using MMA-1 and CRA as a function of p_ω and $p_{\omega|\omega}$ for $\hat{p}_\omega$= $\hat{p}_{\omega|\omega}$=0.001. The figure shows that for large estimation errors MMA-1 outperforms CRA in terms of mean distortion. We observe the biggest difference in the case of independent losses, the difference decreases as the correlation between losses increases. Fig. 13 shows the ratio of the standard deviations for the same scenario. Based on these figures we conclude that in the presence of estimation errors MMA performs better than CRA.

4.3 Performance evaluation considering non-stationary channel

In this section we evaluate how the rate allocation methods shown in Section 2 behave on a non-stationary channel. To study the robustness of the rate allocation methods we use a step increase function of the packet loss probability. For the calculations we used packet level simulations, where each packet corresponds to one piece of information generated by the sender application. We generate the packet loss process according to the Gilbert model. In the considered scenario $p_\omega = 0.003$ and $p_{\omega|\omega} = 0.02$ initially, after the n_1^{st} packet they increase to $p_\omega = 0.05$ and $p_{\omega|\omega} = 0.12$, and after the n_2^{nd} packet they decrease to $p_\omega = 0.005$ and $p_{\omega|\omega} = 0.03$. These particular values for loss probability were taken from measurements shown in [42]. We consider two scenarios, in the first one the application uses the ALI method with $n = 8$ as proposed in [12] to estimate the channel state, and $n_1 = 200$, $n_2 = 600$, while in the second one the application uses the ALI method with $n = 32$ and $n_1 = 200$, $n_2 = 1000$. Based on the estimated channel state the application uses the CRA or the MMA-1 method to set the redundancy rate. The calculated redundancy rate is used by the sender until the next update of the estimate of the loss probability. Figs. 14 and 15 show the averages of the distortions of 5000 simulations respectively.

While CRA achieves a lower mean distortion in the stationary state of the channel, the sudden increase of the stationary loss probability affects its performance more than that of MMA-1: MMA-1 adjusts smoothly to the new channel conditions. We can also observe the smaller variance of the distortion using the MMA-1 method. Comparing MDC and FEC, we see that MDC achieves a lower mean distortion and variance throughout the whole simulation, which is in accordance with results obtained in Section 4.2. Comparing the two figures we conclude that a slower adaptation (increasing value of n) to the changing channel conditions requires more robustness from the rate allocation method to avoid degradation of the quality.

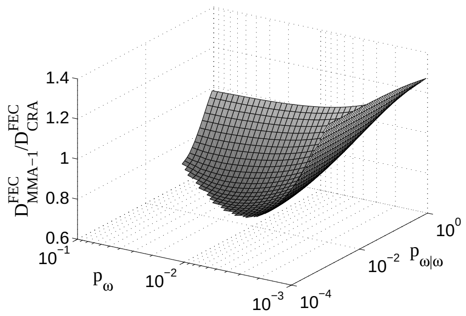

Fig. 12 Ratio of the mean distortions vs. p_{ω} and $p_{\omega|\omega}$ for the MMA-1 and CRA methods when $\hat{p}_{\omega} = 10^{-3}$.

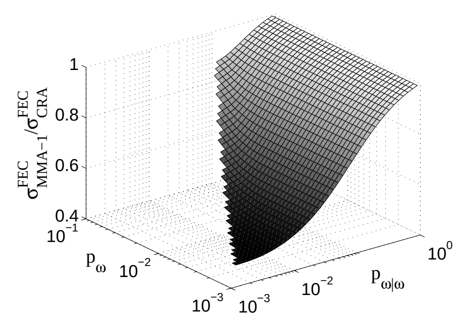

Fig. 13 Ratio of the standard deviations vs. p_{ω} and $p_{\omega|\omega}$ for the MMA-1 and CRA methods when $\hat{p}_{\omega} = 10^{-3}$.

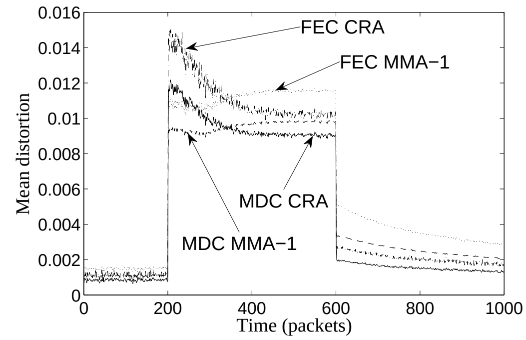

Fig. 14 Mean distortion vs. time for the MMA-1 and CRA methods with MDC and FEC, $n = 8$, $n_1 = 200$, $n_2 = 600$

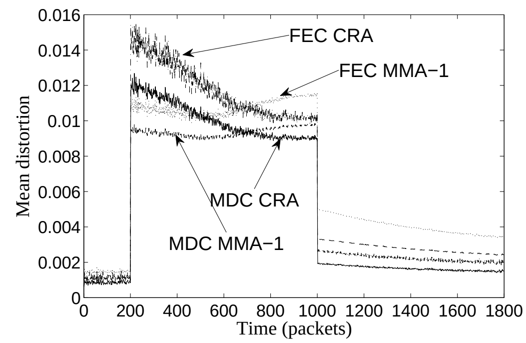

Fig. 15 Mean distortion vs. time for the MMA-1 and CRA methods with MDC and FEC, $n = 32$, $n_1 = 200$, $n_2 = 1000$

5 Min-max-α for H.264 coded video

In this section we show how the min-max-α rate allocation method can be used to tune the error resilience of the H.264 video coder to increase the robustness of the video stream to fluctuations and sudden changes of the channel quality.

5.1 Error resilience in H.264 video

The H.264/AVC video coding standard is the newest in the line of video coding standards, suitable for both low and high quality video communications. For an overview of the standard see [41]. Compared to earlier standards it has a large set of error resilience features as it was developed to be used in error-prone networks. While many of the features, like for example data partitioning, are suitable only for high bitrate applications, others, like periodic intra updates can always be used. Periodic intra updates are an efficient way to combat inter-frame error propagation without the need for large intra coded frames. When periodic intra updates are used, a small portion of each frame, a certain number of macroblocks, is encoded in intra mode. Intra coded macroblocks do not depend on previous frames, and thus stop the propagation of errors. The higher the portion of intra coded macroblocks, denoted by β, the more error resilient will be the video stream. On the other hand, increasing the ratio of intra coded macroblocks increases the encoding distortion at a given available bitrate. The tradeoff between the ratio of intra coded macroblocks, the encoding rate and the mean distortion in the absence of losses is shown in Fig. 16 for the Foreman sequence at 12.5 frames per second in QCIF format. The measure used is the peak signal to noise ratio defined as $PSNR = 10log_{10}(255^2/D)$, where D is the mean distortion per pixel. To conduct the simulations we used the JVT test model encoder and decoder with slight modifications in the decoder to make it able to cope with losses.

We performed simulations over stationary channels modeled with the Gilbert model with different stationary and conditional loss probabilities to determine the distortion-rate curve in the presence of losses. Fig. 17 shows the PSNR as a function of β for seven different channels for the Foreman sequence encoded at 128 kbps, each value is the average of 40 simulations. The figure shows that for each stationary channel there is a particular value of β which minimizes the mean distortion. The values obtained from the simulations can be used to parametrize a distortion rate model of the encoded video, e.g. the one presented in [34],

$$D(R_a, p_{\omega}, \beta) = \frac{\theta(\beta)}{R_a - R_0(\beta)} + D_0(\beta) + p_{\omega}\sigma^2_{u0} \sum_{t=0}^{1/\beta-1} \frac{1-\beta t}{1+\gamma t}.$$

The model can then be used to select the optimal value of β in a changing environment. Fig. 18 shows the CoV of the

PSNR as a function of β, showing that the CoV attains its minimum not necessarily at the same β as the mean.

5.2 Performance on a non-stationary channel

To see how the MMA-α and CRA methods perform on a non-stationary channel we use the step increase function of the loss probability. We set $p_\omega = 0.005$ and $p_{\omega|\omega} = 0.02$ for the first 800 packets and to $p_\omega = 0.03$ and $p_{\omega|\omega} = 0.12$ afterward. We used the Gilbert model to generate 40 loss traces with these parameters. The sender uses the ALI method for loss estimation with $n = 8$ and assumes independent losses.

Fig. 19 shows the PSNR vs. the frame number with the MMA-1, the MMA-25 and the CRA methods for one of the loss traces. Over the 2000 frame interval shown in the figure the mean PSNR of the CRA method is 29.89, the MMA-25 method 29.78 and the MMA-1 method 30.64. The standard deviation of the PSNR is 5.76, 5.42 and 3.44 respectively. Averaged over 40 simulations the mean PSNRs are 30.88, 31.105 and 31.28 respectively, and the averages of the standard deviations are 5.43, 4.91 and 3.53 respectively, which shows the possible benefits of robust rate allocation in a non-stationary environment.

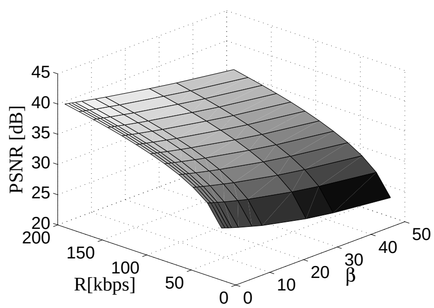

Fig. 16 PSNR vs. bitrate (R) and ratio of I coded macroblocks (β) for the Foreman sequence at 12.5fps, QCIF.

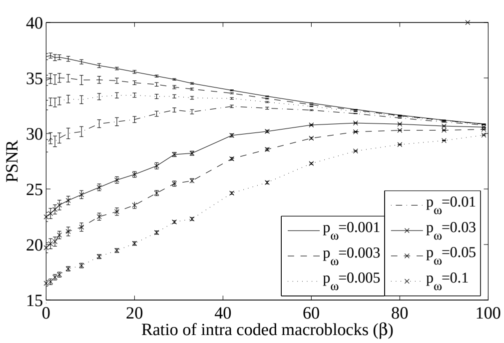

Fig. 17 PSNR vs. ratio of intra coded macroblocks on a stationary channel at 128 kbps at different loss rates.

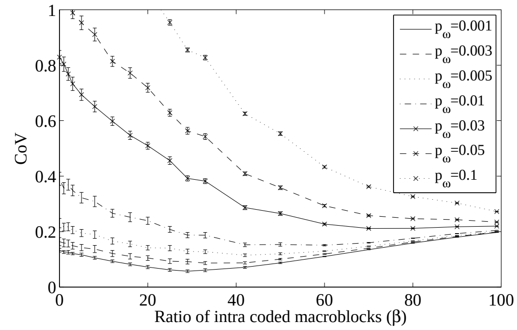

Fig. 18 CoV of the PSNR vs. ratio of intra coded macroblocks on a stationary channel at 128 kbps at different loss rates.

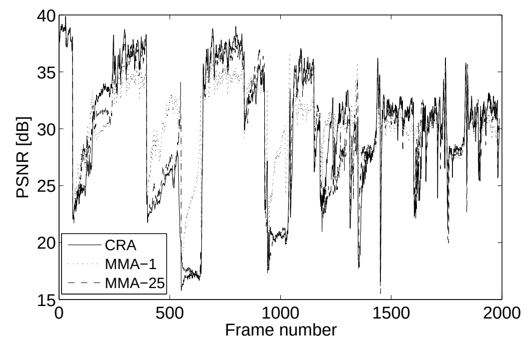

Fig. 19 PSNR on a non-stationary channel at 128 kbps for three rate allocation methods.

6 Conclusions

In this paper we presented a robust rate allocation method for joint source-channel coding that is able to cope with channel estimation errors and changes in the channel state. We applied our method to media-dependent forward error correction and multiple description coding, and compared its performance to the rate allocation method commonly used in the literature. We showed that although in terms of mean distortion min-max-α is suboptimal on a stationary channel, it reduces the variance of the distortion significantly. We studied the effects of short term variations of the channel and estimation errors on the performance of the proposed method. We showed how the proposed method can prevent severe degradations of the quality due to rapid changes on a non-stationary channel. We compared the performance of FEC and MDC and concluded that MDC outperforms FEC under all circumstances, regardless of the rate-allocation method, the errors in the estimate of the channel state and the channel characteristics. We applied the proposed min-max-α rate allocation method to motion compensated video and showed how it improves its error resilience on a non-stationary channel. The proposed min-max-α method can be used in conjunction with different error resilience solutions and objective functions, such as mean opinion score. We believe that due to the characteristics of the human audiovisual perception the proposed robust rate allocation method can provide better perceived quality in an algorithmically efficient way than the approach of minimizing the mean distortion.

Appendix

Proof of Proposition 1: We will prove the proposition with help of the Shannon lower bound, which says that given a continuous random variable, described by a pdf p and difference distortion measure $d(x,y) = L(x-y)$, then

$$R(D) \geq R_{SLB}(D)$$

where

$$R_{SLB}(D) = h(p) + \log a(D) - Db(D)$$

where a(D) and b(D) are solutions to the equations

$$a(D) \int e^{-b(D)L(x)} dx = 1 \tag{24}$$

$$a(D) \int L(x) e^{-b(D)L(x)} dx = D. \tag{25}$$

At a given rate R the lower bound for the second moment of the distortion with respect to the squared distortion measure can not be less than the lower bound for the mean distortion with respect to the quartic distortion measure. Thus we will derive the distortion-rate function for the quartic distortion measure and use it as a lower bound for the second moment of the distortion with respect to the squared distortion measure. ¿From equation (24) we have that

$$a(D) = \frac{b(D)^{1/4}\sqrt{2}\Gamma\left(\frac{3}{4}\right)}{\pi}$$

so that equation (25) evaluates to

$$\frac{1}{4b(D)} = D.$$

Thus using natural logarithms the Shannon lower bound for the quartic distortion measure becomes

$$R_{SLB4}(D) = \frac{1}{2} log(\frac{2\sqrt{\pi e}\sigma^2\Gamma\left(\frac{3}{4}\right)^2}{\sqrt{D}}).$$

Similarly we derive the Shannon lower bound for the squared distortion measure

$$R_{SLB2}(D) = \frac{1}{2} log\frac{\sigma^2}{D}.$$

We can invert both R(D) functions and calculate the ratio

$$\frac{D_4(R)}{D_2(R)^2} = \frac{4e\Gamma\left(\frac{3}{4}\right)^4}{\pi^2},$$

which leads to the proposition.

References

1. Akyildiz, I., Joe, I., Driver, H., Ho, Y.: An adaptive FEC scheme for data traffic in wireless atm networks. IEEE/ACM Trans. Networking **9**(4), 419–426 (2001)
2. Albanese, A., Blömer, J., Edmonds, J., Luby, M., Sudan, M.: Priority encoding transmission. IEEE Trans. Inform. Theory **42**(6), 1737–1744 (1996)
3. Altman, E., Barakat, C., Ramos, V.: On the utility of FEC mechanisms for audio applications. In: Proc. of Quality of Future Internet Services, LNCS 2156, pp. 45–56 (2001)
4. Altman, E., Barakat, C., Ramos, V.M.: Queueing analysis of simple FEC schemes for IP telephony. In: Proc. of IEEE INFOCOM, pp. 796–804 (2001)
5. Bolot, J.C., Fosse-Parisis, S., Towsley, D.: Adaptive FEC-based error control for Internet telephony. In: Proc. of IEEE INFOCOM, pp. 1453–1460 (1999)
6. Dán, G., Fodor, V., Karlsson, G.: Are multiple descriptions better than one? In: Proc. of IFIP Networking, pp. 684–696 (2005)
7. Dán, G., Fodor, V., Karlsson, G.: On the effects of the packet size distribution on FEC performance. Computer Networks **50**(8), 1104–1129 (2006)
8. Eisenberg, Y., Luna, C., Pappas, T., Berry, R., Katsaggelos, A.: Optimal source coding and transmission power management using a min-max expected distortion approach. In: Proc. of Intl. Conf. on Image Proc., pp. 537–540 (2002)
9. Eisenberg, Y., Zhai, F., Pappas, T., Berry, R., Katsaggelos, A.: VAPOR: Variance aware per-pixel optimal resource allocation. IEEE Trans. Image Processing **15**(2), 289–299 (2006)
10. Elliott, E.: Estimates of error rates for codes on burst-noise channels. Bell Syst. Tech. J. **42**, 1977–1997 (1963)
11. Floyd, S., Handley, M., Padhye, J.: A comparison of equation-based and aimd congestion control. http://www.icir.org/tfrc/aimd.pdf (2000). URL www.icir.org/tfrc/aimd.pdf
12. Floyd, S., Handley, M., Padhye, J., Widmer, J.: Equation-based congestion control for unicast applications. In: Proc. of ACM SIGCOMM (2000)
13. Frossard, P.: FEC performances in multimedia streaming. IEEE Commun. Lett. **5**(3), 122–124 (2001)
14. Frossard, P., Verscheure, O.: Joint source/FEC rate selection for quality-optimal MPEG-2 video delivery. IEEE Trans. Image Processing **10**(12), 1815–1825 (2001)
15. García, A.V., Fosse-Parisis, S.: Freephone audio tool. http://www-sop.inria.fr/rodeo/fphone
16. Gilbert, E.: Capacity of a burst-noise channel. Bell Syst. Tech. J. **69**, 1253–1265 (1960)
17. Girod, B.: What's wrong with mean-squared error? Watson, A.B. (Ed), Digital images and human vision, MIT Press, Cambridge, MA (1993)
18. Goyal, V.K.: Multiple description coding: Compression meets the network. IEEE Signal Processing Mag. **18**(5), 74–93 (2001)
19. He, Z., Mitra, S.K.: A unified rate-distortion analysis framework for transform coding. IEEE Trans. Circuits Syst. Video Technol. **11**, 1221–1236 (2001)
20. Hedelin, P., Skoglund, J.: Vector quantization based on gaussian mixture models. IEEE Trans. Speech Audio Processing **8**(4), 385–401 (2000)
21. Jiang, W., Schulzrinne, H.: Modeling of packet loss and delay and their effect on real-time multimedia service quality. In: Proc. of NOSSDAV (2000)
22. Kim, M.Y., Kleijn, B.W.: Rate-distortion comparisons between FEC and MDC based on Gilbert channel model. In: Proc. of IEEE ICON, pp. 495–500 (2003)
23. Kouvelas, I., Hodson, O., Hardman, W., Crowcroft, J.: Redundancy control in real-time Internet audio conferencing. In: Proc. of International Workshop on Audio-Visual Services Over Packet Networks (1997)
24. Mohr, A., Riskin, E., Ladner, R.: Unequal loss protection: graceful degradation of image quality over packet erasure channels through forward error correction. IEEE J. Select. Areas Commun. **18**(6), 819–828 (2000)

25. Ozarow, L.: On a source-coding problem with two channels and three receivers. Bell Syst. Tech. J. **59**, 1909–1921 (1980)
26. Perkins, C., Hodson, O., Hardman, V.: A survey of packet loss recovery techniques for streaming audio. IEEE Network **12**(5), 40–48 (1998)
27. Podolsky, M., Romer, C., McCanne, S.: Simulation of FEC-based error control for packet audio on the Internet. In: Proc. of IEEE INFOCOM, pp. 505–515 (1998)
28. The Multimedia Integrated Conferencing for European Researchers (MICE) Project: RAT: Robust Audio Tool. http://www-mice.cs.ucl.ac.uk/multimedia/software/rat
29. Puri, R., Ramchandran K. Lee, K., Bharghavan, V.: Forward error correction (FEC) codes based multiple description coding for internet video streaming and multicast. Signal Processing: Image Communications **6**(8), 745–762 (2001)
30. Qian, L., Jones, D.: Maxmin disappointment criterion for video broadcasting. In: Proc. of IEEE ICIP (2001)
31. Ramjee, R., Kurose, J., Towsley, D., Schulzrinne, H.: Adaptive playout mechanisms for packetized audio applications in wide-area networks. In: Proc. of IEEE INFOCOM (1994)
32. Rosenberg, J., Qiu, L., Schulzrinne, H.: Integrating packet FEC into adaptive voice playout buffer algorithms on the Internet. In: Proc. of IEEE INFOCOM, pp. 1705–1714 (2000)
33. Servetto, S., Nahrstedt, K.: Video streaming over the public Internet: multiple description codes and adaptive transport protocols. In: Proc. of IEEE ICIP (1999)
34. Stuhlmüller, K., Färber, N., Link, M., Girod, B.: Analysis of video transmission over lossy channels. IEEE J. Select. Areas Commun. **18**(6) (2000)
35. Su, X., Wah, B.: Multidescription video streaming with optimized reconstruction based DCT and neural-network compensations. IEEE Trans. Multimedia **3**(1), 123–131 (2001)
36. Subramaniam, A.D., Gardner, W., Rao, B.: Joint source-channel decoding of speech spectrum parameters over erasure channels using gaussian mixture models. In: Proc. of IEEE Conference on Acoustics, Speech, and Signal Processing, vol. 1, pp. 120–129 (2003)
37. Vaishampayan, V.A.: Design of multiple description scalar quantizers. IEEE Trans. Inform. Theory **39**, 821–834 (1993)
38. Video Quality Experts Group: Final report from the video quality experts group on the validation of objective models of video quality assessment, phase II (2003)
39. Wang, Y., Zhu, Q.: Error control and concealment for video communication: a review. Proc. IEEE **86**(5), 974–997 (1998)
40. Wang, Z., Bovik, A., Sheikh, H., Simoncelli, E.: Image quality assessment: From error visibility to structural similarity. IEEE Trans. Image Processing **13**(4), 600–612 (2004)
41. Wiegand, T., Sullivan, G., Bjontegaard, G., Luthra, A.: Overview of the H.264/AVC video coding standard. IEEE Trans. Circuits Syst. Video Technol. **17**(7) (2003)
42. Yajnik, M., Moon, S., Kurose, J., Towsley, D.: Meaurement and modelling of the temporal dependence in packet loss. In: Proc. of IEEE INFOCOM, pp. 345–352 (1999)
43. Yu, X., Modestino, J., Tian, X.: The accuracy of gilbert models in predicting packet-loss statistics for a single multiplexer model. In: Proc. of IEEE INFOCOM (2005)
44. Zhang, R., Regunathan, S., Rose, K.: Video coding with optimal inter/intra-mode switching for packet loss resilience. IEEE J. Select. Areas Commun. **18**(6), 966–976 (2000)
45. Zhang, Y., Duffield, N., Paxson, V., Shenker, S.: On the constancy of internet path properties. In: Proc. of Internet Measurement Conference, pp. 197–211 (2001)

György Dán received the M.Sc. degree in Computer Engineering from the Budapest University of Technology and Economics, Hungary in 1999 and the M.Sc. degree in Business Administration from the Corvinus University of Budapest, Hungary in 2003. He received his Ph.D. in Telecommunications in 2006 from KTH, Royal Institute of Technology, Stockholm, Sweden. He worked as consultant in the field of access networks, streaming media and videoconferencing 1999-2001. Currently, he is a post-doctoral researcher at KTH, Royal Institute of Technology. His research interests include traffic control and performance evaluation of point-to-point and peer-to-peer multimedia communications.

Viktoria Fodor received her M.Sc. and Ph.D. degrees in computer engineering from the Budapest University of Technology and Economics in 1992 and 1999, respectively. In 1994 and 1995 she was a visiting researcher at Politecnico Torino and at Boston University, where she conducted research on optical packet switching solutions. In 1998 she was a senior researcher at the Hungarian Telecommunication Company. In 1999 she joined the KTH, Royal Institute of Technology, where she now acts as assistant professor. Her current research interests include performance analysis of communication networks and traffic and error control for multimedia communication.

Gunnar Karlsson is a professor since 1998 in the School of Electrical Engineering at KTH, the Royal Institute of Technology, where he is the director of the Laboratory for Communication Networks. He has previously worked for IBM Zürich Research Laboratory and the Swedish Institute of Computer Science (SICS). His Ph.D. is from Columbia University, and the M.Sc. from Chalmers University of Technology. He has been visiting professor at EPFL and ETH Zürich, Switzerland, and the Helsinki University of Technology in Finland.