

SCALABLE MULTIPLE DESCRIPTION VIDEO FOR FAIR STREAMING TO MANY CLIENTS

Jacco R. Taal and R. (Inald) L. Lagendijk
Information & Communication Theory Group
Faculty of Electrical Engineering, Mathematics and Computer Science
Delft University of Technology, The Netherlands

Peer-to-peer networks (P2P) form a distributed communication infrastructure that is particularly well matched to video streaming using multiple description coding. We form M descriptions using MDC-FEC building on a scalable version of the “Dirac” video coder. The M descriptions are streamed via M different application layer multicast (ALM) trees embedded in the P2P network. Client nodes (peers in the network) receive a number of descriptions $m < M$ that is dependent on their bandwidth. In this paper we consider the optimization of the received video qualities, taking into account the distribution of the clients’ bandwidth. We propose three “fairness” criteria to define the criterion to be optimized. Numerical results illustrate the effects of the different fairness criteria and client bandwidth distributions on the rates allocated to the compressed video layers and multiple descriptions.

INTRODUCTION

Peer-to-Peer (P2P) networks and their file swapping P2P applications have become popularized in the past years because of their high download bandwidths and inherent server off-loading. This success has stimulated research into using peer-to-peer networks as infrastructures for streaming video over the Internet[1]. In addition to being clients, the peers in the network then also serve as *application-level* multicast (ALM) nodes. For ALM, no infrastructural changes are required, making fast and flexible deployment possible. The data forwarding model that we consider to be most suitable for video streaming over peer-to-peer networks is known as “bartering”. Chunks of data are exchanged between nodes in a “as fair as possible” manner. The main advantages of Bartering are: 1) leveraging the server load. 2) inherent fault-tolerance due to splitting up data streams.

For P2P video streaming, we apply bartering on multiple descriptions (MD) of the (compressed) video source rather than straightforwardly chopping up a compressed bit stream. In this case, a video encoder generates several (multiple description) streams,

Send correspondence to J.R.Taal@ewi.tudelft.nl. This work has been funded by the Freeband projects CACTUS (www.cactus.tudelft.nl) and I-SHARE (www.freeband.nl)

that are all independently decodable. The different descriptions are put into the different chunks. If more chunks, and hence more descriptions, are received, the decoded video quality improves. The advantages of the MD approach is that video streaming becomes robust against P2P bandwidth variations or failing P2P nodes, which in both cases cause random chunks of data to be unavailable to a video decoder.

In this paper we use the multiple description coding approach MDC-FEC proposed by Puri and Ramchandran[2] (see Figure 1). First, the video is encoded using an M -layer video encoder. Each layer k is encoded at a rate R_k and is then protected by an (M, k) (Reed-Solomon) erasure code. The thus obtained protected data is evenly distributed over the M descriptions. Finally, the individual descriptions form the chunks that are bartered by the P2P nodes. If any m out of M descriptions are available, the first m layers can successfully be decoded, resulting in a quality Q_m .

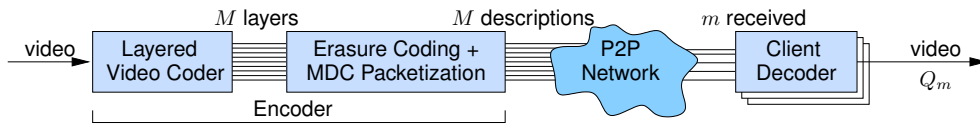


Figure 1: Block Diagram of the multiple description video distribution over a P2P network

In the practical application of the above MDC-FEC approach, several (encoding) parameters have to be chosen, namely the number of layers (and descriptions) and the bit rate at which the individual layers are encoded. In our earlier work [3] we have proposed to use a probabilistic client-bandwidth distribution model that describes clients bandwidth due to physical network limitations (e.g. the various ISDN, cable, ADSL, and LAN connections) and due to network congestion. In this way, the probability for a P2P peer receiving a certain number of descriptions can be calculated as a function of the above mentioned parameters. Using an overall criterion, such as the average compression distortion over all clients, the optimal number of layers/descriptions and the bit rate per layer can be found.

For a discussion of related work we refer to our previous papers [3, 4]. These papers also give an introduction on Application Level Multicast and P2P video streaming. In the remainder of this paper we derive a model for the behavior of our MDC-FEC system. Also we derive an RD-model for the layered Dirac codec. Then we present three different criteria to find the optimal encoding parameters. We will introduce these criteria from the perspective of “fairness” to the clients. Finally we show results of optimizing the fairness criteria with specific settings and assumptions on the client distribution.

MDC VIDEO STREAMING IN PEER-TO-PEER NETWORKS

We use the MDC-FEC structure proposed by Puri *et al.*[2]. Figure 2(a) shows how the layers are grouped and descriptions are generated using erasure codes. The base layer

is encoded at a rate R_1 . The total rate of the first n layers is denoted by R_n . Hence, each successive layer i has a rate $R_i - R_{i-1}$. The rate R_D of each of the M descriptions is then given by:

$$R_D = \sum_{l=1}^M \alpha_l R_l, \text{ where } R_0 = 0, \text{ and} \quad (1)$$

$$\alpha_l = 1/(l(l+1)), \text{ for } l = 1, 2, M-1, \text{ and } \alpha_M = 1/M$$

A node or client in the P2P network with bandwidth R_c receives $m = \lfloor \frac{R_c}{R_D} \rfloor$ descriptions. After decoding these m descriptions, the client experiences a compression distortion of $D(R_1, R_2 \dots, R_m)$ for $m = 1, \dots, M$. In principle, the rate-distortion function for layer m depends on *all* rates R_k of the individual layers $k = 1, \dots, m$ as layered compression induces losses relative to a single layer coder that are dependent on the rates of the (lower) layers. Hence, for each layer we may have a different rates-distortion function. By definition, $D(0) = \sigma_x^2$.

In our work we used *Dirac*, an open-source video coder developed by the BBC[5]. *Dirac* is wavelet-based and its performance is comparable to H.264. In Figure 2 the bit rate – distortion (RD) curve of *Dirac* is shown for the Foreman sequence. We also plotted several RD-curves for two-layer encoding. The branch-point in each curve is the rate R_1 for the first layer. The efficiency loss for the first enhancement layer is clearly visible. For higher layers (not shown) we have observed an insignificant loss of coding efficiency. Consequently, the rate distortion function becomes only dependent on the rate of the base layer R_1 and the rate of the layer m under consideration, i.e., $D(R_1, R_2 \dots, R_m) = D(R_1, R_m)$.

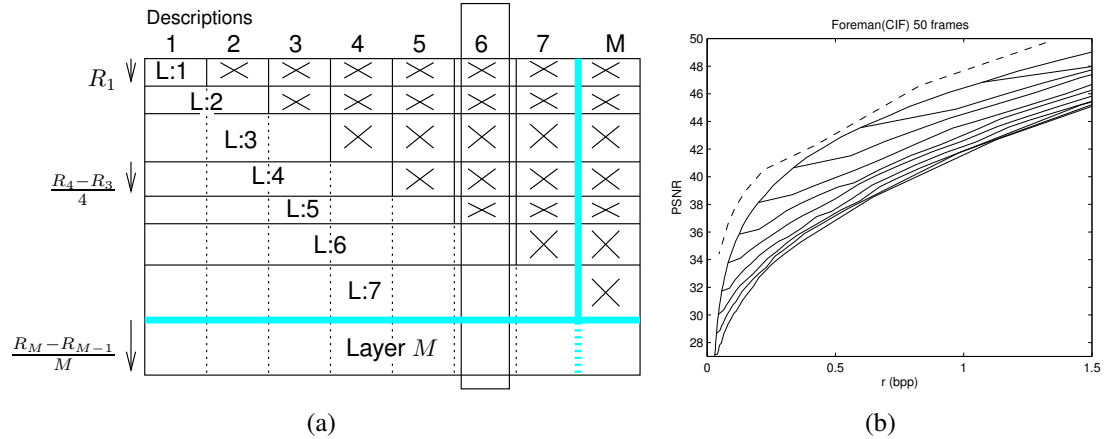


Figure 2: (a) MDC-FEC Packetization scheme. The parts marked with a cross are parity codes. Each column is one description (packet). (b) Rate-Distortion curves for the layered Dirac encoder with the 30Hz CIF-Format Foreman sequence. The dashed line is for the H.264 encoder

In order to be able to find an optimal rate allocation, we need to have an analytic model

of the RD curves. Due to space limitations we will not present the derivation of our model for $D(R_1, R_2 \dots, R_M)$.

RATE ALLOCATION PROBLEM

If we have to serve a large number of different clients, all with a different bandwidth R_c , we have to trade-off quality and redundancy of the descriptions. Making an optimal trade-off is not trivial as clients having different bandwidths will receive a different number of descriptions, and hence experience a different quality and effective bandwidth. On one hand, we wish to offer every client a quality that is as high as possible. On the other hand because of scalability of the system, we cannot offer each client individually an optimal stream. In order to deal with this dilemma in a fair way, two ingredients are needed. In the first place we need to know (or model) the distribution $f_{R_c}(R_c)$ of the bandwidths R_c available to the clients. Secondly, we need to establish a criterion that expresses what we mean by “fair”, and which also lends itself for optimization. The fairness criterion will generally be a function of the clients bandwidth R_c and the bit rates R_k allocated to the layers of the video encoder ($FC(R_c; R_1, R_2 \dots, R_M)$). In the following sections we will introduce and discuss different examples of fairness criteria.

In most cases the rate allocation problem (finding optimal values for $R_1, R_2 \dots, R_M$ given R_D and M) can only be solved numerically. The values of R_D and M , however, can either be optimized numerically but often are chosen practically. The P2P network, for instance, may only support a limited number of descriptions. Or, if we observe that most of the client bandwidths are smaller than B_{\max} , we could choose to select $R_D = B_{\max}/(M + 1)$ so that the client-bandwidth spectrum is equally divided. In the remainder of this paper we only discuss optimizing the rates $R_1, R_2 \dots, R_M$ and we assume that we either have already selected values for M and R_D , or that these values are to be found in an outer optimization loop. When R_D and M are fixed, the maximization function becomes:

$$\{R_1, R_2 \dots, R_M\} = \arg \max \sum_{j=0}^M \int_{b_j}^{b_{j+1}} f_{R_c}(r) FC(r; R_1, R_2 \dots, R_j) dr, \quad (2)$$

where $b_i = i R_D$ for $i = 0, 1, \dots, M$, and $b_{M+1} = \infty$.

MINIMAL MSE FAIRNESS CRITERION

The first – and the most straightforward – criterion we consider is to average the distortion $D(R_1, R_2 \dots, R_m) = D(R_1, R_m)$ over all clients:

$$\hat{D}(R_1, R_2 \dots, R_M) = \sum_{j=0}^M \int_{b_j}^{b_{j+1}} f_{R_c}(r) D(R_1, R_j) dr = \sum_{j=0}^M C_j D(R_1, R_j), \quad (3)$$

where the number of clients receiving $i + 1$ out of M descriptions is C_i , computed as: $C_i = \int_{b_i}^{b_{i+1}} f_{R_c}(r) dr$ for $i = 0, 1, \dots, M$. Given a fixed number of description M and a fixed rate per description R_D , this criterion can be solved numerically using the Lagrange multiplier method, as discussed in the work of Puri *et al.*[2]

$$L(R_1, R_2, \dots, R_M, \lambda) = \sum_{j=0}^M C_j D(R_1, R_j) + \lambda \left(\sum_{j=1}^M \alpha_j R_j - R_D \right) \quad (4)$$

In [2], a method is presented to solve these simultaneous equations such that after optimization $R_i < R_{i+1}$ holds for all rates. Unfortunately, the required conditions can only be verified for a single layer coder that is described by a single RD-function. In our case, however, we have to deal with a layered coder that is described by multiple RD-functions. Consequently, the derivative of $D(R_1, R_2, \dots, R_m)$ with respect to R_i also depends on $R_j, 1 \leq j < i$. For that reason, we cannot use method in [2], and as a result we cannot guarantee that after optimization all $R_i < R_{i+1}$. Currently we deal with monotonicity of R_i as a postprocessing step *after* optimization.

MAXIMAL AVERAGE PSNR FAIRNESS CRITERION

A drawback of the MSE metric described above is that the (inverse) magnitude of the MSE is not a good measure of quality. When the *average* MSE is minimized, most effort is put in minimizing the highest MSE values. Little effort is put in minimizing the lower MSE values, although these still can gain significant amounts of quality. An obvious alternative choice for the fairness criterion is to use the performance measure often used in video compression, namely peak-SNR. When we average the PSNR over all clients, the following criterion is obtained:

$$\widehat{\text{PSNR}}(R_1, R_2, \dots, R_M) = \sum_{j=0}^M \int_{b_j}^{b_{j+1}} f_{R_c}(r) \text{PSNR}(R_1, R_j) dr \quad (5)$$

where $\text{PSNR}(R_1, R_j) = 10 \log_{10} \frac{255^2}{D(R_1, R_j)}$. We can optimize the above criterion in a similar way as the minimum MSE criterion.

WEIGHTED PSNR LOSS (WPL-P) FAIRNESS CRITERION

From the client's point of view, the client is paying for a certain bandwidth R_c and wishes to use this bandwidth as efficient as possible, in particular, the client wishes to obtain maximal quality of the received compressed video stream. Unfortunately, because of the multiple description coding, there is an inherent reduction in quality. As long as there is a balance between the MDC-induced quality reduction and the benefit for the entire P2P video distribution system (of which the client is an integral part), the client is willing to take part in the system.

The third proposed fairness criterion reflects the MDC-induced quality reduction for the individual clients. We consider the system to be fair when the (PSNR) quality of all clients is close to the performance obtained *if* a single description coding system (SDC) had been used. The following *weighted PSNR loss* (WPL) criterion measures the difference between MDC and SDC performance. The p -parameter controls the way in which differences are weighted:

$$\text{WPL}_p(R_1, R_2 \dots, R_M) = \left| \text{PSNR}_{\text{SDC}}(R_c) - \text{PSNR}_{\text{MDC}}(R_1, R_2 \dots, R_{\lfloor \frac{R_c}{R_D} \rfloor}) \right|^p$$

After substitution of this criterion in Eq. (2) and simplification of the resulting expression, we obtain:

$$\widehat{\text{WPL}}_p(R_1, R_2 \dots, R_M) = \sum_{j=0}^M \int_{b_j}^{b_{j+1}} f_{R_c}(r) \left| 10 \log_{10} \frac{D(r)}{D(R_1, R_j)} \right|^p dr. \quad (6)$$

With $p = 1$, this criterion becomes identical to the maximal average PSNR criterion Eq. (5). For $p > 1$ we put more emphasis on larger quality reductions. Minimizing the criterion then results in a solution where the difference between MDC case and the SDC case is more balanced over all clients. Note that for $p \rightarrow \infty$, we effectively minimize the maximum quality reduction.

The Lagrangian for the weighted PSNR loss with parameter p (WPL- p) criterion is

$$L = \sum_{j=0}^M \int_{b_j}^{b_{j+1}} f_{R_c}(r) \left| 10 \log_{10} \frac{D(r)}{D(R_1, R_j)} \right|^p dr + \lambda \left(\sum_{j=1}^M \alpha_j R_j - R_D \right), \quad (7)$$

RATE OPTIMIZATION RESULTS

In this section we present the results of the optimization of above criteria given several different client-bandwidth distributions and different choices for R_D and M . In these simulations we used the general RD model as derived in Section

The first distribution we consider is shown in Figure 3(a). We expressed the clients bandwidths in bit-per-pixel, also note that $f_{R_c}(r > 1.5) = 0$. Furthermore we fix $M = 4$ and $R_D = 0.3$. Using the MSE criterion, this resulted in rate assignment $\mathbf{R} = \{R_1, R_2, R_3, R_4\} = \{0.15, 0.15, 0.18, 0.74\}$. The resulting quality depending on the number of descriptions that is received (and therefore depending on the client bandwidth) is shown in Figure 3(b). This plot shows MDC quality depending on the number of received descriptions in the lower curve. For reference we also plotted the SDC quality $D(r)$ (upper curve). A first observation is that the first layer already obtains a fairly high quality. This has two reasons: a) the RD-function is very steep at lower rates, making it “cheap” to obtain low distortions at low rates; b) since the RD-model correctly models the penalty in coding efficiency in the enhancement layers, a

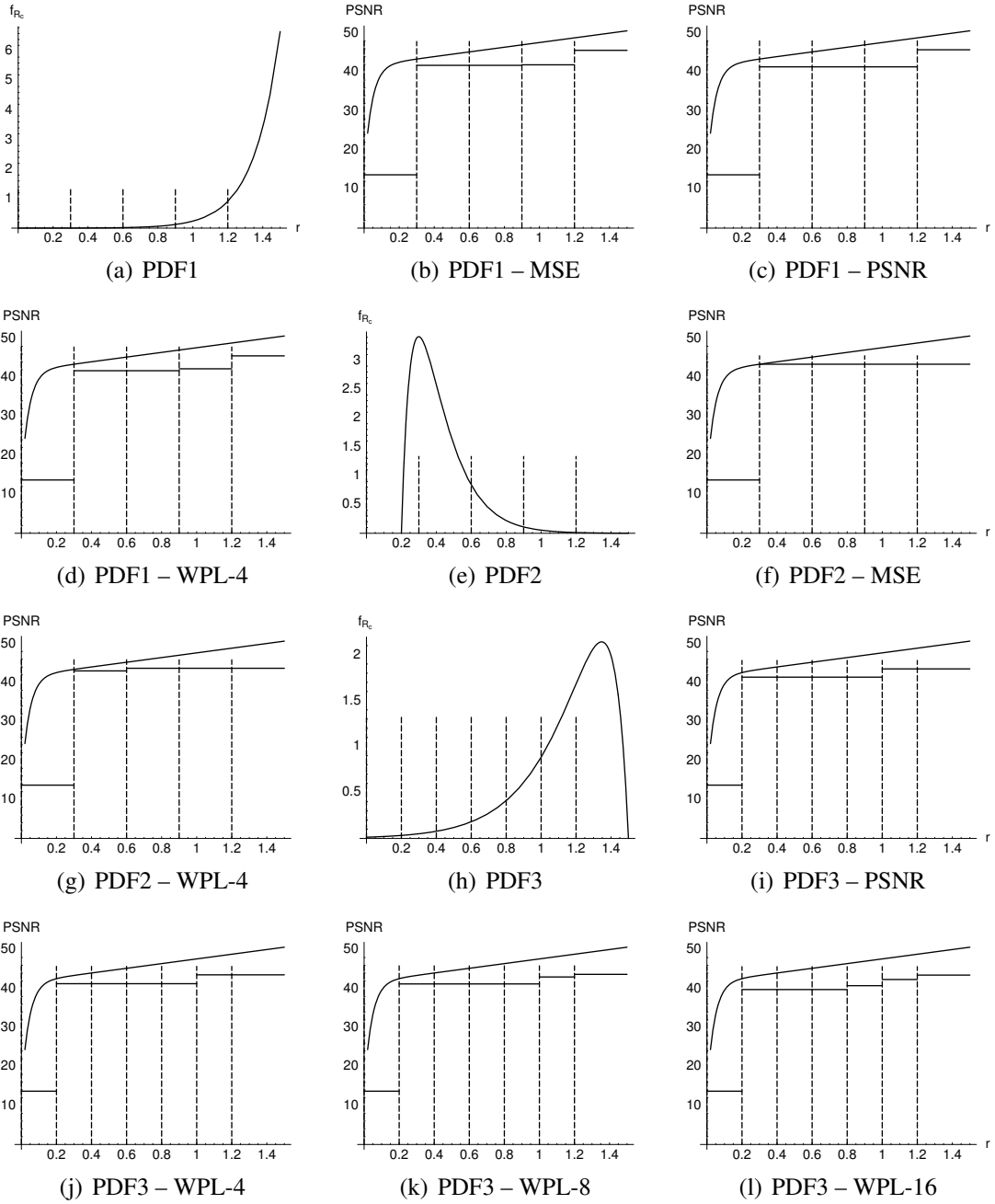


Figure 3: Results for the optimization of the MSE, PSNR and WPL- p criteria for several client distributions (PDF1,2,...)

small base layer implies lower qualities for the higher layers as well. Another observation is that only for clients with bandwidth $R_c > 1.2$, the system really increases the base quality.

When we use the PSNR criterion, the optimal assignment is $\mathbf{R} = \{0.14, 0.14, 0.14, 0.78\}$

(Fig. 3(c)). Although barely noticeable, this criterion favors the higher bandwidths with a little bit higher quality. When we apply the WPL-4 criterion however, $\mathbf{R} = \{0.15, 0.15, 0.32, 0.69\}$, this results in a more “fair” distribution of the quality over all clients. Clients receiving three descriptions do obtain a little higher quality than when receiving only two descriptions, making it worthwhile to participate in the P2P network. Now we consider the distribution in Figure 3(e). This distribution puts emphasis on clients having bandwidths between 0.3 and 0.6. Since $R_D = 0.3$, all clients with $R_c < 0.3$ receive no descriptions at all. For the MSE criterion ($\mathbf{R} = \{0.3, 0.3, 0.3, 0.3\}$) gives a high quality to all clients, but effectively results in a single-layer, single-description solution. With the WPL-4 criterion on the other hand, $\mathbf{R} = \{0.20, 0.31, 0.45, 0.45\}$ and even clients with a $R_c > 0.6$ still gain quality at a small expense of quality for the lower bandwidth clients (Fig. 3(g)).

For the distribution shown in Figure 3(h), we fixed $M = 6$ and $R_D = 0.2$. To gain insight in the effect of the p -parameter, we varied p from 1 (effectively the PSNR criterion) to $p = 16$. Figures 3(i)-3(l) display the results. Remember that a higher p -value puts more emphasis on minimizing larger offsets between the SDC and MDC quality. Using a large p , the criterion goes to a greater extent to prevent large offsets, which results in a system that really offers different qualities for different bandwidths.

CONCLUSIONS

We have introduced three fairness criteria, which in combination with the distributions of the P2P clients’ bandwidth distribution, define the criterion to be optimized. The rate-distortion functions on which the results are based are taken from a layered version of the Dirac video coder. The results we have obtained show that by choosing an appropriate fairness criterion we can obtain a fair mapping between the bandwidth of a client and the quality it will receive. This is very important in a P2P network, where clients will decide on the basis of this mapping, whether they will participate in the P2P streaming network, and how many descriptions they want to receive.

REFERENCES

- [1] M. Castro, P. Druschel, A.-M. Kermarrec, A. Nandi, A. Rowstron, and A. Singh, “Split-stream: high-bandwidth multicast in cooperative environments,” in *Proc. of the 19th ACM symposium on Operating systems principles*, pp. 298–313, 2003.
- [2] R. Puri and K. Ramchandran, “Multiple description source coding through forward error correction codes,” in *Proc. Asilomar Conference on Signals, Systems and Computers*, (Asilomar, CA), Oct. 1999.
- [3] J. R. Taal, J. A. Pouwelse, and R. L. Lagendijk, “Scalable multiple description coding for video distribution in p2p networks,” in *Proc. Picture Coding Symposium*, (San Francisco, CA), 2004.
- [4] J. A. Pouwelse, J. R. Taal, R. L. Lagendijk, D. H. J. Epema, and H. J. Sips, “Real-time video delivery using peer-to-peer bartering networks and multiple description coding,” in *Proceedings International Conference on Systems, Man and Cybernetics*, (The Hague, The Netherlands), pp. 4599–4605, 2004.
- [5] “The DIRAC project.” <http://www.bbc.co.uk/rd/projects/dirac/>.