Hybrid Temporal-SNR Multiple Description Coding for Peer-to-Peer Television

Jacco R. Taal and R.(Inald) L. Lagendijk
j.r.taal@ewi.tudelft.nl, r.l.lagendijk@ewi.tudelft.nl

Information and Communication Theory Group
Delft University of Technology
Mekelweg 4, 2628 CD Delft, The Netherlands

Abstract. In this paper we present Hybrid Temporal/SNR Scalable Multiple Description Coding. For efficient Peer-to-Peer streaming more than two descriptions are required. In our P2P-TV video streaming network, we also need scalability and error resilience. To increase the number of descriptions we use SNR enhancement layers with channel codes (MD-FEC). The robustness against failures in the P2P network, such as packet losses and peer-failure, is increased by either employing odd/even frame MDC or by temporal layering.
We show different hybrid MDC schemes and their respective rate distortion curves and redundancy figures, and discuss the impact of packet-loss and peer-failures on these schemes. Given a peer-to-peer network with certain dimensions, loss rates and failure rates and scalability demand we then optimize the quality of the hybrid MDC streaming over a P2P-TV network.

1 Introduction

In recent years P2P file-sharing has undergone a tremendous boost. More and more, P2P technology is being used for multimedia streaming. Advantages of P2P networks are: increased efficiency, fault tolerance and load balancing. We advocate using Multiple Description Coding (MDC) as an efficient coding technique for P2P streaming. By combining MDC and P2P, we can exploit the efficiency of P2P networks while offering good error-resilience.

Our P2P-TV system generates multiple multicast trees through the P2P network. Each tree carries one description. This way clients have to exchange descriptions with other clients in order to obtain all descriptions and the highest quality. To build efficient multicast trees, more than two trees, and hence descriptions, are required[1]. We assume that the P2P network suffers from peer-failure and packet loss due to congestion. Furthermore, we assume every client has a different download bandwidth, caused by their local router or ADSL line.

With Multiple Description Coding, the source stream is encoded into several parallel descriptions such that each description is sufficient to decode the stream at a reasonable quality and every additional description increases the quality. Descriptions are not prioritized as in layered coding. It is especially this non-prioritized nature of MDC that makes it very useful for P2P transmission, since failures in the P2P network will not have dramatic effects on the picture quality.

Practical MDC coding systems have to match to the network conditions in order to obtain optimal behavior according to some criterion (mostly minimal distortion over all clients) In previous work[2, 3] we discussed the optimization of a MDC codec based on layered coding system with channel codes for several different criteria, such as minimal distortion, maximum average quality and optimal fairness for all clients. In this work we will also apply these results. Crucial in MDC is the required amount of redundancy. The amount of redundancy defines how the MDC-encoded video stream is resilient to errors. Therefore, different network conditions demand different amounts of redundancy. Most papers present a comparison of MDC with another coding method and draw a general conclusion on the applicability of MDC, without looking at the required amount of redundancy, which should follow from their network assumptions. MDC methods differ in rate-distortion behavior and the error-resilience they offer. In fact there is no single MDC method that is superior under all network conditions. Therefore it is sensible to investigate a combination of MDC methods.

Our general Hybrid Temporal/SNR MDC method (HYBRID-MDC)— is a combination of a temporal MDC method and MD-FEC. The most familiar temporal MDC method is ODD-EVEN-MDC MDC or splitstream[4]. In this paper we also present another temporal MDC technique based on temporal layers as can be obtained from the MCTF structure of H.264. MD-FEC is a popular MDC method based on (SNR) Layers and Reed-Solomon Erasure Codes. We believe a hybrid method is able to deal with more different conditions of the P2P network.

In this paper we evaluate a HYBRID-MDC method and show the performance in a fictive P2P network. In

Section 2 we present the model for our P2P-TV network. Section 3 discusses several MDC methods for making two and four descriptions. Here we also present our Hybrid MDC method. Section 4 shows simulations and comparisons of HYBRID-MDC under different P2P network conditions. We conclude with a Discussion (Section 5).

2 Peer-to-Peer Television

In [1], we presented a framework for peer-to-peer streaming of video, called P2P-TV. The advantages of using P2P networks for video transmission are 1) the smart routing facilities that can be implemented at application level; 2) redundancy and fault-tolerance by using multiple application level multicast trees; 3) better distribution of the traffic over the network, instead of a concentration at servers. Our aim is to offer a scalable and error-resilient video stream to the clients. Since most P2P networks are based on the fact that nodes are actively taking part in the network and are forwarding packets to others, we have to make it worthwhile for users to participate. This means we have to offer scalability since users have different (A)DSL speeds and they expect a good quality corresponding to whatever speed they have. Of course we have to offer error-resilience for a smooth watching experience.

Our P2P-TV application implements a peer-to-peer network which offers multiple redundant streaming-trees, tree-repair. The video coder implements Multiple Description Coding. Clients can — dependent on their bandwidth — subscribe to any number of descriptions. Suppose our MDC video coder supports M different descriptions. We then model the demand for descriptions as the probability that a client requests m out of M descriptions Q_m. Due to the streaming latency constraint, we have to assume packet losses will occur when there is congestion. Whenever packets arrive too late to be displayed in time, they are considered lost. We assume a constant packet loss probability of P_{PL}. Furthermore we assume our P2P network is dynamic in the sense that peers are joining and leaving the network erratically which makes that the streaming trees have to be repaired or changed. We assume a peer-repair-time of T_{rep}(s) and the mean time between peer failure T_F, which results in the probability that any packet gets lost due to peer-failure: $P_{\mathrm{PF}} = \frac{T_{\mathrm{rep}}}{T_F}$. Which results in a total packet loss probability of

$$P_{\mathrm{P}} = P_{\mathrm{PL}} + (1 - P_{\mathrm{PL}}) \frac{T_{\mathrm{rep}}}{T_F} \qquad (1)$$

3 Scalable and Error Resilient MDC

The most disastrous visual effect of losing data in video coding is due to the temporal prediction structure of the codec. When a frame is missing or part of a frame is impaired, the error may also impair other dependent frames. The ODD-EVEN-MDC MDC method[4][5] tackles exactly this problem. When one frame is lost, and only one description is affected, the decoder will halve the frame rate and only display the unaffected frames, until the next intra-coded frame arrives. Of course more advanced repair or concealment techniques can be applied. A problem with the ODD-EVEN-MDC scheme is that it cannot be extended to –say– 4 or 8 descriptions. Receiving only one or two descriptions, will result in a very jerky display.

As already discussed we like to have more descriptions for scalability. MD-FEC, which is based on layered coding and Reed-Solomon erasure codes, is easily extensible to more descriptions. As Layered Coding technique we have chosen for H.264/SVC.

The new H.264/SVC Scalable Video Codec, is a very efficient scalable coding technique which has a very low penalty on the PSNR, as compared to single-layer coding. H.264/SVC can generate multiple SNR layers. Due to the MCTF prediction structure, the enhancement layers also exploit temporal correlation, which results in a good performance, but also less error-resilience. Without additional robustness measures, missing data immediately causes effects in the temporal domain.

We propose a combination of temporal multiple description coding and MD-FEC to achieve the following goals:

1. more than two descriptions,
2. good error resilience,
3. scalability.

3.1 Temporal Multiple Description Coding

ODD/EVEN — With the ODD-EVEN-MDC method, each description is obtained by separately encoding all odd frames and all even frames. The elegance of this scheme lies in the fact that from an encoder point of view it is extremely simple and very nicely handles temporal propagation destroyed pictures. The decoder becomes more complex, when it implements concealment of destroyed areas or frames. Although not clearly present, ODD-EVEN-MDC introduces redundancy since both descriptions contain a basic (low frame rate) representation of the original data. The redundancy is in fact caused by a less efficient temporal prediction. The drawback of this method is that this amount of redundancy cannot be controlled. Figure 1 shows the interleaved prediction coding structure for ODD-EVEN-MDC MDC.

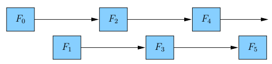


Fig. 1. ODD-EVEN-MDC MDC

BFRAME-MDC — The BFRAME-MDC-MDC method tackles the temporal dependency between frames and is, in fact, a temporal layering method. A temporal base layer is made with a lower frame rate. The temporal enhancement is made of all missing frames, which can be encoded as (hierarchical) B-frames, hence the name. This can be done very efficiently using MCTF. Suppose we have a sequence of 30 fps. The base layer consists of all lower temporal resolutions up to threshold of for instance 15 fps. The temporal enhancement layer contains the remaining high-pass MCTF frames. We generate Multiple Description Packets by combining the base layer of a GOP with one or two frames of the enhancement layer. Figure 2 shows an example of an MCTF prediction structure. The gray Frames belong to the base Layer. The white to the enhancement layer. Fig. 2(b) shows the MDC packet structure for 2-Description MDC, and Fig.2(c) for 4-description MDC.

The ODD-EVEN-MDC and the BFRAME-MDC method are comparable in the sense that both limit the temporal effect of errors and both have no real control over the redundancy.

3.2 SNR-layering and MD-FEC

The MD-FEC scheme is already widely discussed in literature[6], so we will only focus on its application in our Hybrid Temporal/SNR MDC (HYBRID-MDC) scheme. MD-FEC [6], is based on layered coding and channel codes to implement unequal error protection. On each layer k we apply an (M, i) channel code. When arbitrary i out of M descriptions are received, the first layers $1, \ldots, i \leq M$ can be decoded. Figure 3 shows the general packetization structure of MD-FEC.

By changing the rates within each layer, we effectively control the error-resilience and the redundancy.

Figure 4 shows a comparison between the three methods ODD-EVEN-MDC, BFRAME-MDC and MD-FEC for two descriptions and for different description rates R_{mdc}. When all two descriptions are present, the three

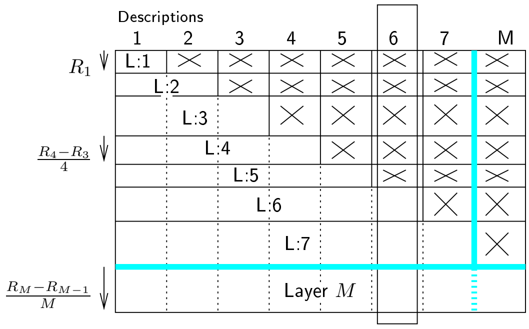


Fig. 3. MD-FEC packetization structure for M descriptions. Each layer L_i is split up in i pieces, on which (M, i) Reed-Solomon code is applied. Each description contains one piece of each layer.

schemes are comparable. When one description is lost, clearly the MD-FEC method outperforms the others. The performance ODD-EVEN-MDC is worse than the others, since we employed no advanced error concealment technique at the decoder. We simply repeat the previous (correct) frame from the other description. An advanced technique might improve the results, but we expect that the ODD-EVEN-MDC method never surpasses the BFRAME-MDC method.

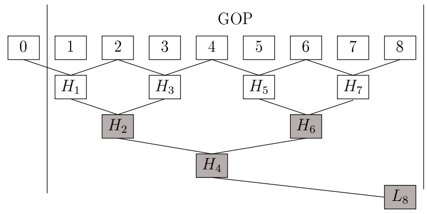
3.3 Hybrid MDC

The HYBRID-MDC method consists of a combination of MD-FEC and BFRAME-MDC. We consider a system with 4 descriptions. The HYBRID-MDC case is basically the same as the MD-FEC method with 4 layers (and 4 descriptions). The base layer however is replaced by descriptions of the BFRAME-MDC method. In such a way that, each hybrid descriptions contains one of the two bframe descriptions and. The BFRAME-MDC encoded base layer adds extra error-resilience to prevent errors in the MCTF predictions.

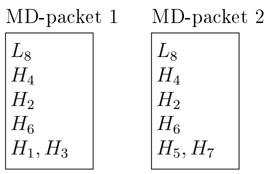
4 Simulations

We now take the P2P network into account. We compare our methods against each other under different network conditions.

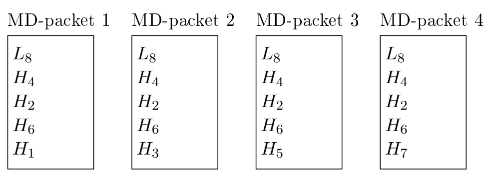
First we only consider packet losses. Under high packet-loss probabilities MD-FEC suffers. When all descriptions (of a frame) are lost, a large part of the group of pictures (GOP) cannot be correctly decoded, resulting in a higher distortion. The BFRAME-MDC method is more resilient to these errors, since all packets contain a low-frame-rate version of the complete GOP, and



(a) Example of a MCTF Prediction Structure for a GOP of 8.

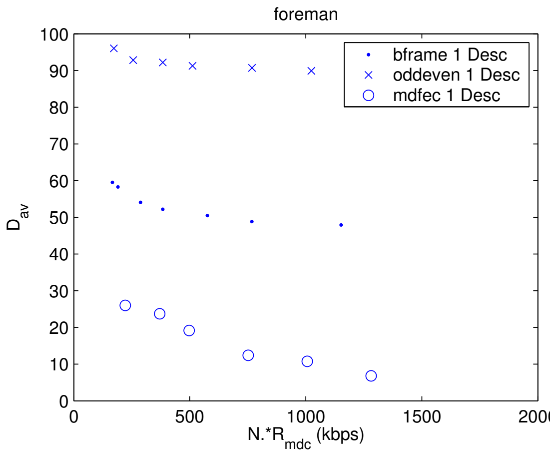


(b) Packet scheme with 2 desc.

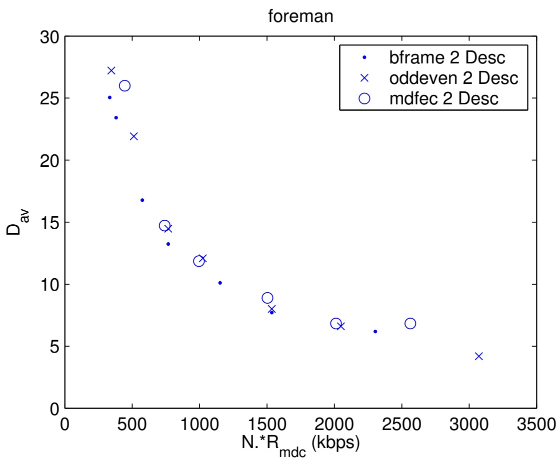


(c) Packet scheme with 4 desc.

Fig. 2. Prediction structure and Packetization structure of BFRAME-method.

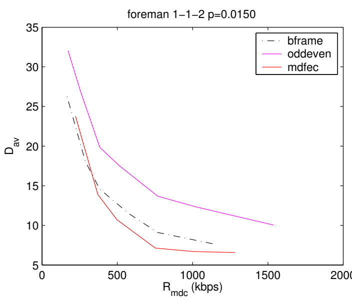


(a) 1 out of 2 descriptions

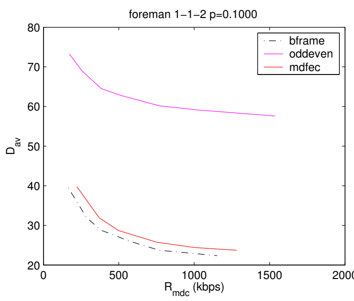


(b) 2 out of 3 descriptions

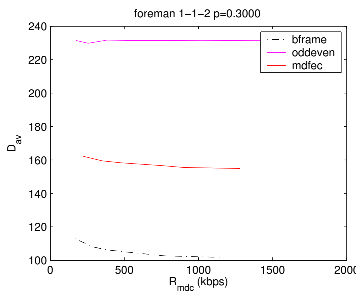
Fig. 4. Comparison of ODD-EVEN-MDC and BFRAME-MDC-mdc and MD-FEC, for 2 descriptions



(a) $P = 1.5\%$



(b) $P = 10\%$

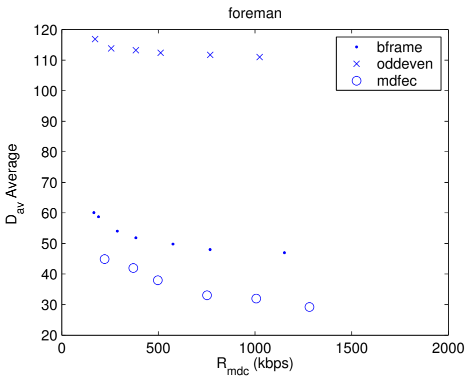


(c) $P = 30\%$

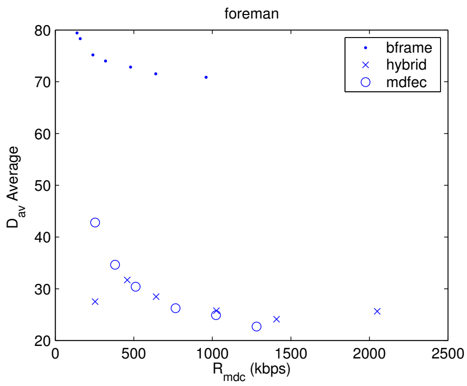
Fig. 5. Performance under different packet loss ratios.

a missing packet only results in a single missing frame, the drawback of course is that BFRAME-MDC has an intrinsically higher redundancy than MD-FEC.

For the P2P scenario we take a Packet loss probability of 20% and a $T_{\rm rep} = 0.5$(s) and mean time without any peer in the multicast tree failing of $T_F = 30s$. Furthermore we asume an equal demand for all numbers of description. i.e. There are as many clients that receive one description as there are that receive two descriptions. Under these conditions we calculate the average distortion over all clients. Figure 6 shows the achieved distortions under these circumstances, for different description bitrate $R_{\rm mdc}$.



(a) 2 description scenario at 10% loss



(b) 4 description scenario at 30% loss

Fig. 6. Peer-to-peer scenarios.

5 Conclusions & Discussion

In this paper we have introduced the BFRAME-MDC and the HYBRID-MDC method. The BFRAME-MDC method generates 2 symmetric descriptions, exploiting the MCTF prediction structure of H.264. We compared the BFRAME-MDC method against the MD-FEC method and found that the BFRAME-MDC offers more error-resilience, especially at higher loss rates, this gives better results. The HYBRID-MDC method, generates 4 descriptions based on a combination of MD-FEC and the bframe method. The HYBRID-MDC method outperforms the MD-FEC method for higher loss rates. In the case of peer-to-peer streaming, the BFRAME-MDC method may prove to be very useful and efficient, when the network suffers from congestion, which results in packet losses. We plan to investigate more cases of HYBRID-MDC, with different numbers of layers and descriptions. Also, more ways of assembling Multiple Description packets from all SNR-layers and temporal layers are possible.

References

1. Pouwelse, J.A., Taal, J.R., Lagendijk, R.L., Epema, D.H.J., Sips, H.J.: 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 (2004) 4599–4605
2. Taal, J.R., Pouwelse, J.A., Lagendijk, R.L.: Scalable multiple description coding for video distribution in p2p networks. In: Proc. Picture Coding Symposium, San Francisco, CA (2004)
3. Taal, J.R., Lagendijk, R.L.: Fair rate allocation of scalable multiple description video for many clients. In: Proc. of SPIE, Visual Communications and Image Processing. Volume 5960. (2005) 2172–2183
4. Castro, M., Druschel, P., Kermarrec, A.M., Nandi, A., Rowstron, A., Singh, A.: Splitstream: high-bandwidth multicast in cooperative environments. In: Proc. of the 19th ACM symposium on Operating systems principles. (2003) 298–313
5. Apostolopoulos, J.: Error-resilient video compression through the use of multiple states. In: Proceedings Int. Conf. on Image Processing, 2000. (2000)
6. Puri, R., Ramchandran, K.: Multiple description source coding through forward error correction codes. In: Proc. Asilomar Conference on Signals, Systems and Computers, Asilomar, CA (1999)