

# A SIMPLE MULTIPLE DESCRIPTION CODING SCHEME FOR IMPROVED PEER-TO-PEER VIDEO DISTRIBUTION OVER MOBILE LINKS

Catherine Lamy-Bergot<sup>\*</sup>, Benjamin Candillon<sup>\*</sup>, Béatrice Pesquet-Popescu<sup>‡</sup> and Benjamin Gadat<sup>\*</sup>

(<sup>\*</sup>) Thales Communications, EDS/SPM, Colombes, France.

(<sup>‡</sup>) Ecole Nationale Supérieure des Télécommunications, TSI Department, Paris, France

## ABSTRACT

This paper presents an adaptation of peer-to-peer diffusion strategies of video over mobile links based on the use of multiple description coding. The system architecture, including the used overlay algorithm (the unstructured GIA algorithm) and video coding with multiple description coding (MDC), are presented as well as their interest in this context. Simulation results illustrate the approach, showing the interest of using MDC vs. a single description coding solution. The proposed simulation framework involves an H.264/AVC video codec with two descriptions and a simple adaptation of the overlay algorithm.

**Index Terms**— video coding, H.264/AVC, Flexible Macroblock Ordering (FMO), multiple description coding (MDC), peer-to-peer diffusion, GIA algorithm, mobile links.

## 1. INTRODUCTION

Peer-to-peer (P2P) data sharing and corresponding content distribution solutions have played for a few years a key role in the Internet growth, representing a large fraction of the traffic in the network. Slowly becoming a standard way of distributing and sharing information, P2P networks are offering different types of applications, from the well-known and widely used file sharing, in particular for multimedia content, with applications such as Gnutella or BitTorrent, to the less-represented high performance computing through resources sharing like SETI@HOME project.

With their capability to drastically change the way people are using Internet, P2P networks and soon their expected mobile declinations, are representing a key challenge for audio and audio-video content distribution. Indeed, the limitations of wireless and mobile P2P networks are obvious: the network links are much less dependable, due to significant error rates over the wireless channels, and to variations of the communication channel quality due to mobility and frequent disconnections caused by peer mobility out of range.

In parallel to the research efforts done to enhance the link reliability in particular by means of error correction

techniques, the last decades have seen the emergence of an alternative solution answering mainly to the situation in which immediate data retransmission is either impossible (*e.g.* network congestion or broadcast applications) or undesirable (*e.g.* in streaming applications with delay constraints). The specific joint source-channel coding technique known as *Multiple Description Coding* (MDC) corresponds to the introduction of redundancy at the source level in order to create multiple correlated streams that can be transmitted over different channels and decoded independently, even though leading to suboptimal quality if only partially received.

In a P2P network, several peers could store the different descriptions or parts of descriptions of the entire video, and a client could request from several peers their descriptions, to decode and combine them for rendering the original video. Using MDC over P2P networks in an approach that has already been proposed for specific applications, such as the improvement of TV stream delivery [1] or in a more general framework for adaptive video streaming [2]. It presents multiple advantages: it is very robust to server/peer breakdown, since a single description has a limited impact on the reconstructed quality and the same description can be stored by several peers; it reduces the load on each serving node, since each video file is split among different peers; and finally it allows using reduced uplink bandwidth for each serving node, as each description has a lower rate than the entire video. We focus here on showing that MDC schemes are useful per se in the P2P context. Indeed we will show that very simple operations can be done to take advantage of the use of multiple descriptions over classical single description schemes to ensure both a quicker access to a lower quality version of the requested content, and protection against unaccessible elements due to unavoidable disconnections of mobile peers.

This paper is organized as follows. Section 2 introduces the principle of P2P networks and details some of the existing strategies proposed for content distribution over a mobile link, in particular the selected GIA overlay algorithm. Section 3 then presents the principles and interest of multiple description coding and describes the video coding technique used. In Section 4, experimental results illustrate how MDC can help to combat losses in the P2P network. Finally, conclusions are drawn and perspectives are proposed in Section 5.

---

This work was partially supported by the French Agence Nationale de la Recherche through project RNRT DITEMOI (06 TCOM 003-\*).

## 2. P2P NETWORKS AND SYSTEM ARCHITECTURE

Peer-to-peer architectures relies on large-scale network overlays that offer the mechanisms allowing to discover and route the data from its storage peer to the requesting one, with a request finding its path among various peers up to the desired file. Different types of overlay mechanisms exist, generally classified according to the degree of centralization they still may require. Typical P2P networks are purely decentralised systems where all peers assume the same role, acting both as clients and servers. In this kind of systems, resources are shared without any centralised control or a priori knowledge of information location. However, it is common to find hybrid P2P systems, where some well-known nodes assume explicitly the role of servers (for example for receiving requests or managing a list of users). In this variant of P2P systems, these well-known nodes must be always contacted prior to any transfer of resources. A second important distinction between the different overlays is the way peers are organised, for which two cases can be found: structured and unstructured. Unstructured overlays, as for instance GIA [3] or DYMO [4], do not define a specific organisation for the nodes but react to peer requests to discover the underlying graph. Floods or random walks can be used to evaluate at each visited node the capacity to satisfy the query locally or on known neighbour nodes. In the negative case, neighbour nodes are visited, until a given timeout or a limiting threshold are met. The drawback is the necessity to perform numerous searches to find nodes storing the requested data, in particular when this requested data is rare in the P2P network. Structured overlays have been developed to answer to this problem of data discovery performance. By imposing constraints on the graph itself, and possibly on data placement, they can offer much quicker discovery of the requested data. However, when considering mobile networks, where nodes are extremely transient and mobile, these constraints on the graph are quite difficult to meet and, generally, will favor unstructured overlays in deployed data sharing applications.

For video sharing over mobile peer-to-peer networks, we consequently propose to consider a decentralized and unstructured overlay: GIA [3]. Its token-based flow control algorithm allows to take into account the fact that peers have each specific capacity constraints, and the use of random walk requests instead of the more traditional flooding approach helps to provide path random diversity, which will be interesting for MDC streams. The requests made by the GIA rely on keywords identifying the content present in the network and looked for by a peer. To adapt it to the case of transient P2P clients and variable wireless links capabilities, we have cut the different video files in chunks and allocated a specific identification key to each chunk (key to be in practice transmitted with its corresponding chunk): the routing algorithm working on the keys and not on the files themselves, it allows an effortless adaptation to link degradations or disconnections. The

GIA overlay routine is then used by the node IP forwarding engine when no information is available in the routing table to decide of request transmissions and message exchanges, as illustrated by Figure 1. Furthermore, to better take advantage of the properties of MDC streams, the file request procedure was modified. Instead of sending out randomly requests of any key of the complete file, requests are grouped at the description level: requests concerning a new description can only be dispatched after requests for the chunks of a given description have all been sent out.

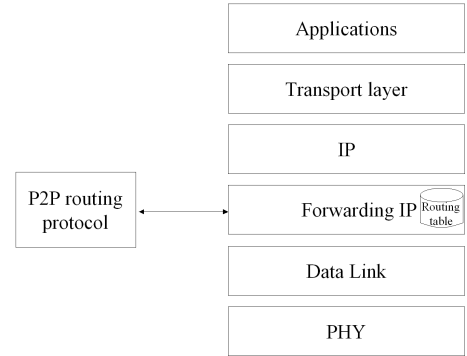


Fig. 1. Protocol layers in a considered P2P node.

## 3. MULTIPLE DESCRIPTION CODING

Multiple description coding denotes any source coding technique leading to several correlated but independently decodable (preferably with equivalent quality) bitstreams, called *descriptions*. Each of these streams are often sent over independent channels. In the simplest scenario, the channel works in an ON-OFF mode; when an error occurs in one channel, the corresponding stream is considered unusable and is thus discarded at the decoder side. The case of two descriptions is represented in Fig. 2.

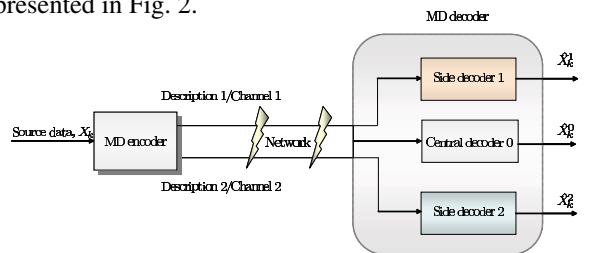


Fig. 2. Two-channel scheme for multiple description coding.

It corresponds to the case where we have to transmit the source  $X_k$  and we know that two independent channels are simultaneously available, in an error-prone network. Therefore the two descriptions are generated by the MD encoder and each of them is sent over its corresponding channel. The MD decoder handles two different situations: in the first one, errors have occurred on one of the channels and the decoder ignores the data coming from it, delivering an approximate version of  $X$ ,  $\hat{X}_k^j$  with  $j \in \{1, 2\}$ ; in the second situation both channels were unaffected by errors and a *central decoder* produces a (usually) better version of  $X$ ,  $\hat{X}_k^0$ . The reception of

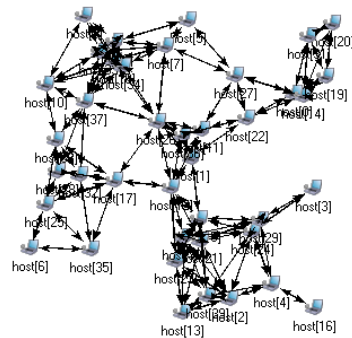
only some of the descriptions will allow a low quality decoding, while the full decoding will lead to optimal quality. An ingredient enabling the success of an MDC technique is the path diversity, since its usage balances the network load and reduces the congestion probability. Another ingredient is the limited amount of redundancy introduced at source level between the descriptions, such that an acceptable reconstruction can be achieved from any of the bitstreams.

Without explicitly introducing additional redundancy, some methods exploit the correlation existing in natural images or video, for example by splitting the successive frames of a sequence in odd and even ones and allocating them to the two descriptions [5], or by separating the polyphase components of a transform and sending them separately [6]. In the same class of methods enters a particular case of the Flexible Macroblock Ordering (FMO), which is an error resilience tool available in the H264/AVC video coding standard. FMO allows to assign macro-blocks (MB) to slices or Network Abstraction Layer units (NALs) in an order different from the scan order inside the picture. Each slice group can be partitioned into one or more slices, such that a slice is a set of consecutive macroblocks inside a particular slice group (but not necessary inside the picture). The way MBs are grouped inside a slice group can obey different strategies. A famous case is the so-called two-slice groups *dispersed* mode, splitting the MBs in two sets, following the pattern of a chessboard. Another case is the separation of odd lines in a description and even lines in another description. Further assume that, during transmission, the packet containing the information of the first slice group gets lost. Since every lost MB has several spatial neighbors that belong to the second slice group, error-concealment mechanisms can be efficiently employed. Experiments have shown that, in video conferencing applications with CIF-sized pictures, the visual impact of the losses can be kept so low that it takes a trained eye to identify the lossy environment. The coding efficiency for such FMO modes is slightly lower than for a usual H264/AVC encoding, because the prediction mechanisms have to be adapted to non-neighboring blocks. This is however the price for allowing independent decoding of each description. We used this very simple MDC technique in our experiments, as FMO presents the advantage of being available even in basic H.264/AVC codecs.

#### 4. SIMULATION RESULTS

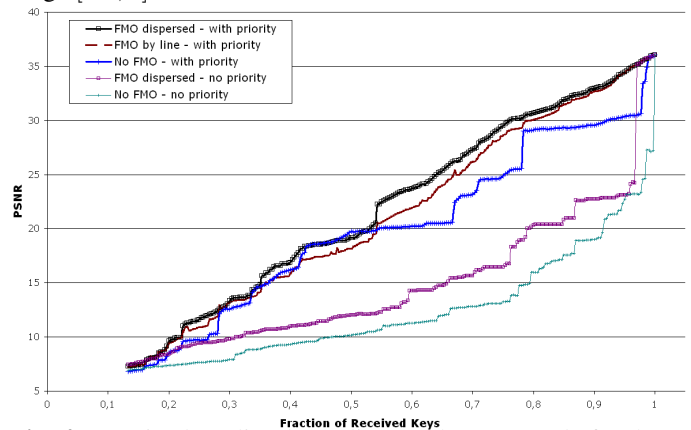
Simulation results have been performed using the OMNeT++ discrete event simulator [7], in which the GIA overlay was introduced from the Oversim package [8]. The major modifications we made consists in the improvement of Oversim token management procedure and the introduction of a retrieval procedure with description priority, as explained in Section 2.

A compilation of 11 ITU-T QCIF at 15 Hz reference sequences totalling 3361 frames (Akiyo, Carphone, Children,



**Fig. 3.** Considered realisation of the P2P network within the OMNeT++ framework.

Container, Foreman, Hall Monitor, Mobile Calendar, Mother and daughter, Stefan, Table tennis, Trevor) has been used to obtain a sequence length representative of a video clip duration. The H.264/AVC encoder [9] was then used either in FMO with two slice groups to present two correlated descriptions as described in the previous section, or with the classical single description coding mode, which offers a slightly better initial video quality for the same overall bitrate, due to a reduced overhead. As illustrated in Fig. 3, the P2P simulated network consists of 40 hosts, which evolve in a rectangular area of dimensions 300 units on x-axis and 270 units on y-axis. The connectivity depends on the relative distance between hosts, the maximum range for connection being set in our simulations to 70 units. The nodes are mobile in both vertical and horizontal directions, with a movement every 5 seconds (to be compared to an average simulation time greater than thousands of seconds), according to a uniform law in the range  $[-2; 2]$  units.



**Fig. 4.** Received quality depending on the P2P mode for the ITU-T reference sequence compilation.

Fig. 4 presents the obtained results with, for each fraction of received keys given on the x-axis, the corresponding obtained visual quality expressed in terms of PSNR (Peak Signal to Noise Ratio) for the sequences compilation. A first observation is that the sequences reconstructed with chunks of data received in random order perform both quite badly until having received almost all data. This is explained by the fact that the partially received streams consist of NAL units

randomly located in the stream, which provide a decoded sequence desynchronised with respect to the original sequence. This phenomenon, which is worsened by the fact that concealment is traditionally done forward in time and not backward, leading to bad results when the first elements of a sequence are not received. Logically, the curve for the FMO dispersed mode then jumps to a better PSNR as soon as it has totally completed a first description, here when about 95% of the keys have been received. These results enhance the interest of working with a priority system for the key request, which was achieved with a pool of key requests done in logical temporal order for the non-FMO encoded sequence, and in logical temporal order of even and then odd descriptions for the FMO dispersed and by line encoded sequences. The interest of using two descriptions vs. a single one is obvious: very quickly, the FMO encoded sequence outperforms the non-FMO encoded one by several dBs in PSNR, due both to the fact that the entire sequence can be decoded in FMO mode when about 55% of the keys have been received. This 55% figure is due to the token approach and time to receive information, which means that some chunks of data corresponding to the odd description can be received before the final ones corresponding to the even description. The obtained results are also related to the efficiency of the used temporal concealment technique based on the Boundary Matching Algorithm (BMA) [10].

To better understand the interest of the multiple description approach, let us focus on two cases: 80% (resp. 60%) of the keys are received, corresponding in our simulations to only about 32% (resp. 24%) of the total time needed to retrieve the totality of keys. We obtain in both cases a PSNR difference of 3 to 4 dB. In the second case, the FMO dispersed sequence totalling 3261 received frames presents 24 dB, while the standard sequence PSNR is only 20 dB, with 2336 frames received. Clearly, with more than one thousand frames not received, the standard sequence presents major holes in its reconstructed sequence, with complete parts of the stream lost, while the FMO encoded sequence is almost complete, except for a few frames at the end of the sequence. Another interest of the MDC approach is that the non-FMO mode succeeds in totally decoding the sequence only when about 95% of the keys have been received, which implies to retrieve almost all elements from the P2P network, including those rarely represented or stored very far from the end-user. In practice, this would imply much greater time needed to obtain a perceived quality as good as the one obtained with MDC in priority retrieval mode (in our simulations, approximately the double of the time needed for FMO sequences).

## 5. CONCLUSIONS AND FUTURE WORK

In this paper, we have studied a practical implementation of a mobile P2P video streaming system with GIA overlay, and highlighted the interest of using a multiple description coding technique in this context. In particular, the FMO in dispersed

mode chessboard pattern is used as a simple instantiation of the MDC, since it does not require any specific change of the existing H264/AVC codecs. The obtained results first highlight the interest of having a coupled approach, with a priority system for the key request instead of a random one, and an MDC encoding of the content. Using FMO encoded sequences with priority keys retrieval allows a large improvement in terms of bandwidth usage (depending on the quality level wished by the user, the transmission can be cut before the completion of the file transmission) but foremost, in terms of time spent, as only one third of the total time is in our simulations needed to retrieve 80% of the keys.

A solution to greatly improve the network performance in terms of delay and capacity usage could then be to add error correction to the various keys. This would allow to compensate for the non-received last keys and yield an (almost) perfect visual quality with a very reduced time delay.

## 6. REFERENCES

- [1] J. Pouwelse, J. Taal, R. Lagendijk, D. Epema and H. Sips, "Real-time video delivery using peer-to-peer bartering networks and multiple description coding," in *Proc. of IEEE SMC'04*, pp. 4599–4605, Oct. 2004.
- [2] E. Akyol, A. Tekalp and M. Civanlar, "A flexible multiple description coding framework for adaptive peer-to-peer video streaming," *IEEE Journal of Selected Topics in Signal Processing*, pp. 231–245, Aug. 2007.
- [3] Y. Chawathe, S. Ratnasamy, L. Breslau, N. Lanham and S. Shenker, "Making Gnutella-like P2P systems scalable," in *Proc. of Sigcomm'03*, pp. 407–418, Aug. 2003.
- [4] I. Chakeres and C. Perkins, "Dynamic MANET on-demand (dymo) routing," Internet Draft (13), Mobile Ad hoc Networks Working of IETF, April 2008.
- [5] J. Apostolopoulos, "Reliable video communication over lossy packet networks using multiple state encoding and path diversity," in *Proceedings of VCIP'01*, pp. 392–409, Jan. 2001.
- [6] N. Franchi, M. Fumagalli, R. Lancini and S. Tubaro, "Multiple description video coding for scalable and robust transmission over ip," *IEEE Trans. on Circ. and Syst. for Video Tech.*, pp. 321–334, March 2005.
- [7] "OMNeT++: discrete event simulation system," v3.3, <http://www.omnetpp.org>, 2007.
- [8] I. Baumgart, B. Heep and S. Krause, "Oversim: A flexible overlay network simulation framework," in *Proc. of IEEE GI '07*, pp. 79–84, May 2007.
- [9] "Joint verification model for h.264 (jm12.1)," 2007. [Online]. Available: <http://iphome.hhi.de/suehring/tml>
- [10] W.-M. Lam and A. Reibman, "Recovery of lost or erroneously received motion vectors," in *Proc. of IEEE ICASSP'93*, pp. V417–V420, April 1993.