An MDC-based video streaming architecture for mobile networks

Claudio Greco, Giovanni Petrazzuoli, Marco Cagnazzo, Béatrice Pesquet-Popescu

TELECOM-ParisTech, TSI department
46 rue Barrault, F-75634 Paris Cedex 13
Paris, FRANCE

{greco,petrazzu,cagnazzo,pesquet}@telecom-paristech.fr

Abstract—**Multiple description coding (MDC) is a framework designed to improve the robustness of video content transmission in lossy environments. In this work, we propose an MDC technique using a legacy coder to produce two descriptions, based on separation of even and odd frames. If only one description is received, the missing frames are reconstructed using temporal high-order motion interpolation (HOMI), a technique originally proposed for distributed video coding. If both descriptions are received, the frames are reconstructed as a block-wise linear combination of the two descriptions, with the coefficient computed at the encoder in a RD-optimised fashion, encoded with a context-adaptive arithmetic coder, and sent as side information. We integrated the proposed technique in a mobile ad-hoc streaming protocol, and tested it using a group mobility model. The results show a non-negligible gain for the expected video quality, with respect to the reference technique.**

Index Terms—**Video coding, multiple description, legacy coder, image interpolation, robust coding, mobile ad-hoc networks.**

I. INTRODUCTION

Multiple Description Coding (MDC) is a framework allowing an improved immunity towards losses on error prone channels, thus providing a trade-off between coding efficiency – in terms of compression ratio for a given quality – and robustness.

The basic principle of MDC is that the encoder, given an input signal (image, audio, video, *etc.*), is able to produce a set of *descriptions*, *i.e.*, a set of sub-streams which are mutually refinable, but – unlike scalable coding – also independently decodable. Each description provides a low, yet acceptable, quality; while, as any further description is received, the quality of the reconstruction increases, independently on which description is received [1]. The decoding unit used when all descriptions are available is referred to as *central decoder*, while any decoding unit used when a non-empty subset of the descriptions is available is referred to as *side decoder*.

Several solutions have been proposed for video MDC: separation of the even and odd frames, to be encoded independently with recovery of missing frames by interpolation from the nearest decoded frames [2]; redundant wavelet transform, which also allows scalability [3]; or periodical insertion of redundant pictures to increase loss resilience and reduce error propagation [4].

However, while many authors have proposed to design MD codecs from scratch, others have pointed out that an MD codec based only on pre- and post-processing, with the use of legacy coders, reduces significantly the development time, hence the development cost [5–7], even though it may come at the price of sub-optimal performance. Within this context, we recently proposed an MDC scheme for double description coding which is entirely based on pre- and post-processing [8]. In this work, we propose an improved version of this scheme, and integrate it into a cross-layer framework for video transmission on wireless ad-hoc networks [9].

The rest of this paper is organised as follows: in Sec. II we introduce the reference MDC technique, while in Sec. III we propose some improvements; then, in Sec. IV, we address the problem of the transmission of the encoded stream over a wireless ad-hoc network. In Sec. V, we present our experimental results using a mobility model particularly adapted to the ad-hoc scenario. Finally, in Sec. VI we draw conclusion and point out some possibilities for future work.

II. MULTIPLE DESCRIPTION CODING REFERENCE SCHEME

In this section, we present the MDC scheme as proposed in our previous work.

The original video sequence is split up into even and odd frames. Then, each sub-sequence is encoded independently with a legacy video coder to produce the two descriptions. Any video coder could be used, but here we shall employ H.264, which is the most recent standard for video coding. Side decoding is performed using an H.264 decoder on the received description, then reconstructing the missing frames using the temporal interpolation technique proposed in the DISCOVER project [10, 11] and based on the assumption of constant velocity motion. When both decoded descriptions are available, central decoding is performed as a block-wise convex combination of the sub-sequences. The relative weight α of each block of the received frame with respect to the corresponding block in the interpolated frame is computed and quantised at the encoder to minimise the distortion between the block in the original frame and the convex combination; then the sequence of weights α is sent along with the descriptions as side information (see Fig. 1).

In order to reduce the bitrate needed to transmit the sequence of weights α, a context-based coding is adopted, with the context E defined as the distortion between received blocks
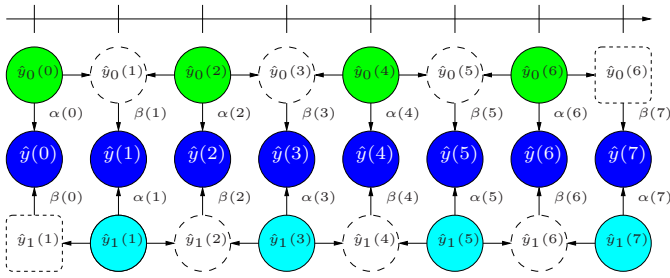
Figure 1. Structure of central decoder. Solid circles represent received frames; dashed circles represent interpolated frames. Horizontal arrows represent interpolation. Vertical arrows represent weighted sum. With $\hat{y}_n(k)$ we denote the k-th frame of description n. For each k, $\beta(k) = 1 - \alpha(k)$.

and interpolated blocks. Since the number of possible contexts is very high, in order to avoid *context dilution*, we perform a context quantisation [12]: the distortion values are grouped into quantisation clusters defined by a convex quantisation function $Q(E)$. The thresholds of the quantisation intervals are found by minimising the mutual information $I(\alpha; E|Q(E))$.

III. PROPOSED IMPROVEMENTS

In this section, we introduce the proposed improvements for the technique described in Sec. II.

A. High-order motion interpolation

The reference technique performs linear interpolation by using the two closest frames (previous and next) w.r.t. the frame that has to be interpolated. This technique works well for constant-velocity movement, but when between these frames the acceleration is not null, a high order motion interpolation (HOMI) is necessary to better model the movement [13].

The main idea consists in using four frames, instead of two, for the interpolation. Let us call I_k the frame that has to be estimated; the frames that will be used for the interpolation are I_{k-3}, I_{k-1}, I_{k+1} and I_{k-3}. Frames I_{k-3} and I_{k+3} have been chosen because the sequence is sub-sampled, and each side description has only either the even or the odd frames of the original sequence.

First, the interpolation method proposed in DISCOVER [10, 11] (which consists in a linear interpolation) is applied to frames I_{k-1} and I_{k+1}. The result is a pair of motion vector fields $\widetilde{\mathbf{u}}(\cdot)$ and $\widetilde{\mathbf{v}}(\cdot)$ (black dashed vectors in Fig. 2). These vector fields are such that, for each point $\mathbf{p}$, the missing frame may be estimated as:

$$\widetilde{I}_k(\mathbf{p}) = \frac{I_{k-1}(\mathbf{p} + \widetilde{\mathbf{u}}(\mathbf{p})) + I_{k+1}(\mathbf{p} + \widetilde{\mathbf{v}}(\mathbf{p})))}{2}$$

Let us define the block of the frame I_k centred in $\mathbf{p}$ as $B_k(\mathbf{p})$. We presume that the reference technique may be affected by an error, consisting in an incorrect placing of the block $B_k(\mathbf{p})$ in $\widetilde{I}_k$. Hence, we try to displace the block keeping into accounts frames I_{k-3} and I_{k+3} as well. In order to do that, we match the blocks centred in the points $\mathbf{q}^{(-1)} = \mathbf{p} + \widetilde{\mathbf{u}}(\mathbf{p})$ in I_{k-1} and $\mathbf{q}^{(+1)} = \mathbf{p} + \widetilde{\mathbf{v}}(\mathbf{p})$ in I_{k+1} with blocks in frames I_{k-3} and I_{k+3} respectively. Let us call $\widehat{\mathbf{u}}(\cdot)$

and $\widehat{\mathbf{v}}(\cdot)$ the corresponding motion vector fields. Let us define $\mathbf{q}^{(-3)} = \mathbf{q}^{(-1)} + \widehat{\mathbf{u}}(\mathbf{q}^{(-1)})$ and $\mathbf{q}^{(+3)} = \mathbf{q}^{(+1)} + \widehat{\mathbf{v}}(\mathbf{q}^{(+1)})$.

We suppose that the blocks $B_{k-3}\left(\mathbf{q}^{(-3)}\right)$, $B_{k-1}\left(\mathbf{q}^{(-1)}\right)$, $B_{k+1}\left(\mathbf{q}^{(+1)}\right)$, and $B_{k+3}\left(\mathbf{q}^{(+3)}\right)$ are actually the same block, which is moving; hence, we interpolate its trajectory, and find the position of its centre in time instant k, that we indicate by $\mathbf{q}$. The motion vector field from k to $k-1$ is $\mathbf{u}(\mathbf{q}) = \mathbf{q}^{(-1)} - \mathbf{q}$ and the one from k to $k+1$ would be $\mathbf{v}(\mathbf{q}) = \mathbf{q}^{(+1)} - \mathbf{q}$ (red dashed vectors in Fig. 2). In order to avoid holes and overlapping blocks, since $\mathbf{q}$ might not correspond to the centre of any block in I_k, we assume that $\mathbf{u}(\mathbf{p}) \approx \mathbf{u}(\mathbf{q})$ and $\mathbf{v}(\mathbf{p}) \approx \mathbf{v}(\mathbf{q})$ (green dashed vectors in Fig. 2), justified by the fact that $\mathbf{p}$ and $\mathbf{q}$ are very close to each other, *i.e.*, they belong to the same object. Then, the missing frame can be estimated for each point $\mathbf{p}$ as:

$$\widetilde{I}_k^*(\mathbf{p}) = \frac{I_{k-1}(\mathbf{p} + \mathbf{u}(\mathbf{p})) + I_{k+1}(\mathbf{p} + \mathbf{v}(\mathbf{p})))}{2}$$

In principle, this technique could be extended to an arbitrary number of frames preceding and following I_k; however, we limited the number of positions from which we interpolate in order to reduce the dependency among frames, which could affect the robustness of the decoding in a lossy transmission scenario. Moreover, if a too large number of frames is used, the interpolated trajectory may be incorrect and unreliable, due to possible mismatches during the block matching.

B. RD-optimised estimation of linear combination coefficients

In the reference technique [8], the sequence of weights α is computed at the encoder to minimise the distortion $D(\alpha)$ between the block in the original frame and the convex combination, and the resulting rate $R(\alpha)$ for the sequence depends on its conditional entropy given the quantised context. We hereby propose the following improvement: first, we estimate $H(\alpha|Q(E))$ with a preliminary test; then, the conditional entropy is used as an estimation of the coding cost $\widetilde{R}(\alpha)$ and, for each block, we find the optimal value α^*, defined as:

$$\alpha^* \triangleq \arg\min_{\alpha} \left\{ D(\alpha) + \lambda \, \widetilde{R}(\alpha) \right\},$$

where the Lagrangian parameter λ can be chosen equal to the one used by the hybrid coder to encode the sequence.

IV. TRANSMISSION OVER WIRELESS AD-HOC NETWORK

In order to validate our technique, we decided to integrate it in a suitable scenario that properly exploits its properties. Thus, in this section, we discuss the streaming of a video content, encoded with the proposed technique, over a mobile ad-hoc network. A mobile ad-hoc network (or MANET) is a dynamic network of mobile devices inter-connected by wireless links, self-organised in a mesh topology [14]. The set of properties offered by MANETs – such as flexibility, ease of deployment, robustness, *etc.*– makes them suited for real-time transmission in environments without preexisting infrastructure, such as military or disaster-relief applications [15]. In these scenarios, given both the inherent lossy nature of MANETs and the real-time constraints of video streaming,
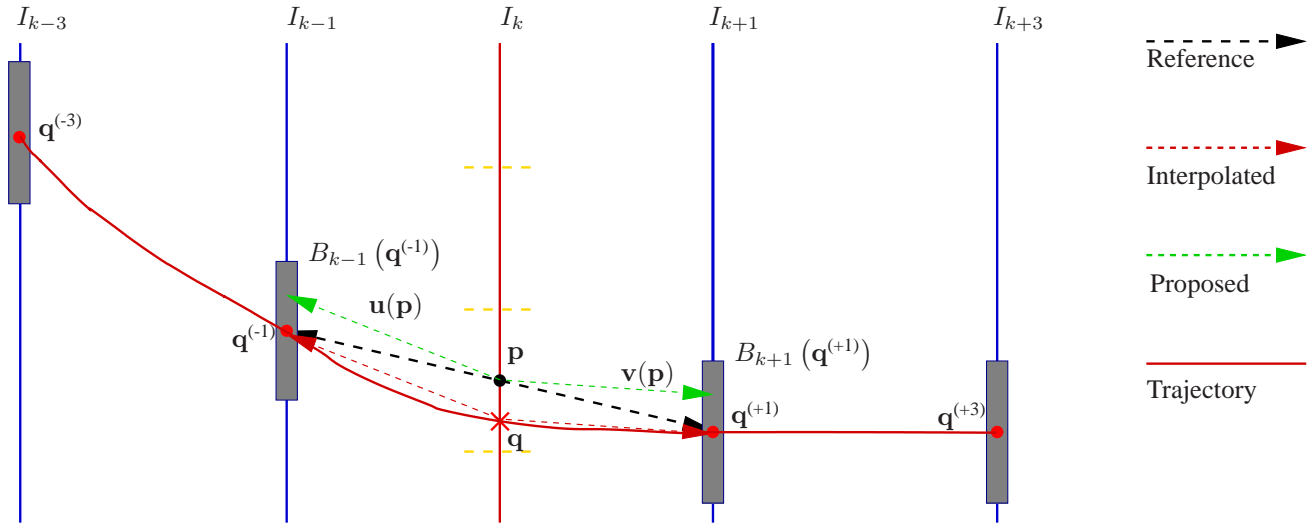
Figure 2. Proposed interpolation method for motion estimation. The position of the block in the current frame is estimated by interpolating its trajectory in two previous and two following frames.

MDC is a very viable solution, since it provides a trade-off between coding efficiency – in terms of rate/distortion – and robustness towards losses.

A. Transmission protocol

For real-time video streaming, solutions at Application layer perform poorly over ad-hoc networks because they are unaware of the links existing among nodes and their states, thus producing an overlay mismatched w.r.t. the topology of the network. On the other hand, solutions at Data-Link layer are unaware of the specific logic and state of the application, and are thus hardly adapted to its needs. Having a shift towards a cross-layer approach been pointed out as needed to overcome these limitations [16], we recently proposed [9, 17] a novel protocol called ABCD, inherently designed for video streaming in ad-hoc wireless networks, which exploits the broadcast property of the medium in a cross-layered fashion.

The ABCD protocol provides a multi-tree overlay network, with a different tree for each description. Moreover, the protocol is able to quickly adapt to topology changes (movement, nodes' departure or arrival) with a small number or exchanged messages. The main feature of ABCD is the modified 802.11 MAC layer, which implements a reservation mechanism for reliable broadcast, which greatly reduces the collision probability, thus allowing better performance than the standard 802.11 w.r.t. the rate *vs.* diffusion area trade-off typical of multi-hop broadcast in wireless networks [18].

B. Mobility model

The node mobility model is one of the key factors when evaluating the performance of an application over mobile ad-hoc networks. The most commonly used mobility model is the Random Way-point Model (RWM) [19]. In RWM, each node selects a random destination and a random speed, then moves to the selected destination at the selected speed. The destination reached, the node stays for a random pause time,

then reiterates the process by selecting a new destination, speed, and pause time. However, RWM fails to describe the co-herence among the nodes' movements, which is to be expected in military and disaster-relief scenarios [20]. Therefore, we preferred to use the Reference Point Group Mobility (RPGM) to model the movements of the nodes [21]. RPGM is a generalised version of RWM in which each node is statically assigned to a group, each group having a logical centre; each node has a randomly assigned reference point, *i.e.*, a relative position to the centre of its group. The reference points follow the same movement vectors of the centre (which is determined according to an RWM model) and keep their relative position w.r.t. it. The nodes then set their destination point randomly in a neighbourhood of their reference point. This model has been chosen because it provides a relatively simple framework, but is able to accurately describe realistic scenarios where different groups carry out different tasks over the same area, *e.g.*, in a disaster recovery. In our implementation, to add realism to the model, each group is centred on a node (*group leader*), to whose movement the other nodes respond, with a random delay and a random variation of their speed.

V. EXPERIMENTAL RESULTS

A. Comparison over lossless transmission

To encode the test sequences in our proposed scheme and in the reference scheme (*i.e.* the one proposed in [8]) we used as legacy coder the H.264/AVC reference software JM [22], version 17.0. We selected a set of QPs (22, 25, 28, 31, 33, and 36) in order to compare the RD performance of the two methods.

The rate-distortion performance comparison for side decoding, which is the most affected by our proposed technique, of the video sequence "Stefan" (CIF, 30 fps) is shown in Fig. 3.

A more concise comparison over a set of sequences is given in Tab. I in terms of Bjontegaard metric of the proposed
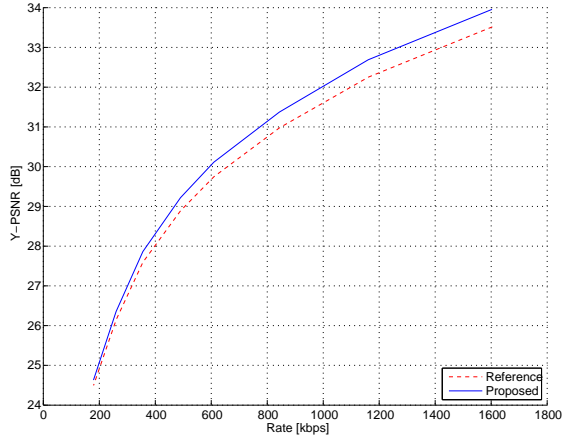
Figure 3. RD-comparison between the reference and the proposed technique, for video sequence "Stefan", CIF, 30 fps (Side decoder).

Figure 4. Performance versus packet loss rate comparison for fixed bitrate (1.0 Mbps).

technique w.r.t. the reference technique [23]. Notice that, since sequences are encoded with a fixed QP, the resulting bitrate is higher for sequences with a higher motion content.

We notice that the quality of the central reconstruction for the two techniques is slightly improved. We explain this fact by observing that, with a proper choice of the α coefficients, the central decoder is quite robust to possible errors in the interpolated frames. For the side decoders, on the other hand, we see that in some cases the reference technique can be significantly improved by the high-order motion interpolation. For instance, for the sequences "Foreman" and "Football", there is a gain in terms of Y-PSNR of ~ 0.2 dB on average, corresponding to an average reduction of the bitrate of $\sim 5.4\%$. It is worth mentioning that, since the side reconstructions of the reference and the proposed technique differ only on the interpolated frames, the gain on those frames is actually the double; therefore, in the above mentioned cases, the average quality gain for the interpolated frames is ~ 0.4 dB.

B. Comparison over lossy transmission network

First, a performance comparison with the reference method as a function of the packet loss rate is shown in Fig. 4. An encoded video sequence, obtained as a concatenation of the sequences in Tab. I at the highest quality, is affected by packet losses, modelled as independent and identically distributed Bernoulli random variables with success probability p equal to the loss rate. The same packet loss rate could actually lead to different frame loss rates, since not only the frame or frames in the packet will be lost, but also all the frames predicted upon them (however, it is always only one description that is affected). To reduce the impact of loss propagation, on lossy channels, the stream is usually encoded with a closed GOP, and the size of the GOP must be relatively small ($8 \sim 16$ frames). In order to avoid a bias due to the GOP structure, the results presented here are obtained averaging a number of simulations with the same value of packet loss rate. Both techniques, for
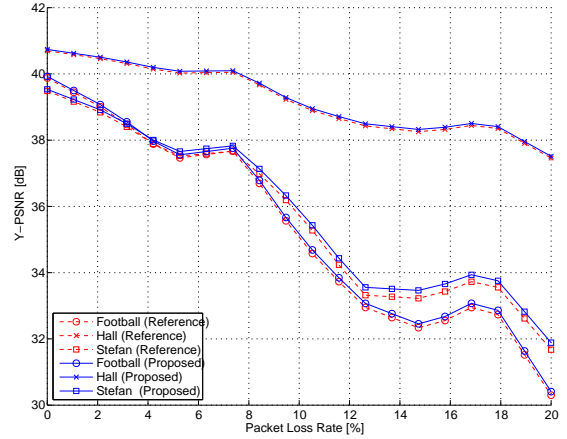
each image of the decoded sequence, use central decoding whenever both description are received, side decoding if only one description is received, and concealment (freeze of the last decoded image) if both descriptions are lost. As expected, sequences with higher motion content are more affected by packet loss; however, our technique slightly but consistently outperforms the reference method.

Then, a performance comparison with the reference method in a simulated scenario is illustrated in Fig. 5 and 6. The encoded sequence in this test is transmitted over a mobile ad-hoc network (see Tab. II) using the ABCD protocol described in Sec. IV-A. The mobile nodes' interfaces parameters are based on the specifications of the ORiNOCO 11 b/g card [24]. The mobility model detailed in Sec. IV-B generates a packet loss patter typical of MANETs, where losses may be due to the fact that the network is no longer connected (as a group of nodes moved away from the source) or to collisions, caused by a high local density (as groups of nodes move towards each others). The figures show the Probability Density Function (PDF) and the Cumulative Distribution Function (CDF) of the visual quality of a received frame, estimated with the Parzen window method [25]. The PDF of the improved method is more concentrated towards high PSNR values, as it was to be expected after results shown in Tab. I. However, here we give a finer resolution picture of the two methods. The two PDFs are quite close for high values of the Y-PSNR — which correspond to the use of the central decoder. However, for lower values of the Y-PSNR, the proposed method outperforms the reference technique since very low values of Y-PSNR are less probable.

VI. CONCLUSIONS

In this work, we presented an MDC scheme for double description based on high-order motion interpolation (HOMI), integrated it into a cross-layer framework for video streaming over mobile ad-hoc networks. The results show a consistent

Sequence	Bitrate range [kbps]	Y-PSNR Gain (Central)	Bitrate Variation (Central)	Y-PSNR Gain (Side)	Bitrate Variation (Side)
foreman	139 ∼ 349	+0.04 dB	−0.9 %	+0.15 dB	−3.8 %
foreman	195 ∼ 502	+0.04 dB	−0.9 %	+0.19 dB	−5.1 %
foreman	248 ∼ 719	+0.04 dB	−0.9 %	+0.22 dB	−6.4 %
football	271 ∼ 714	+0.03 dB	−0.7 %	+0.16 dB	−5.9 %
football	391 ∼ 993	+0.03 dB	−0.6 %	+0.17 dB	−5.7 %
football	505 ∼ 1355	+0.03 dB	−0.5 %	+0.18 dB	−5.4 %
stefan	356 ∼ 843	+0.03 dB	−0.5 %	+0.12 dB	−3.0 %
stefan	490 ∼ 1161	+0.03 dB	−0.6 %	+0.14 dB	−3.5 %
stefan	609 ∼ 1603	+0.04 dB	−0.6 %	+0.15 dB	−3.8 %
coastguard	218 ∼ 767	+0.03 dB	−0.8 %	+0.11 dB	−3.5 %
coastguard	352 ∼ 1140	+0.03 dB	−0.7 %	+0.13 dB	−3.5 %
coastguard	490 ∼ 1614	+0.03 dB	−0.6 %	+0.14 dB	−3.5 %
city	178 ∼ 484	+0.01 dB	−0.2 %	+0.06 dB	−1.3 %
city	267 ∼ 674	+0.01 dB	−0.2 %	+0.07 dB	−1.8 %
city	346 ∼ 920	+0.01 dB	−0.2 %	+0.09 dB	−2.2 %
hall	120 ∼ 273	+0.03 dB	−0.6 %	+0.05 dB	−1.0 %
hall	161 ∼ 393	+0.03 dB	−0.8 %	+0.05 dB	−1.5 %
hall	200 ∼ 597	+0.03 dB	−0.9 %	+0.05 dB	−2.0 %
akiyo	82 ∼ 180	+0.04 dB	−0.6 %	+0.04 dB	−0.8 %
akiyo	109 ∼ 246	+0.04 dB	−0.6 %	+0.05 dB	−1.0 %
akiyo	135 ∼ 331	+0.04 dB	−0.7 %	+0.06 dB	−1.2 %
flower	437 ∼ 1087	+0.01 dB	−0.1 %	+0.02 dB	−0.3 %
flower	616 ∼ 1471	+0.01 dB	−0.1 %	+0.02 dB	−0.4 %
flower	785 ∼ 1939	+0.02 dB	−0.2 %	+0.02 dB	−0.4 %
mobile	479 ∼ 1160	+0.01 dB	−0.1 %	−0.01 dB	+0.3 %
mobile	660 ∼ 1636	+0.01 dB	−0.1 %	−0.01 dB	+0.2 %
mobile	832 ∼ 2245	+0.02 dB	−0.2 %	−0.01 dB	+0.1 %

Table I

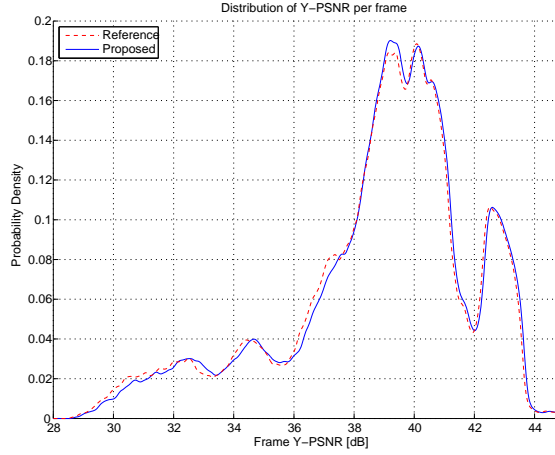BJONTEGAARD METRIC OF THE PROPOSED TECHNIQUE WITH RESPECT TO THE REFERENCE OVER VARIOUS SEQUENCES.

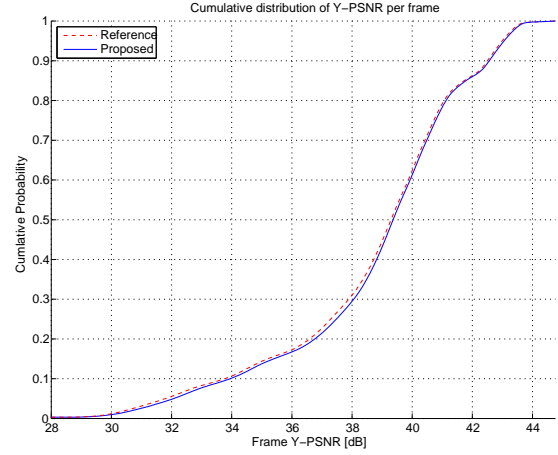Figure 5. Probability density function of the Y-PSNR of the decoded frames(Bitrate 1.0 Mbps, ∼ 25 % of packets lost).

Figure 6. Cumulative distribution function of the Y-PSNR of the decoded frames (Bitrate 1.0 Mbps, ∼ 25 % of packets lost).

Number of nodes	100
Number of groups	10
Playground size	100 m × 100 m
Nominal range	25 m
Average speed	2.0 m/s
Average pause time	2.0 s
Simulation time	600 s

Table II

SIMULATION PARAMETERS. THE NETWORK INTERFACE CONTROLLERS ARE BASED ON THE SPECIFICATIONS OF THE ORINOCO 11 B/G CARD.

gain over a set of test sequences, in some instances as high as 0.2 dB of video quality for a fixed bitrate, and a bitrate reduction of 6 % for a fixed quality.

Our future work will focus on providing the technique with a tuning mechanism for the redundancy between the sub-streams, *i.e.*, by different quantisation of the two sub-streams, or by a variable block size for the α sequence.

REFERENCES

[1] V. K. Goyal, "Multiple description coding: compression meets the network," *IEEE Signal Processing Mag.*, vol. 18, no. 5, pp. 74–93, 2001.

[2] J. G. Apostolopoulos, "Reliable video communication over lossy packet networks using multiple state encoding and path diversity," in *Proc. of Intern. Symp. Visual Comm. and Image Proc.*, 2001.

[3] C. Tillier, T. Petrişor, and B. Pesquet-Popescu, "A motion-compensated overcomplete temporal decomposition for multiple description scalable video coding," *EURASIP J. Image Video Proc.*, vol. 1, pp. 1–12, 2007.

[4] I. Radulovic, P. Frossard, Y.-K. Wang, M. Hannuksela, and A. Hallapuro, "Multiple description video coding with H.264/AVC redundant pictures," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, no. 1, pp. 144–148, 2010.

[5] S. Shirani, M. Gallant, and F. Kossentini, "Multiple description image coding using pre-and post-processing," in *IEEE ITCC '01: International Conference on Information Technology: Coding and Computing*, 2001, pp. 35–39.

[6] E. Kozica, D. Zachariah, and W. Kleijn, "Interlacing intraframes in multiple-description video coding," in *Proc. of IEEE Intern. Conf. Image Proc.*, vol. 4, 2007, pp. IV–277.

[7] D. Wang, N. Canagarajah, D. Redmill, and D. Bull, "Multiple description video coding based on zero padding," in *IEEE ISCAS '04: International Symposium on Circuits and Systems*, vol. 2, 2004, pp. II–205.

[8] C. Greco, M. Cagnazzo, and B. Pesquet-Popescu, "H.264-based multiple description coding using motion compensated temporal interpolation," in *Proc. of IEEE Worksh. Multim. Sign. Proc.*, Saint-Malo, France, October 2010.

[9] C. Greco and M. Cagnazzo, "A cross-layer protocol for cooperative content delivery over mobile ad-hoc networks," *Inderscience Intern. J. of Commun. Networks and Distrib. Syst.*, vol. 7, no. 1–2, pp. 49–63, 2011.

[10] C. Guillemot, F. Pereira, L. Torres, T. Ebrahimi, R. Leonardi, and J. Ostermann, "Distributed monoview and multiview video coding," *IEEE Signal Processing Mag.*, vol. 24, pp. 67–76, September 2007.

[11] X. Artigas, J. Ascenso, M. Dalai, S. Klomp, D. Kubasov, and M. Ouaret, "The DISCOVER codec: architecture, techniques and evaluation," in *Proc. of Pict. Cod. Symp.*, 2007.

[12] M. Cagnazzo, M. Antonini, and M. Barlaud, "Mutual information-based context quantization," *Signal Proc.: Image Comm. (Elsevier Science)*, vol. 25, no. 1, pp. 64–74, 2010.

[13] G. Petrazzuoli, M. Cagnazzo, and B. Pesquet-Popescu, "High order motion interpolation for side information improvement in DVC," in *IEEE ICASSP '10: International Conference on Acoustics Speech and Signal Processing*, pp. 2342–2345.

[14] M. Frodigh, P. Johansson, and P. Larsson, "Wireless ad-hoc networking: the art of networking without a network," *Ericsson Review*, vol. 4, pp. 248–263, 2000.

[15] M. Gerla, "From battlefields to urban grids: new research challenges in ad-hoc wireless networks," *Elsevier J. on Pervasive and Mobile Comput.*, vol. 1, no. 1, pp. 77–93, 2005.

[16] M. van der Schaar and N. Sai Shankar, "Cross-layer wireless multimedia transmission: challenges, principles, and new paradigms," *IEEE Trans. Wireless Commun.*, vol. 12, no. 4, pp. 50–58, 2005.

[17] C. Greco, M. Cagnazzo, and B. Pesquet-Popescu, "ABCD : Un protocole cross-layer pour la diffusion vidéo dans des réseaux sans fil ad-hoc," in *Proc. of GRETSI*, Bordeaux, France, September 2011, [Accepted for publication].

[18] A. El Fawal, J. Le Boudec, and K. Salamatian, "Multi-hop broadcast from theory to reality: practical design for ad-hoc networks," in *Proc. of IEEE/ACM Intern. Conf. on Autonomic Comput. and Commun. Syst.*, 2007.

[19] J. Broch, D. A. Maltz, D. B. Johnson, Y.-C. Hu, and J. Jetcheva, "A performance comparison of multi-hop wireless ad-hoc network routing protocols," in *Proc. of ACM Intern. Conf. on Mobile Comput. and Networking*, 1998, pp. 85–97.

[20] B. Zhou, K. Xu, and M. Gerla, "Group and swarm mobility models for ad hoc network scenarios using virtual tracks," in *MILCOM '04: Military Communications Conference*, vol. 1, 2004, pp. 289–294.

[21] X. Hong, M. Gerla, G. Pei, and C. Chiang, "A group mobility model for ad hoc wireless networks," in *Proc. of ACM Intern. Conf. on Modeling Anal. and Simul. of Wireless and Mobile Syst.*, 1999, pp. 53–60.

[22] K. Sühring, "JM reference software release 14.2," Source Code, September 2008. [Online]. Available: http://iphome.hhi.de/suehring/tml/

[23] G. Bjontegaard, "Calculation of average PSNR differences between RD-curves," in *ITU-T VCEG Meeting*, Austin, TX, USA, April 2001.

[24] "Proxim ORiNOCO 11b/g client PC card specifications," 2006. [Online]. Available: http://www.proxim.com/downloads/products/cp/DS_0806_11b_gPCCard_A4HR.pdf

[25] E. Parzen, "On estimation of a probability density function and mode," *Ann. of Math. Stat.*, vol. 33, no. 3, pp. 1065–1076, 1962.