International Journal of Computer Applications (0975 – 8887)
Volume 66– No.13, March 2013

Scalable Video Streaming Techniques over Cooperative Relay Networks

Sajid Nazir
School of Engineering
King's College
University of Aberdeen
Aberdeen, AB24 3UE, UK

Fabio Verdicchio
School of Engineering
King's College
University of Aberdeen
Aberdeen, AB24 3UE, UK

GorryFairhurst
School of Engineering
King's College
University of Aberdeen
Aberdeen, AB24 3UE, UK

ABSTRACT

This paper investigates a communication method using unequal loss protected layered video and explores use of this over a cooperative relay network. The method is adapted to prioritized video transmission to explore two prioritized layering (scalable) video techniques. Cooperative communication networks can reduce power consumption and benefit wireless networks, where intermediate nodes act as relays of transmitted data, increasing the coverage and network throughput. The methods are therefore evaluated in the context of a number of different relay collaboration strategies using a finite state Markov chain (FSMC) loss model for the wireless channel. Random linear codes are ideally suited to multiple path transmission because the aggregated received symbols over multiple paths assist decoding at the receiver. The results illustrate the advantages for these application scenarios using each considered strategy and analyze the impact in a relay-based network.

General Terms

Video communications, Relay networks, Scalable video.

Keywords

Expanding window fountain codes, Random linear codes, Co-operative relay networks.

1. INTRODUCTION

There has been a significant growth and interest in methods for loss-tolerant communication of video over multi-hop (cooperative) wireless networks. In cooperative communications, the network nodes collaborate to transmit information from a source node to a destination node (receiver) over multiple parallel paths [1]. This not only results in a reduced power consumption (important for low power wireless devices), but may also increase communication reliability/throughput. Many relay collaboration strategies have been proposed, but the analysis in this paper is limited to the comparison of Amplify-and-Forward (AF) and Decode-and-Forward (DF) strategies.

The paper considers two (related) video coding formats. First, the H.264 Advanced Video Coding (AVC) standard [2], a state-of-the-art coding scheme that is gaining widespread use in many applications. Second, the more recently introduced scalable extension of H.264/AVC, termed H.264/Scalable Video Coding (SVC) [3] which adds a set of scalability options to H.264/AVC. One such scalability option is SNR (or quality) scalability. SNR scalability supplements a base layer by defining one or more enhancement layers. Each layer can be used by a receiver to progressively add to the quality of the reconstructed video.

The data-partitioning (DP) feature of H.264/AVC allows partitioning the encoded video stream into prioritized layers, each of decreasing importance for video reconstruction. This creates a layered video coded output, which is comparable to the SNR scalability feature of H.264/SVC. DP has been used in [4] in combination with error protection.

For real-time video streaming applications, packet forward error correction (FEC) is a favoured approach to mitigate the effects of network packet loss. This is preferable to packet retransmission, because it avoids complexity and delay, eliminates the need to return acknowledgement data to the sender [5], [6], [7]. This paper restricts the analysis to the case without feedback.

Packet FEC is, especially suited to multicast/broadcast as specified in FECFRAME which is a standard for using FEC codes to provide protection against packet loss [8]. Random linear codes (RLC) [9] offer near-capacity performance, even for short lengths of codeword, but suffer from high decoding complexity of a Gaussian Elimination (GE) decoder as codeword length increases.

The degree of protection provided by the FEC codes could be equal for different layers, which is termed as an Equal Loss Protection (ELP) scheme. ELP treats all layers equally with no prioritized transmission for the important (base) layer packets. An Unequal Loss Protection (ULP) scheme may provide a higher degree of protection, where "important" (prioritised) video packets are more protected. ULP schemes may also be designed to prioritise important packets according to the prevailing channel condition.

A physical layer based solution with fountain codes is proposed in [10]. SVC transmission [11] has been compared with multiple description coding using FEC with Raptor codes, but only DF is considered. A comparison of AF and DF with Turbo codes [12] concludes that the DF schemes have better performance on the average.

This paper bases analysis on the expanding window (EW) method [13], with RLC applied to streaming of layered video over a multi-hop relay network. In this approach, video data is partitioned into two prioritized windows/layers, based on its importance for video reconstruction. A range of selection probability is assigned for each window to afford a selected

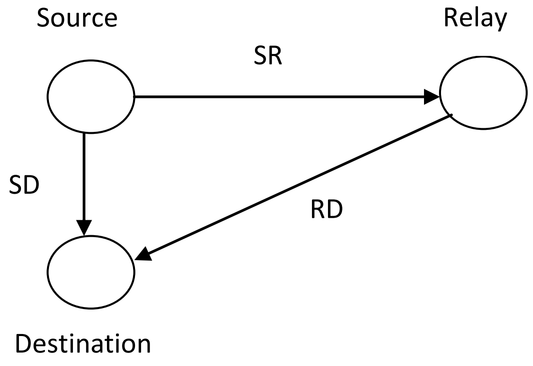

Fig 1: Simple 3-node relay model

degree of protection to the video based on the decoding importance of each layer.

For video broadcast/multicast applications, it can also be advantageous to relate the degree of protection to the encoding layer depending on the prevailing channel conditions, i.e., the available capacity and the packet loss ratio (PLR). This concept has been employed in [14] for video transmission over DVB-H channel.

Our simulation results show that application layer EW RLC can be effectively used to protect video. The key contributions are: (1) Analysis of AF [15] and DF schemes for streaming layered video with EW-RLC over relay networks, such as LTE-A [16] (2) Comparison of PSNR performance of H.264/AVC and H.264/SVC for SNR scalability. Although analysed in the context of a single relay node, the results can be extended to include more than one relay node.

The remainder of this paper is structured as follows: Section 2 provides the necessary background. Section 3 describes the proposed system. The video configuration and results for H.264/AVC and H.264/SVC are presented in Section 4 and 5 respectively. Discussion and analysis are provided in Section 6. Finally, Section 7 provides the conclusion.

2. BACKGROUND

This section briefly reviews the operation of multi-hop relay networks, layered (scalable) video, and random linear codes.

2.1 Multi hop Relay Networks

The nodes in a multi-hop relay network typically communicate over unreliable wireless channels and have limited transmission power. The analysis uses a simple 3-node relay model, which comprises a source (S), relay (R), and a destination (D), as shown in Fig. 1. We assume the destination is able to simultaneously receive transmissions from S and R and that the relay is half-duplex that is it cannot transmit and receive at the same time. The channel from the source-to-relay (SR) and relay-to-destination (RD) was assumed to be better than the direct channel between the source and destination (SD). If the SD channel has low loss, then there may be no requirement for relay collaboration. However, if the direct channel SD experiences loss, then the relay may enter the collaboration state, where it tries to improve the received video quality by additional decoding of data received via the relay at the destination node.

The relay collaboration could use an AF scheme [12], where the relay acts as a repeater, simply re-transmitting each received packet towards the destination. The same code is transmitted on the channel SD, hence if both packets are received by the destination node, it does not offer a decoding advantage. At the destination, decoding is attempted at the end of transmission.

For the relay collaboration schemes using DF, the relay continues to accumulate the packets until such time that the received data is decodable. As soon as a relay completes decoding, it re-encodes the video and starts transmitting. Early decoding at the relay is advantageous since it reduces end-to-end delivery delay when poor channel conditions result in loss on the direct link, SD, from the source to destination.

2.2 Scalable/Layered Video

H.264/AVC is widely deployed [2] in multimedia applications. It provides many error-resilience features to mitigate the effect of lost packets. One scheme, available in the extended profile is DP [17], which supports partitioning of a slice in up to three partitions (NAL units), based on the importance of the encoded video syntax elements for video reconstruction.

Partition A contains the most important data comprising slice header, quantization parameters, and motion vectors. Partition B contains the intra-coded macroblocks residual data, and partition C contains inter-coded macroblocks residual data. This allows the network transport to assign a different protection to the different partitions based on their importance.

The decoding of DP A is always independent of DP B and C. However, if DP A is lost the remaining partitions cannot be utilized. The decoding of DP B is possible without DP C, but not the other way around. The Constrained Intra Prediction (CIP) parameter must be set in the H.264/AVC encoder to make DP B independent of DP C.

Quality scalability of H.264/SVC has a coarse-grain quality scalable coding (CGS) feature that can be considered as a special case of spatial scalability, with identical picture sizes for the base and enhancement layers. This supports a few selected bit rates in a scalable bit stream.

In general, the number of supported rate points is identical to the number of layers. When the relative rate difference between successive CGS layers decreases, then CGS becomes less efficient.

2.3 Random Linear Codes

The RLC class of rateless codes has recently become popular [9]. RLC is applied over a source message to produce encoded symbols as random linear combinations of source symbols with coefficients randomly selected from a given finite field [18]. When used as a packet level AL-FEC solution, RLC is simple to implement and provides near-optimal erasure codes for sufficiently large finite field used for creating linear combinations of source symbols (one-byte field GF(256) is usually sufficient [9]). This makes RLC an attractive code as a universal FEC/network coding solution for emerging wireless communication systems, such as LTE-A, and WiMAX. The major limitation of using RLC is the decoding complexity of Gaussian Elimination (GE) decoding, which is polynomial in the number of symbols. However, for short source messages, the decoding complexity is acceptable.

3. THE SYSTEM MODEL

3.1 EW RLC

Expanding window fountain (EWF) codes [13] are a class of ULP fountain codes [19] based on the idea of creating a set

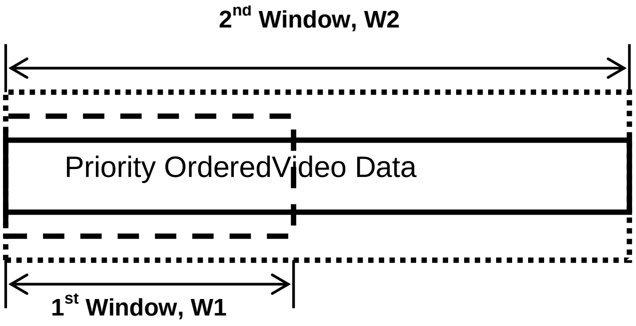

Fig 2: Expanding Window Structure

of "nested windows" over the source block. The rate-less encoding process is then adapted to use this windowing information while producing encoded packets. The EWF concept is used to create EW RLC [20] over two windows.

A set of windows is defined over the groups of source symbols of unequal importance generated as a result of DP H.264/AVC video or CGS of H.264/SVC to obtain source blocks amenable to ULP encoding.

To define a window over a subset of source symbols, the DPs of a particular type have to be aggregated over the entire source data. Coding is performed over a progressively increasing source block subset window aligned with these "most to least importance" subsets. The layout of a window structure with two importance classes is shown in Fig. 2. The first window (W1) has the most important subset of encoded data. The window W2, in addition to its own data also encloses the data of W1. The subset data of W1 is hence the most protected. The size and structure of a window depends upon the elements meeting a particular set of criteria from a specific subset window. The number of windows is governed by the aggregation scheme employed to group the encoded elements. The window with most important data is termed here as higher priority layer (HPL) and the one with least important data as low priority layer (LPL). Thus, HPL and LPL are analogous to the base and enhancement layers as in H.264/SVC.

The decoding of a window is the same as for RLC decoding, in that, a window is considered recoverable if the receiver collects at least the same volume of linearly independent encoded symbols from the window (or the windows contained in it) as were sent in the window [20].

As an initial step in the encoding process of EW RLC, a window is first selected from which the RLC encoded symbol is to be generated. This selection of a window is determined by the probability of selection (PS) of a window, which is a pre-assigned parameter set depending on the relative importance of the layers and the rate available. After a window is selected, standard RLC encoding is performed over only the source packets contained in the particular window [20].

For the decoding process, the received symbols are gathered at the destination. For successful decoding of the entire transmitted video, both the HPL and LPL must be decodable. The decoding of HPL alone yields usable but poorer quality video at the destination. However, decoding of only the LPL is not of value and results in a decoding failure.

Table 1. Relative Partition Sizes- Paris sequence

Partition	Size (bytes)	Size (%)	Cumulative PSNR
IDR	22,281	27.52	-
DP-A	12,838	15.86	-
DP-B	97	0.12	30.32
DP-C	45,732	56.50	39.16
Total	80,948	100	39.16

Table 2. Relative Partition Sizes- Football sequence

Partition	Size (bytes)	Size (%)	Cumulative PSNR
IDR	23,374	28.92	-
DP-A	22,823	28.24	-
DP-B	2893	3.58	25.39
DP-C	31,731	39.26	32.62
Total	80,821	100	32.62

Table 3. Packetization of Layers

Sequence	Number of Packets		
	HPL	LPL	Total
Paris	35	45	80
Football	49	31	80

3.2 Measurement Setup

The well-known video sequences "Paris" and "Football" [21] were used for the simulations. This choice was made because the Paris sequence has a high spatial complexity and the Football sequence has a high temporal complexity.

Both sequences were in the CIF format with 25 frames per second (fps), and the CIP flag set. The Group of Pictures, GOP, size was 16 frames, with an IPPP... structure. The total number of video packets for each video configuration was about the same, to enable a comparison of the schemes.

The video was packetized into packets of 1024 bytes. A packet header of 6 bytes was assumed, anticipating the use of packet header compression, common on wireless links. Each simulation was repeated 1000 times and the results averaged.

A set of simulations was performed using EW-RLC to compare the performance of AF and DF for transmission over the relay model. A finite-state Markov chain (FSMC) [22] channel model was used to model the network links.

The ELP scheme, with no priority for HPL is compared to the ULP schemes. For ULP schemes, the probability of selection (PS) for HPL was varied among the values of 0.50, 0.60, 0.80 and 1.0. This created four ULP schemes. With a PS of 0.60, the symbols were generated from the HPL with a 60% probability, hence, affording better protection to HPL. The size of the HPL also governs selection of an appropriate PS. The PS should be higher for ULP compared to the size (percentage) of HPL in comparison with the LPL.

The simulations considered the case where there was a better channel condition for the relay than using the direct path. The PLR for the RD channel was constant at 0.025. The PLR for the direct SD channel was varied within the set: 0.05, 0.10, 0.20, 0.30, 0.40, and 0.50.

Loss concealment is common for packet video receivers. We considered a simple method that replaces a missing frame (or an entirely missing GOP) with the last frame previously decoded.

4. VIDEO TRANSMISSION WITH H.264/AVC

4.1 System Configuration

The DP feature was used to create encoded data with partitions. These partitions were used to form two layers of packet video that were protected with ULP using EW-RLC.

The breakdown of the video into the constituent partitions for the Paris and Football sequences are respectively shown in Tables 1 and 2. After encoding, the IDR frame together with DP A and B were placed in the HPL, and the remaining portion, that is, DP C, was placed in the LPL.

The packetization details for each sequence are given in Table 3. This table shows that the number of encoded packets for the HPL using the Football sequence is substantially higher compared to the Paris sequence. This implies that the same ULP strategy does not perform equally when encoding both sequences, demonstrating that performance is linked to video content structure.

4.2 Simulation Results

The results for the Paris sequence are shown in Figure 3. The scheme AF50 denotes an AF scheme with a PS of 0.5. The notation AF-ELP denotes protection over the entire video data, treating all data with the same priority. The results show that the performance of the ELP scheme is much lower than when using ULP. DF-ELP has the poorest performance, because it is unable to decode using the relay, and hence preventing relay collaboration.

Two notable cases are AF100 and DF100. These exactly overlap, but provide a consistent PSNR contribution of 30.32 dB irrespective of the PLR. With an over-protection for HPL the performance is limited to the PSNR achieved by only receiving the HPL (see Table 1).

The performance of DF improves with an increase in PS of the HPL, such that DF80 provides significantly better performance. This is attributed to the higher protection offered by the HPL, the probability of an early decoding of HPL at the relay node increases, resulting in an early start of the relay collaboration phase, reducing time for decoding at the destination.

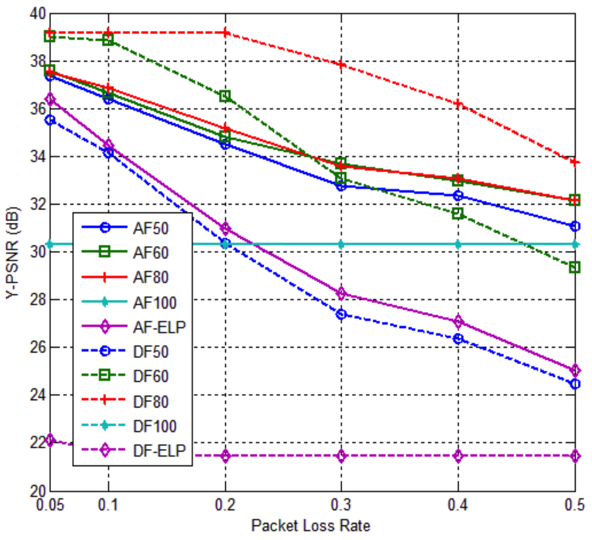

Fig 3: PSNR vs. PLR for Paris Sequence (AVC)

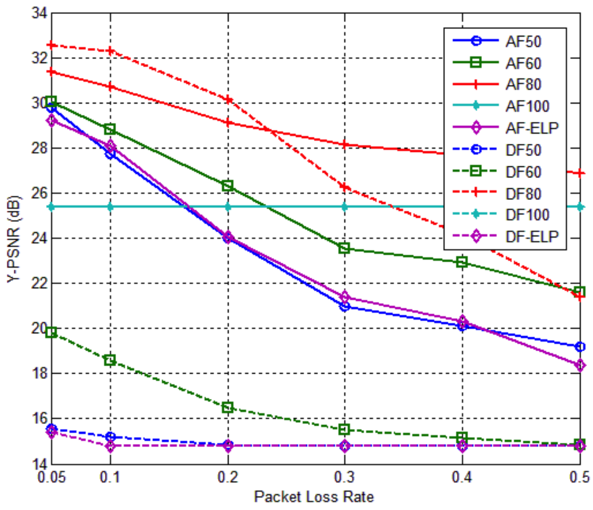

Fig 4: PSNR vs. PLR for Football Sequence (AVC)

The results for the Football sequence are shown in Figure 4. The performance of DF schemes was generally reduced at a low PS, where the HPL was much smaller than the LPL in the Football sequence; hence successful decoding of HPL was delayed using the DF scheme. This resulted in a sharp decrease in performance at a high PLR, with no relay participation, especially when using a DF scheme with a low PS. The AF100 and DF100 were able to maintain performance across the range of PLR.

5. VIDEO TRANSMISSION WITH H.264/SVC

5.1 System Configuration

CGS was used to create scalable video. Two layers of video data are created, which could then be protected, by providing unequal protection by using EW-RLC. A Quantization parameter (QP) was used to tailor the size of the base and

Table 4. Relative Partition Sizes- Paris sequence

Layer	Size (bytes)	Size (%)	Cumulative PSNR
HPL	34,814	43.44	27.91
LPL	45,337	56.56	30.32
Total	80,151	100	30.32

Table 5. Relative Partition Sizes- Football sequence

Layer	Size (bytes)	Size (%)	Cumulative PSNR
HPL	29,738	36.64	32.34
LPL	51,441	63.36	32.34
Total	81,179	100	37.14

Table 6. Packetization of Layers

Sequence	Number of Packets		
	HPL	LPL	Total
Paris	30	51	81
Football	34	45	79

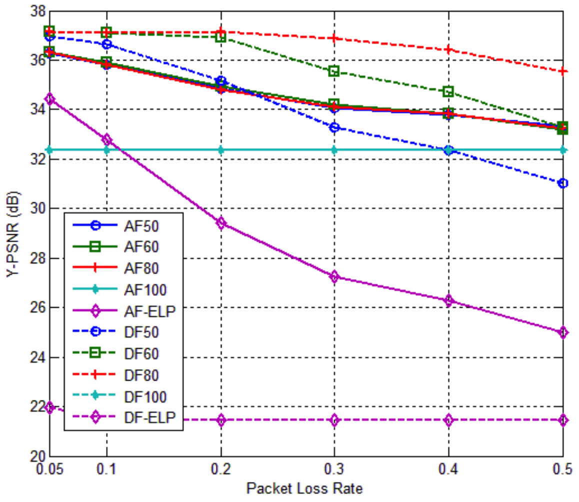

Fig 5: PSNR vs. PLR for Paris Sequence (SVC)

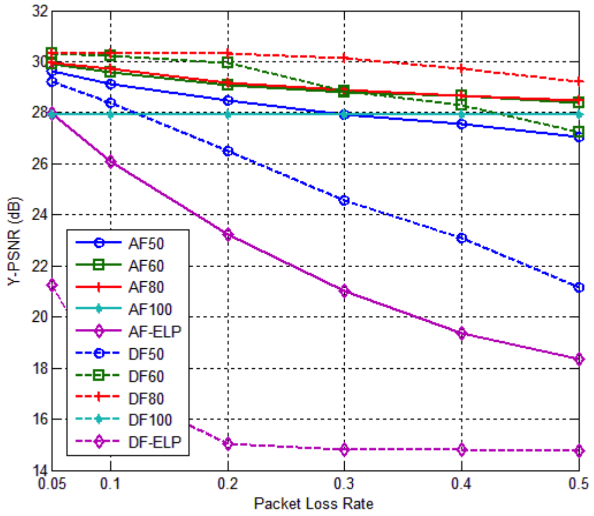

Fig 6: PSNR vs. PLR for Football Sequence (SVC)

enhancement layers. This also affects the resulting PSNR achieved for each layer.

The breakdown of video data into the two layers for the Paris sequence and Football sequence are respectively shown inTables 4 and 5. The packetization details are shown in Table 6.

5.2 Simulation Results

The results for the Paris sequence are shown in Figure 5, showing that the performance of the ELP scheme is poorer than for the ULP schemes. The schemes AF100 and DF100 maintain a consistent PSNR contribution of 32.34 dB irrespective of the PLR, which is the same as the PSNR achievable using only the HPL. The performance of the DF schemes was generally better than using AF. However, as previously, the performance of DF schemes improved at a higher PS for the HPL. Again, DF80 provides the best overall performance.

The results for the Football sequence are shown in Figure 6. These results confirm the results obtained with Paris sequence. However, the performance of DF schemes have slightly poorer performance than for the Paris sequence. This degradation is attributed to the larger HPL.

6. COMPARISON AND ANALYSIS

The performance of H264/AVC and H.264/SVC can be compared using Figures 3 and 5. However, the difference in HPL sizes must also be considered. Tables 1 and 4 show that the size of HPL for H.264/AVC is larger with a lower PSNR compared to H.264/SVC. H.264/SVC also imposes a loss in coding efficiency, so the PSNR for H.264/AVC is slightly better than when using H.264/SVC.

An interesting effect arises when comparing AF and DF schemes with increasing PLR. This results from the contribution of the HPL compared to combined decoding of both HPL and LPL. This effect is observed for AF60 and DF60 in Figure 3, where the value for DF60 drops below that for AF60 at a higher PLR. The effect is also seen in Figure 3 and 5, i.e., the results for DF80 for H.264/AVC is better at

Table 7. Effect of PS on Layer Decoding

HPL Size (%)	PS (%)	Decoding (%), (PE = 0.3)					
		HPL		HPL+LPL		Failure	
		AF	DF	AF	DF	AF	DF
30	30	43	38	21	4	36	58
	40	78	60	21	37	1	3
	50	81	17	19	83	0	0
70	70	37	2	22	0	41	98
	80	76	47	23	1	1	52
	90	77	91	23	7	0	2

low PLR, but the performance of H.264/SVC is better for a higher PLR. Similar analysis was observed using AF. Theseconclusions are also substantiated with the results for the Football sequence.

The effect of HPL size on the performance of the AF and DF schemes is discussed next. The HPL size is chosen to be 30 and 70 (out of a total of 100) video packets for each of the two considered configurations. To quantify the effect, the PE was kept at 0.3; while the PS was increased in steps and the resulting performance of the AF and DF schemes was measured, as summarised in Table IV. The decoding (%) implies a count of the cases when for instance only the HPL is decoded, or the GE fails to recover any useful FEC packets (the column titled failure). Fig. 4, shows that a larger HPL size adversely impacted the performance of the DF scheme.

The performance of DF was better as the PS increased (Table 7). Hence, the PS used for the HPL needs to be selected based on the size of the HPL in the video GOP data. This effect can similarly be seen when using the AF scheme. It is possible that future systems can select the PS using an adaptive scheme that assigns the PS depending upon the size of the HPL (for each GOP) increasing the video quality throughout.

DF80 offered the best choice over the entire range of PSNRin a scenario where no feedback mechanism was available.

The advantage of H.264/AVC is that once the DP have been created, it is possible to selectively drop selected partitions to yield a base layer of the required rate for transmission.

The use of relays as envisaged in emerging wireless standards brings consistent video services to the end user. Our comparison of the two scalable video standards using ULP over a relay network has not previously been analysed in the literature but are consistent with other results for relay transmission with AF and DF using other packet FEC codes [12]. We suggest that scalable video transmission over relay networks could be adapted to the channel conditions (e.g. PLR). Relay transmission may be optimised by favouring transmission of base layer packets, thus improving overall video quality at a higher PLR.

Although study did not explore use of a feedback channel from the destination, if this is available the method could be extended to result in substantial power savings. The advantages of DF over AF become significant with an early decoding at the relay, which in turn depends on selecting an appropriate ULP scheme and packet FEC code.

7. CONCLUSIONS

This paper compared the transmission of scalable video data using two most popular video coding standards over a relay network. It shows that the reliability, throughput and coverage of the wireless nodes can be increased using rate-less codes in conjunction with Expanding window random linear codes. Unequal loss protection schemes were seen to perform better than using an equal loss protection scheme. In some cases of extremely high loss, it was seen to be advantageous to just transmit the HPL for a particular Group of Pictures.

8. REFERENCES

[1] Molisch, A.F.,Mehtra, N.B.,Yedidia, J.S.,and Zhang, J., Cooperative relay networks using fountain codes,*IEEE GLOBECOM 2006*, pp. 1-6, 2006.

[2] Wiegand, T., Sullivan, G.,Bjøntegaard, G., and Luthra, A.,Overview of the H.264/AVC video coding standard,*IEEE Trans. Cir. Sys. Video Techn.*,vol. 13, no. 7, pp. 560-576, Jul. 2003.

[3] Schwarz, H., Marpe, D., and Wiegand, T., Overview of the scalable video coding extension of the h.264/avc standard,*IEEE Trans. Circ. and Syst. for Video Techn.*, vol. 17, no. 9, pp. 1103–1120, 2007.

[4] Rouzbeh, R., Fleury, M., Altaf, M., Simak, H., and Ghanbari, M., H.264 video streaming with data-partitioning and growth codes, *IEEE Int. Conf. on Image Processing(ICIP)*, Cairo, Nov 7-10. 2009, pp. 909-912

[5] Liu, Y., A Low Complexity Protocol for Relay Channels Employing Rateless Codes and Acknowledgement,*IEEE Intl Symposium on Info Theory*, pp. 1244-1248, 2006.

[6] Qin, S., Feng, G., and Zhang, Y., Cooperative Communications for Reliable Data Transport with Fountain Codes,*Journal of Communications*, Vol 5, No 4 (2010), 340-347, Apr 2010.

[7] Ahmad, S., Hamzaoui, R, and Al-Akaidi, M., Adaptive Unicast Video Streaming With Rateless Codes and Feedback,*IEEE Trans. Circuits and Systems for Video Technology*, Vol 20, pp 275-285, 2010.

[8] Watson, M., Begen, A., and Roca, V., Forward Error Correction (FEC) Framework, IETF RFC 6363, October 2011.

[9] Wang, M., andLi, B., How practical is network coding? *IEEE IWQoS 2006, Intl. Workshop on Quality of Service*, pp.274-278.

[10] Zhang, Y.Y.,Ma, Y., andTafazolli, R., Improved coded cooperation schemes for wireless communication*WiOpt 2008 - 6th Intl. Symposium on Modeling and Operation*, pp.533-538, April 2008.

[11] Park, J.A.,Zhao, Z.,Suh, D.Y., andOstermann, J., Effective Relay Communication for Scalable Video Transmission, *World Academy of Science, Engineering and Technology*, pp. 1156-1160, 2010.

[12] Yu, M.,and Li, J., Is amplify-and-forward practically better than decode and-forward or vice versa?in*Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 3, March 2005, pp. 365–368.

[13] Vukobratovic, D.,Stankovic, V.,Sejdinovic, D.,Stankovic, L., andXiong, Z., Expanding window fountain codes for scalable video multicast, *IEEE Trans. Multimedia*, vol. 11, no. 6, pp. 1094–1104, 2009.

[14] Nazir, S., Stankovic, V.,and Vukobratovic, D., Expanding Window Random Linear Codes for data partitioned H.264 video transmission over DVB-H network, in *Proc. IEEE Int. Conf. on Image Processing (ICIP)*, Sep 2011, pp. 2205-2208.

[15] Li, J.,Hu, Z.,and Wang, Y., DF/AF cooperative relay in LTE-A, *IEEE Vehicular Tech. Conf.*, pp. 1-5, 2010.

[16] Yang, Y.,Hiu, H.,Xu, J., andMao, G., Relay technologies for WIMAX and advanced mobile system,*IEEE Comm. Magazine*, Oct. 2009.

[17] ITU-T Recommendation H.264 (03/2010) Advanced video coding for generic audiovisual services.

[18] MacKay, D.J.C., Fountain codes,*IEE Proc. Comm.*, vol. 152, pp. 1062–1068, 2005.

[19] Mitzenmacher, M., Digital fountains: A survey and look forward, in *Proc. IEEE Inf. Theory Workshop 2004*, pp. 271–276, Oct. 2004.

[20] Vukobratovic, D.,and Stankovic, V., Unequal error protection random linear coding for multimedia communications, in *Proc. MMSP-2010 IEEE Multimedia and Signal ProcessingWorkshop, Saint Malo, France*, October 2010.

[21] http://www2.tkn.tu-berlin.de/research/evalvid/cif.html.

[22] Zhang, Q., andKassam, S., Finite-State Markov model for Rayleigh fading channels, *IEEE Trans. Commun*, vol. 47, no. 11, pp. 1688-1692, Nov. 1999.