ALD: Adaptive Layer Distribution for Scalable Video

Jason J. Quinlan
Department of Computer Science
University College Cork, Ireland
j.quinlan@cs.ucc.ie

Ahmed H. Zahran
Electronics and Electrical Communications Department
Cairo University, Egypt
azahran@eece.cu.edu.eg

Cormac J. Sreenan
Department of Computer Science
University College Cork, Ireland
cjs@cs.ucc.ie

ABSTRACT

Bandwidth constriction and datagram loss are prominent issues that affect the perceived quality of streaming video over lossy networks, such as wireless. The use of layered video coding seems attractive as a means to alleviate these issues, but its adoption has been held back in large part by the inherent priority assigned to the critical lower layers and the consequences for quality that result from their loss. The proposed use of forward error correction (*FEC*) as a solution only further burdens the bandwidth availability and can negate the perceived benefits of increased stream quality.

In this paper, we propose Adaptive Layer Distribution (*ALD*) as a novel scalable media delivery technique that optimises the trade-off between the streaming bandwidth and error resiliency. ALD is based on the principle of layer distribution, in which the critical stream data is spread amongst all datagrams thus lessening the impact on quality due to network losses. Additionally, ALD provides a parameterised mechanism for dynamic adaptation of the scalable video, while providing increased resilience to the highest quality layers. Our experimental results show that ALD improves the perceived quality and also reduces the bandwidth demand by up to 36% in comparison to the well-known Multiple Description Coding (*MDC*) technique.

Categories and Subject Descriptors

H.5.1 [**Multimedia Information Systems**]: Video; C.2.5 [**Local and Wide-Area Networks**]: Internet

General Terms

Design, Performance, Experimentation

Keywords

Scalable Video, Adaptive Video, Media Streaming, Lossy Networks, Layered Coding, Error Resilience, Layer Distribution

1. INTRODUCTION

MMSys '13, February 26-March 1, 2013, Oslo, Norway.

Recent years have featured a dramatic rise in the volume of video streaming traffic over the Internet and mobile networks. This increase contributes to a widely acknowledged bandwidth "crunch" at the network edge. In mobile networks, the competition for scarce wireless resources has a multi-fold impact on streaming applications. First, datagram losses may occur in the buffers of bottleneck links. Such losses, when added to typical link handover losses in wireless systems, reduce the streamed video quality. Second, the shared nature of wireless resources produces noticeable resource variation during video sessions. Hence, using adaptive video streaming schemes that can scale to the available bandwidth evolves as a crucial need in wireless systems.

Multi-bitrate streaming and layered coding are the two main streaming models that support video scalability. Multi-bitrate stream encoding is a mechanism by which a media clip is encoded as several streams each with a different bitrate. Early adoption of this mechanism limited the video choice of the user to only one of the available bit-rates. Subsequent change in quality required user video re-selection. Recent multi-bitrate technologies such as HTTP streaming enable stream quality adaptation by separating the stream into video sections, commonly known as segments. These segment points permit the decoder to switch between the different qualities based on a number of settings, such as user requirements or network conditions. Examples of this technique include 3GPP Packet-switched Streaming Service (PSS) [5] and Dynamic Adaptive Streaming over HTTP (DASH) [17]. The main limitations of a multi-bitrate scheme include large storage and bandwidth requirements.

In layered coding, a high quality media clip is fragmented into N layers, which consist of a single base layer and numerous enhancement layers. The base layer is generally constructed from the key frames and supports coarse minimal quality. The reception of the subsequent enhancement layers increase the viewable quality by providing an increase in temporal, spatial or quality dimensionality. Thus stream quality adaption of layered coding is provided by means of layer selection. An example of layered coding is Scalable Video Coding (*SVC*) [15], an extension to the H.264/MPEG-4 Part 10 or AVC (Advanced Video Coding) compression standard. A major limitation to layered coding is that the technique implements a prioritised encoding hierarchy such that the increase in quality delivered by an enhancement layer is subject to the availability at the decoder of all lower layers that the enhancement layer is dependent upon [19]. In this manner, the loss of a lower layer prohibits the decoding of all higher enhancement layers. More seriously, the loss of the base layer invalidates the video decoding.

To compensate typical packet losses, Multiple Description Coding (*MDC*) [7] introduces redundancy to the transmitted video. On doing so, MDC creates a set of descriptions, each consisting of a group of *s*ections, which represent slices of SVC layers. Although MDC improves the perceived video quality, the introduced redundancy burdens the transmission medium by increasing the amount of data being transmitted over the network, i.e. increasing the *t*ransmission byte-cost.

In this paper, we propose Adaptive Layer Distribution (*ALD*), a framework for scalable video content delivery based on three different components:

- Section thinning (*ST*), by which ALD provides a significant reduction in transmission byte-cost in comparison to MDC,
- Improved error resiliency (*IER*), that enhances the user quality of experience, and
- Section packetisation, that targets reducing the impact of packet losses on the video quality. This technique is applicable to any description-based encoding, such as MDC.

ALD optimally determines the degree of section thinning, such that the total average streaming byte-cost is minimised. Our simulations show that ALD components result in a comparable quality to that achievable by MDC, but with a significant reduction in transmission byte-cost.

The remainder of the paper is organised as follows. Section 2 presents relevant background and the literature review. Section 3 presents ALD components and design, while Section 4 is dedicated to the comparative performance evaluation of ALD. Finally, Section 5 concludes the paper.

2. BACKGROUND AND RELATED WORK

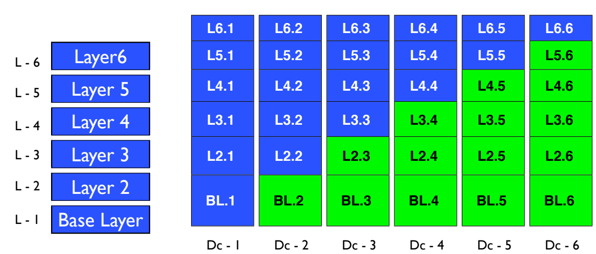
Generally, a video is identified as a scalable stream when an original high quality version of the video can be encoded into a set of sub-streams such that a combination of one or more of these sub-streams can be used to replay the video, at varying quality levels. The varying fidelity can be achieved over one or more dimensions including temporal, spatial, and quality scalability. The temporal scalability is determined through using different frame rates, spatial scalability is defined by changing the frame resolution, and quality scalability is achieved by changing the picture quality (e.g. pixel size). These scalable techniques can be used individually or in combination to create scalable videos. Typically, a number of subsequent frames are gathered to form a collection known as a Group Of Frames (*GOF*) (sometimes called a GOP, a Group Of Pictures). Each GOF is encoded using one or more of the aforementioned scalability approaches.

SVC employs one or more of these scalability schemes to encode a video stream into a set of layers, including a base layer and a set of enhancement layers as shown in Figure 1a. A critical issue in the SVC design is the interdependence between different layers, such that for viewing a specific quality, the corresponding topmost enhancement layers and *all* lower layers should be received. This constraint impacts the performance of SVC over lossy networks.

To overcome the prioritised hierarchy of SVC, Multiple Description Coding (*MDC*) is proposed as an error-resilient layered video coding with equal sub-streams importance. To achieve equal importance, MDC segregates the N SVC layers into M MDC descriptions, where the receipt of any description yields a minimal quality stream (i.e. base layer only), and the cumulative receipt of *a*ny number of descriptions incrementally increases the quality of the stream decoded at the device. Several variations of MDC have been proposed, including Sub-Sample, Quantisation, Transform and Forward Error Correction (FEC) [1] . Each implementation provides beneficial reasons for its use, yet only MDC-FEC supports an adaptive streaming encoding with dynamic error-resilient, low computational complexity, and flexible bandwidth adaptability for future heterogeneous devices [4].

MDC-FEC [12], henceforth in this paper referred to as MDC, utilises unequal error protection (*UEP*) [6] and layer partitioning to allocate equal importance to each layer in the original SVC stream data. UEP uses FEC proportional to the priority of the layer importance such that the lower the SVC layer, the greater the percentage of FEC. Layer partitioning provides a mechanism to separate the original layer data into an equal number of sections, corresponding to the original layer value, e.g. layer three is partitioned into three equal byte-sized sections. In this manner, MDC utilises layer partitioning to generate numerous sections per layer and FEC to extend these sections over a number of descriptions.

Typically FEC can provide either systematic or non-systematic encodings. Systematic schemes encode the original symbols as part of the transmitted stream, while non-systematic schemes encode and transmit the original symbols as new symbols. Raptor codes [16] proposes that a systematic encoding, with encoded symbols interspersed among the original symbols provides a greater level of decodability. To support the incremental increase of stream quality per additional description, N layers == M descriptions, Figure 1a shows an example of a six layer SVC scheme and it corresponding MDC encoding in Figure 1b. To highlight the comparability between original SVC data and FEC data, the original critical SVC data is shown in blue, dark shade, and the FEC sections are shown in green, light shade. It is important to note that in reality all layer sections are either a combination of FEC and original data, assuming a systematic encoding, or all FEC data, assuming a non-systematic encoding. As can be seen in the MDC-FEC image, the base layer features the highest percentage of FEC, while each of the enhancement layers is sub-divided into sections based on layer value and contains a proportionally smaller quantity of FEC. It is important to note that the increase in error resilience yielded by MDC, significantly increases its bandwidth requirements.



(a) SVC (b) MDC-FEC

Figure 1: An example of a 6 Layered SVC stream encoded as MDC-FEC

To this end, it is worth noting that datagram loss would impact the transmission unit of the encoding algorithm. In this context, a transmission unit refers to an SVC layer or an MDC description. That is to say that a loss of a datagram belonging to a specific SVC layer would invalidate the ability to decode such layer for the transmitted GOF even if other datagrams belonging to the same layer are received. Similarly, the decoder at the receiver would fail to reconstruct an MDC description if one of its datagrams is lost. In the case of using SVC, a loss of one or more datagrams from the GOF may result in a decrease in the play-out quality, or, in the worst case, prevent the client application from playing the video (loss occurs in base layer). In MDC, the loss of one or more datagram(s) from the GOF would never prevent the video play-out unless only one description is initially transmitted to the end user.

Table 1: Notation and Definitions

N	The number of SVC layers per Group of Frame
L_l	Byte-size of SVC Layer l
l	Integer value corresponding to the layer number of L_l
STF	Section Thinning Factor
D_c	A complete description, containing sections from layers 1 to N
q	Number of MDC descriptions required to decode Layer q
IER	Increased Error Resilience for a given layer

Several mechanisms have been proposed to reduce the transmission byte-cost of MDC while maintaining achievable quality from MDC description allocation. These include:

- adjusting the levels of FEC, such as Adaptive FEC [10];
- optimising FEC resilience, such as permitting FEC on a higher layer to be composed from SVC and FEC data from lower layers [8];
- modifying the layer allocation per MDC description, such as transmitting the base layer as a separate MDC description [3];
- modifying the base layer to create two individual descriptions [21], and
- optimising transmission cost by reducing the number of higher quality layers being transmitted [2].

In the following section, we present components of our ALD solution that reduce MDC overhead, improve error resiliency, and reduce the impact of packet losses.

3. ADAPTIVE LAYER DISTRIBUTION

In this section we introduce Adaptive Layer Distribution (ALD), a novel layered media technique that optimises the trade-off between streaming bandwidth demands and error resiliency. The proposal of ALD is motivated by two main objectives: reducing the transmission byte-cost overhead and maintaining a consistent play-out quality over lossy networks. In this context, play-out consistency refers to reducing the frequency of transitions in play-out quality due to packet losses. In order to realise these goals, ALD introduces the concepts of section thinning, improved error resiliency, and section-based application packetisation as detailed in the following subsections.

3.1 Section Thinning

3.1.1 Layer Section Allocation

As illustrated in Figure 1b the level of additional FEC data in MDC is proportionally high compared to the initial level of SVC data, thus leading to a large increase in transmission byte-cost relative to SVC. An MDC description section from layer l contains $\frac{1}{l}$ of the layer size, while a single MDC description, as shown in Equation (1), contains one section from layers 1 to N.

$$MDC\_D_c = \left(\sum_{l=1}^{N} \frac{L_l}{l}\right) \tag{1}$$

Thus the total transmission byte cost of MDC required to decode quality layer q can be seen as

$$MDC\_D(q) = MDC\_D_c * q \tag{2}$$

While the total FEC overhead for MDC quality layer q can be characterised as

$$MDC\_D(q) - \sum_{l=1}^{q} L_l \tag{3}$$

Note that layer l defines a specific layer within the encoding and transmission of SVC, while quality, or layer quality, q defines the viewable quality achievable by decoding a number of descriptions.

ALD section thinning is motivated to reduce the percentage of FEC data per layer, thus leading to a significant reduction in transmission cost for ALD in comparison to MDC. Section thinning reduces the byte-size of each layer section by increasing the number of ALD descriptions. The formation of the ALD sections follows the same footsteps of MDC section formation, but the section size in each scheme corresponds to a different share of the original SVC layer. In ALD, each section layer-share is scaled by an additional *section thinning factor* (STF) such that an ALD description section from layer l contains $\frac{L_l}{(l+STF)}$ of the layer size. Thus a single ALD description, as shown in Equation (4), contains one section from layers 1 to N, but each section byte-size is smaller. Thus leading to a smaller transmission byte-cost per ALD description.

$$ALD\_D_c = \left(\sum_{l=1}^{N} \frac{L_l}{(l+STF)}\right) \tag{4}$$

Thus the total transmission byte cost of ALD required to decode quality layer q is

$$ALD\_D(q) = ALD\_D_c * (q + STF) \tag{5}$$

While the total FEC overhead for ALD quality layer q can be characterised as

$$ALD\_D(q) - \sum_{l=1}^{q} L_l \tag{6}$$

Thus, if STF > 0, the transmission cost of an ALD description is less than the cost of an MDC description (because ALD contains less FEC data), but more ALD descriptions are required to decode the same quality layer q.

It is important to note that ALD with an STF value of zero equates to the same layer section byte allocation, number of descriptions and transmission byte-cost as MDC. Thus ALD with an STF value equal to zero is exactly MDC. Figure 2 illustrates the representation of the six-layer SVC video from Figure 1a, using ALD with an STF value equal to three. As shown in the figure, each layer is further extended over the three additional descriptions in comparison to the original MDC.

There are a number of points to note when you compare MDC, Figure 1b, and ALD, Figure 2:

1. As previously highlighted, each MDC description is capable of providing base layer quality, thus mandating MDC to allocate the entire SVC base layer to each MDC description.

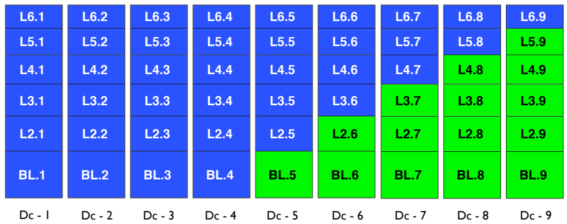


Figure 2: ALD GOF for six-layers, with STF = 3

As illustrated in Figure 1b, BL-1 from Dc-1 is the original (blue) SVC base layer, while BL-2 to BL-6, inclusive, are the additional FEC base layer sections. Thus, in this example, leading to six base layer sections being transmitted, or 600% of the original SVC base layer transmission cost.

While in Figure 2, by utilising STF, it can be seen that the original blue (dark) SVC base layer data is distributed over more ALD descriptions, BL-1 to BL-4 in our example, consequently reducing the byte cost of each ALD description base layer section to 25% of the original SVC base layer. Thus, in this example, leading to a transmission byte cost of 225% of the original SVC base layer transmission cost.

Once this mechanism for section thinning is applied to each layer in the transmitted stream, the transmission byte-cost of ALD is less than MDC. It can be seen that the original blue (dark) SVC data for each layer is shared over more ALD descriptions than MDC descriptions (excluding the highest layer in both schemes where no FEC occurs), thus leading to a reduction in transmission byte-cost, irrespective of encoding rate.

2. The number of FEC sections per layer is consistent between MDC and ALD, but the FEC section byte-allocation in ALD is smaller.

3. A greater number of ALD description, three from the Figure, are required before base layer decoding is achievable. For a device that only needs to view at low-quality this has implications in terms of having to receive more descriptions than with MDC. This is discussed in the next section.

So clearly, the optimal choice of STF is an important design issue that will be introduced later in this paper.

3.1.2 Quality Transmission Cost

Generally, multiple users may be interested in viewing the same video at different qualities, depending on several factors such as the available resources and device capabilities. Section thinning realises significant savings for users interested in receiving high quality video. On the contrary, if a user is interested in receiving low video quality, ALD may result in a larger overhead in comparison to MDC, as additional (STF) ALD-descriptions have to be received in order to decode the base layer. As previously defined, only q MDC descriptions, Equation (2), are required to decode quality layer q in comparison to $(q + STF)$ ALD descriptions, Equation (5), to realise the same video quality.

Hence, the difference in the amount of transmitted data, or total relative overhead $D(q)$, for a single user between ALD and MDC for video quality q, can be calculated as

$$D(q) = ALD\_D(q) - MDC\_D(q) \tag{7}$$

Note that negative total overhead implies that ALD is more bandwidth efficient than MDC for the selected quality level q. Future work will investigate mechanisms to reduce the transmission byte-cost increase for lower layer streaming.

3.1.3 Optimal STF Selection

As previously mentioned, multiple users may be interested in viewing the same video at different qualities, thus ALD provides a mechanism for optimal STF in streaming scenarios for both unicast, single user with one quality requirement, and multicast, numerous users with possibly differing requirements. Multicast provides two options for ALD transmission: i) Each quality layer q is transmitted as a separate entity, thus implementing a multi-bitrate scheme (this option overly increases transmission cost) ii) each ALD description is transmitted as a single multicast stream, thus allowing users to subscribe to (q + STF) descriptions to receive the required q quality layer (this option reduces transmission cost, as only the maximum requested quality layer, (max[q] + STF) descriptions, are transmitted thus permitting multiple users access the same descriptions for their respective q' quality layers).

Let p_q denote the percentage of clients interested in viewing a video with quality level q. In a unicast scenario, this would be based on the requirements of a single user, while in multicast, would consider the needs of numerous users and their varied demands. Thus, the expected total overhead can be estimated as

$$E\{D(q)\} = \sum_{q=1}^{N} p_q D(q). \tag{8}$$

In our design, we choose an $\overline{STF_O}$ value that minimises the expected total overhead and can be expressed as

$$\overline{STF_O} = \arg\min_{STF} E\{D(q)\} \tag{9}$$

Note that the optimal $\overline{STF_O}$ would vary depending on different factors including the number of layers and the size of each layer. To this end, it is worth noting that other strategies can be used to determine the STF value.

3.2 Improved Error Resiliency (IER)

The main objective of the IER component is to enhance the streaming quality by ensuring a smooth play-out with fewer quality transitions. Clearly, the FEC overhead of higher layers in MDC is inversely proportional to the layer-level. For the top-most layer, no FEC is considered. Hence, the loss of any MDC description results in an immediate downgrading of the stream quality. Similarly, further proportional reductions in the stream quality for the same GOF is dependent on the cumulative loss of additional descriptions.

IER reduces the number of non-redundant sections of higher layers by distributing the higher layer data over one less section allowing for an additional FEC section; i.e. the layer share per an ALD description would increase from $1/(l + STF)$ to $1/(l + STF - ER)$, where ER represents the added IER factor. IER may be applied to any number of the highest layers at the expense of additional FEC overhead for the selected layers. However, it is typically applied to the few top-most layers to reduce the incurred FEC overhead. Thus IER-1 can be viewed as providing one additional section of FEC to the highest layer, while IER-2 provides one additional section to the top two highest layers, such that in general IER-n allocates one additional section to the n highest layers.

3.3 Section Packetisation

This component reduces the impact of packet loss on any description-based scalable video, such as MDC and ALD. The application transmission unit for MDC is its description. For purpose of illustration, we use a single GOF example from the widely-used video clip known as crew.yuv, encoded as a six-layer SVC stream. Table 2 shows the byte-size of each layer for the selected frame.

Table 2: GOF SVC Layer sizes

Layer	1	2	3	4	5	6
Layer Size	1442	1577	1601	1546	1255	3372

In today's Internet, the maximum packet size observed is usually limited by that of the Ethernet frame, which has a maximum payload of 1,500 bytes. We assume a packet payload of 1,440 bytes, allowing for overhead due to headers of network, transport and streaming media protocols. We assume that the GOF frames are transmitted over Ethernet packets. On transmitting this frame, eleven Ethernet packets would be required using SVC where the transmission unit is an individual layer. The same frame would require eighteen packets when encoded using MDC in which the description represents the application transmission unit. On losing any of these datagrams, the application would not be able to decode the *entire* frame to the highest quality. In order to reduce the impact of losses on the stream quality, we propose to use two packetisation mechanisms, called section-based description packetisation and section distribution.

3.3.1 *Section-Based Description Packetisation*

With section-based description packetisation (SDP), we propose using sections as application transmission units instead of the entire description for description-based layered coding techniques such as MDC and ALD. As a consequence the description is composed of a number of sections, with each description section transmitted as a single unit, thus limiting the affects of packet loss to individual section while allowing partial description re-use. Partial description re-use in this instance means the availability at the device of one or more layer sections from a single description. The probability of loss affecting all sections from a single description, or all sections from a single layer, is low, while the probability of partial description re-use is high.

SDP improves the possibility of higher stream quality by mitigating lower layer loss thus increasing the availability of a sufficient number of lower layer sections.

SDP can be applied in several ways as follows:

- *Option 1 - Individual layer sections* - this option transmits each layer section as a separate group of one or more datagrams. This option may increase the number of datagrams being transmitted, depending on the original encoding but maximises the number of sections available during decoding. Using the example frame, it can be seen that for each MDC description six datagrams are required for transmission as shown in Figure 3. This option increases the number of datagrams and in some instances creates datagrams not containing a full data payload. Consequently the overhead due to packet headers and processing is higher.

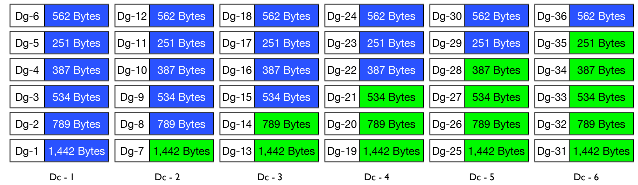


Figure 3: MDC-SDP Option 1 - with six descriptions (D_c) consisting of six datagrams (D_g)

- *Option 2 - Minimising datagrams quantity* - this option groups layer sections together to fully occupy each transmitted datagram, thus mitigating the problems with Option 1. Figure 4 illustrates this option for the example frame. This option reduces the number of transmitted datagrams. However, the loss of a datagram may cause the loss of numerous layer sections. To reduce the probability of stream degradation due to datagram loss, only one section for each layer should be included within a datagram. If a datagram were to contain numerous sections for one specific layer, then the loss of that specific datagram may aversely affect the decoding of that layer and all enhanced layers that rely upon it. It can be seen that for each description, D_c, three datagrams, D_g, are required for transmission.

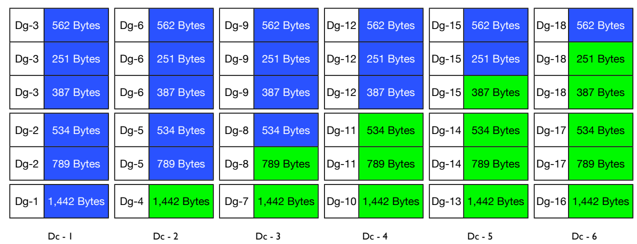


Figure 4: MDC-SDP Option 2 - with six descriptions (D_c) consisting of three datagrams (D_g)

It is worth noting that the blue (dark) sections are the critical SVC data and the green (light) sections the FEC section allocation. It can be seen that in Figure 3 and 4 that the base layer consumes a single datagram, D_g-1 in description one; in Figure 3 each section is allocated to an individual datagram while in Figure 4, a section from layer two and three are allocated to D_g-2 and a section from layer four, five and six are allocated to D_g-3. A total of thirty six datagrams are transmitted over the network with Option 1 in which only twenty one specific datagrams are required for maximum stream quality. In Option 2, eighteen datagrams are transmitted among which only ten specific datagrams are required for maximum stream quality. Thus Option 1 increases the probability of maximising stream quality in the presence of high levels of datagram loss.

To this end, we also define an $\overline{STF_E}$ value that maintains a level of error resilience per ALD description and can be expressed as a lower bound on the number of datagrams per description. For example, $\overline{STF_E}$ can be chosen such that a minimum of two datagrams should be used to packetise each ALD description. Hence, the loss of one of these packets would not completely affect an entire description. This approach would sustain high levels of video quality. Hence, the chosen $\overline{STF}$ value can be defined as

$$\overline{STF} = \min\{\overline{STF_O}, \overline{STF_E}\} \tag{10}$$

3.3.2 Section Distribution

Due to the transient nature of the Internet, network traffic can be affected by both *individual* and *burst loss* states corresponding to a single datagram loss or numerous consecutive datagram losses. Packet losses at lower layers have a negative impact on scalable video due to inter-layer dependency. As shown above, by manipulating the stream packetisation, we can increase stream quality and consistency. With this in mind, we propose Section Distribution (SD), a mechanism to further distribute the description sections over the datagrams used to transmit a ALD description to further reduce the impact of losing critical sections. SD is utilised to distribute each ALD section per description over a number of datagrams, thus limiting packet loss to only a segment of each ALD section. SD is consistent with the well-known Interleaving [6] technique, which is widely used to combat the effect of burst loss.

We first determine the number of datagrams, denoted as R, required to transmit an ALD description. Each ALD section, denoted as S_j, is spread over the R datagrams. Hence, an ALD datagram would carry subsections from different layers. In this manner, the loss of an individual datagram, will result in a partial loss from each layer. Thus the quality of the packetised stream is limited only by the percentage of lost packets rather than the specific carried description or layer. Additionally, the probability of losing critical sections is reduced since lower layers enjoy greater redundancy.

Furthermore, on using section distribution, ALD datagrams per frame would be identical in size, thus providing datagram equality. Also should a larger GOF rate be utilised, then SD will provide data equality for all frames within the GOF. In [11], the authors highlight that datagrams of dissimilar processing times, produce dissimilar transmission times. Such that by maintaining such datagram byte-size equality, the order of packet delivery is improved. Thus SD datagram equality results in a consistent delivery in network transmission.

The SDP and SD claims made in this section are generally applicable to videos that have different number of SVC layers, as highlighted later in the evaluation section, where the chosen layer size has been increased to eight.

3.3.3 Transmission Unit Stream Quality Loss Rate

In Table 3, we show the transmission cost, in terms of bytes, for SVC, MDC, both options of MDC-SDP and by utilising Eq (10) an ALD with an STF of 3. Thus increasing the number of ALD descriptions to nine. Table 3 also presents the number of datagrams per frame and highlights the best case (B/C) and worst case (W/C) maximum viewable layer based on the loss of a specific number of datagrams. It is worth noting that the SVC data transmission byte-cost for all versions of MDC are equal, but the total transmission byte-cost for each scheme will vary, dependent on the number of datagrams being transmitted and the increased byte-cost of datagram headers. In this section to provide a simplified example, we evaluate the SVC data element only.

As the number of lost datagrams increases, and dependent on which datagram is lost, the quality of the stream can remain high or degrade significantly. As can be seen, SVC is severely affected by datagram loss. The worst case (W/C) for all four lost datagrams, highlights the loss of a datagram from the base layer, while the best case (B/C) is based on consecutive losses from the highest quality layer down, i.e. layer six is composed of three datagrams, such that B/C will remain at quality level layer 5 until the fourth datagram is lost, when the quality reduces to quality layer four.

MDC-FEC is similar in that each description is composed of three datagrams, such that the B/C remains consistent over three datagram losses, and reduces quality to layer four when the fourth datagram is lost. W/C is based on the loss of a single datagram from distinct descriptions, thus incrementally reducing quality for each additional datagram lost. The increase in viewable quality is consistent with the level of additional error resilience added to the original SVC data, but this increase in viewable quality requires an additional approximately 13,000 bytes of transmission bandwidth.

Table 3: Example transmission byte-costs for SVC, MDC, MDC-SDP (both options) with viewable quality as datagram loss increases

Scheme	SVC	MDC-FEC	MDC-SDP-1	MDC-SDP-2	ALD
Transmission Cost	10,793	23,790	23,790	23,790	15,273
# of Datagrams	11	18	36	18	18
One Lost Dg	B/C = 5 W/C = 0	B/C = 5 W/C = 5	B/C = 6 W/C = 5	B/C = 6 W/C = 5	B/C = 5 W/C = 5
Two Lost Dg	B/C = 5 W/C = 0	B/C = 5 W/C = 4	B/C = 6 W/C = 4	B/C = 6 W/C = 4	B/C = 5 W/C = 5
Three Lost Dg	B/C = 5 W/C = 0	B/C = 5 W/C = 3	B/C = 6 W/C = 3	B/C = 6 W/C = 3	B/C = 4 W/C = 4
Four Lost Dg	B/C = 4 W/C = 0	B/C = 4 W/C = 2	B/C = 6 W/C = 2	B/C = 6 W/C = 2	B/C = 4 W/C = 4
...	...	...	...	...	...
Six Lost Dg	B/C = 3 W/C = 0	B/C = 4 W/C = 0	B/C = 6 W/C = 0	B/C = 6 W/C = 0	B/C = 3 W/C = 3

Consistent with MDC-FEC, both options of MDC-SDP achieving the same W/C viewable quality, again based on a single lost datagram from distinct descriptions. Both options of MDC-SDP achieve the maximum B/C over all four lost datagrams, as loss can be confined to the green FEC section datagrams. Thus highlighting the benefits offered by section based description packetisation.

As previously stated, ALD employs the section distribution (*SD*) technique for datagram packetisation, thus achieving datagram equality. As highlighted in the Table 3, this equality produces a uniformity in the B/C and W/C achieved by ALD. As each of the nine ALD descriptions is composed of two datagrams, achievable viewable quality is incrementally reduced once two additional datagrams are lost.

A loss rate of six datagram is illustrated to highlight that with the loss of six datagrams, the transmission cost of ALD over the network is less than the transmission cost of SVC with no datagram loss. This offers a comparison of the B/C and W/C quality achieved by SVC and ALD for similar transmission byte-cost. It is important to note that while the B/C of ALD is less than SVC, the W/C of ALD is better, thus highlighting the balance offered by ALD between transmission cost and achievable consistent quality.

Note that by implementing the previously highlighted IER technique on the highest layer, layer six. The viewable quality layer value for both B/C and W/C achieved by ALD-IER for the loss of one or two datagrams is six. Thus maximum quality can be achieved for a very minor increase in transmission byte-cost, 47 bytes per ALD description.

4. PERFORMANCE EVALUATION

In this section, we assess the performance of ALD with respect to SVC, MDC and MDC-SDP (option 1), furthermore just referred to as MDC-SDP. Transmission is simulated over a lossy network and the maximum stream quality level achievable at the device is measured. The goal of our evaluation is to determine how viewable video quality, for each streaming scheme, is affected by defined levels of datagram loss. Thus we do not implement a retransmission mechanism in our evaluation, as in any case to do so would add unwanted delay and network overhead. We begin by outlining the evaluation framework utilised to encode the media, simulate datagram loss, and generate the results.

4.1 Evaluation Framework

4.1.1 Media Encoding

As efficient encoding is the initial step in providing a quantitive evaluation, Joint Scalable Video Model (*JSVM*) [13] software is utilised for our encoding requirements.

We use eight distinct quality layers for each simulated video. Additionally we use a GOF value of one to focus on a one dimensional overview of how loss affects scalable media. This value eliminates the interdependence of I, P and B frames. In this manner, the maximum quality of a frame is dependent only upon itself, and is not affected by the quality of neighbouring frames. The effects of a larger GOF size would limit the achievable quality of a frame to the maximum quality of its dependent frame.

In this regard, JSVM v9.19, based on eight-layer and one frame per GOF, is utilised to create a multi-layered SVC-compliant stream. This stream is encoded from a widely used raw ten-second 4CIF 30fps media clip, crew.yuv, which consists of a number of astronauts walking down a corridor while waving. All well-known YUV clips utilised in the evaluation section have a total viewing time of ten seconds and were obtained from the well known Leibniz Universität Hannover video library [20].

Table 4: SVC structure, defined by JSVM, for an eight-layer, 30 fps, scheme with a GOF of 1 frame

Layer	Resolution	Bitrate	DTQ
0	176x144	345.60	(0,0,0)
1	176x144	826.80	(0,0,1)
2	352x288	1381.70	(1,0,0)
3	352x288	2069.00	(1,0,1)
4	352x288	2755.00	(1,0,2)
5	704x576	4293.00	(2,0,0)
6	704x576	5345.00	(2,0,1)
7	704x576	6796.00	(2,0,2)

Table 4, highlights the JSVM SVC structure command line output for the eight-layer, single frame GOF, encoded stream. The only heading that may not be immediately clear is the JSVM defined **DTQ**, which equates to three degrees of scalability: resolution (**D**ependency ID), frame (**T**emporal level) and PSNR (**Q**uality level). This example JSVM encoding yielded a maximum PSNR of 38.52, a maximum cumulative bitrate of 6,788 bytes and an encoding time of approximately 13.89 seconds.

JSVM contains a mechanism for extracting actual trace data from the encoded SVC stream, which provides a means for us to determine the transmission cost and datagram requirements per frame, for SVC, MDC, MDC-SDP and ALD. Table 5 utilises the trace data and highlights the transmission cost, average size of each layer or description, number of layers/description per-frame, the number of datagrams required and percentage byte size with reference to SVC for each of the adaptive schemes. This trace data is utilised by ns-2 in the simulation section.

By applying Eq (10), based on the transmission costs of SVC in Table 5, an STF value of 6 can be determined. The STF value specifies that six additional descriptions are required, thus increasing the number of ALD descriptions to fourteen. Note that the transmission byte-cost of MDC compared with SVC is 197%, while the ALD transmission byte-cost compared to SVC is 127%, a reduction of approximately 36% when ALD is compared with MDC.

Table 5: Transmission byte-cost as per the eight layer JSVM streaming trace data

Scheme	SVC	MDC	MDC-SDP	ALD
Transmission (bytes)	8,487,766	16,681,768	16,681,768	10,760,302
Layer or Description $\overline{Size}$	3,536	6,950	6,950	2,561
Transmission Units (TU)	Layer	Description	Description	Description
TU Quantity	8	8	8	14
# Datagrams	6,993	12,456	20,728	9,380
Value compared to SVC	100 %	$\approx$ 197 %	$\approx$ 197 %	$\approx$ 127%

Note also that the number of datagrams for MDC-SDP has increased to 20,728 while the transmission byte-cost (in terms of payload) has remained the same.

4.1.2 Simulation

Network Simulator 2 (*ns-2*) [18] is utilised to simulate a lossy channel between a streaming server and client. For this setup we simulate a packet erasure channel, whose packet error rate varies from 1% to 10%. In this simulation, UDP is the transport protocol. Selected 10-second videos at 30 frame per second (*fps*) are transmitted over the aforementioned channel.

myEvalSVC [9], an open source tool for evaluating JSVM stream trace data in ns-2, presents a means of dynamically determining bitrates, based on the JSVM trace data, and simulating real-time packetisation, over a lossy network, in ns-2. In our evaluation myEvalVid is utilised to create streaming packetisation models for each of the scalable schemes, namely SVC, MDC, MDC-SDP and ALD. As JSVM encodings are specific in both scalability (revolution, frame, quality) and quantity of layers, minor modifications are made to the original myEvalSVC scripts, so as to operate with our encoding of an eight layer streaming model.

Due to the multi-datagram requirements of each of the streaming schemes and possible out of order delivery of datagrams at the device, myEvalVid provides mechanisms so as to calculate the maximum achievable stream quality at the device. Comprehensive simulations are performed and once complete, the myEvalVid trace files are analysed to determine the maximum, per-frame, stream quality at the client. This trace is then saved to an achievable quality (*AQ*) trace file, for each streaming scheme.

4.1.3 Result Generation

To generate results for each of the adaptive schemes, PSNR [14], a widely used pixel quality differentiation mechanism, is utilised to correlate the values in the AQ trace files to the quality of the original YUV media clip. To determine the effects of loss on the quality, the AQ trace file values are first converted to a YUV file. myEvalVid and JSVM do not contain a reliable mechanism for this form of YUV modification, so a new program, modPSNR.exe, is created based on the original JSVM source code.

A by-product of the initial JSVM SVC encoding is the creation of individual YUV files for each quality layer of the original SVC stream. As PSNR calculations require that the resolution of evaluated streams are equal, JSVM is utilised to up-sample the lower resolution streams using a set of integer-based 4-taps filters derived from the Lanczos-3 filter. modPSNR utilises these adjusted YUV files to create a modified YUV file, based on the AQ trace file, which is consistent to the level of ns-2 simulated loss. modPSNR also supports basic error concealment by which non decodable frames are substituted by duplicating the previous frame. As only SVC and MDC are affected by non-decodable frames, more advanced error concealment mechanisms would unfairly increase the PSNR value of these schemes.

Finally JSVM is utilised to ascertain the PSNR value of the modified YUV, in comparison to the original YUV file.

4.2 Simulation Results

4.2.1 Effects of Section Thinning Value Selection

We shall begin by highlighting how the choice of STF affects both the transmission cost of each quality layer and maximum achievable PSNR for the highest quality layer.

Transmission Cost:

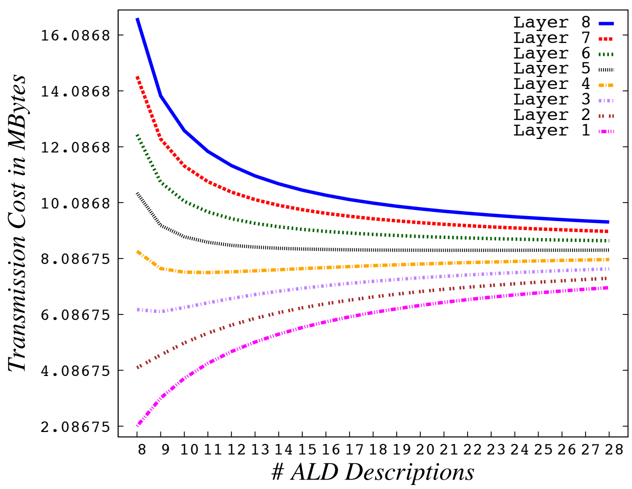


Figure 5: Transmission cost of the crew media clip for each distinct media quality, as the STF value and associated number of ALD descriptions increase

Figure 5 illustrates the transmission cost of each of the different achievable video qualities, as the STF value increases. It can be seen that as the number of descriptions increases, the transmission cost of streaming the higher quality videos decreases, while the transmission cost of streaming lower quality videos increases. The increase in the number of ALD descriptions affects the transmission byte-cost of both the lower quality and higher quality videos in opposite manners. As mention in Section 3.1.1, ALD with an STF of zero equates to MDC. Thus in Figure 5, ALD with a description number of 8, illustrates the per layer transmission cost of MDC.

As the value of STF increases, more descriptions are required to decode the base and lower layer quality sub-streams, thus more bandwidth is consumed for the lower quality video. The opposite is true for the higher quality layers, as STF value increases, the bandwidth cost is reduced. As the value of STF increases, this mandates a modification in the transmission byte-cost of the stream. In a scenario where all quality layers are transmitted and fidelity is determined at the device by the number of descriptions required to decode a specific quality layer, an increase in STF delivers an overall reduction in transmission bandwidth costs, due to a reduction in the number of bytes per section, as well as a decrease in the overall percentage of FEC per layer. Thus Figure 5 can be viewed as means of highlighting the transmission trade off between the different quality videos for a number of STF values, namely 0 to 20.

Also the increase in transmission cost of the base quality (minimum quality video) is slower than the drop of the transmission cost of the maximum quality. It is important to note that in description-based streaming scheme (such as ALD and MDC), sections of other layers are transmitted with the layer that is requested. Such that in the case of the base layer quality, not only is the base layer section transmitted, but also sections from all other layers.

Achievable PSNR:

Figure 6 illustrates the adjustment in achievable PSNR, for the highest quality video, as the value of STF changes from six to five. This adjustment increases the number of ALD descriptions to fourteen. It can be seen that there is a direct correlation between the reduction in transmission byte-cost and maximum achievable PSNR as STF increases.

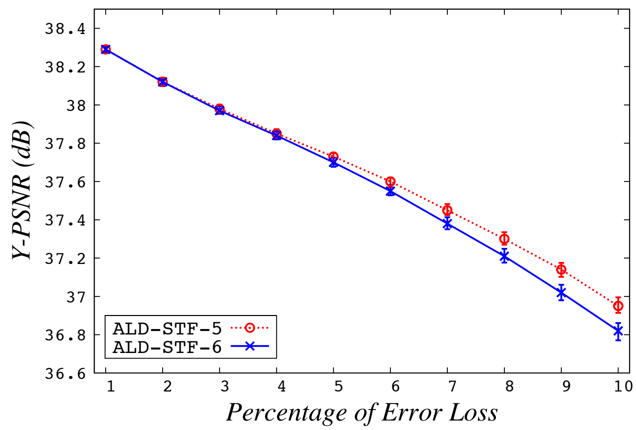


Figure 6: Mean Y-PSNR values and 95% Confidence Interval error bars with incremental datagram loss for crew with a revised STF value of 6 and the initial value of 5

The adjustment in STF increases the transmission byte-cost by approximately 0.3MB, while also increasing PSNR by approximately 0.1dB at 10% datagram loss. Thus STF provides a mechanism for determining the optimal balance between transmission byte-cost and achievable quality.

It is important to note that the highest levels of transmission byte-cost reduction are achieved with low values of STF, as shown in Figure 5. Once a slow gradation in transmission cost is achieved, as occurs at ALD description number fifteen in Figure 5, further transmission cost reduction supports little additional benefit to achievable stream quality.

4.2.2 Impact of Section Thinning

To better understand the degree of variation that can occur in adaptive media streaming, Figure 7 shows a two-second snapshot of the ns-2 simulation and the viewable quality transitions that were analysed for SVC, MDC, MDC-SDP and ALD with a datagram loss rate of 10%.

It can be seen that SVC and MDC feature the highest frequency of variation and as such would provide a media stream with frequent variation in video quality. MDC-SDP and ALD also contain

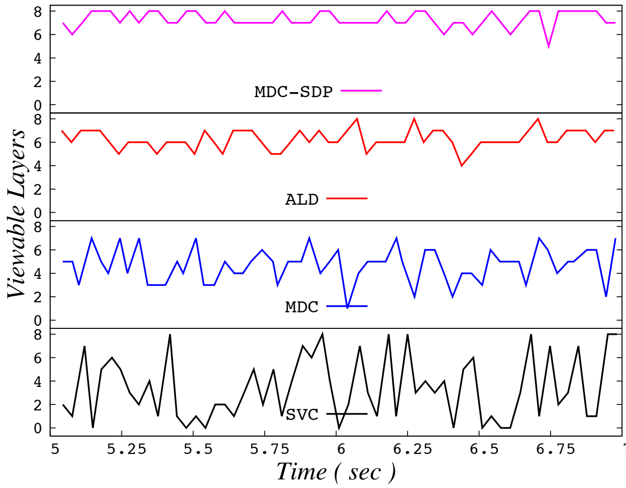


Figure 7: Two second example of viewable quality transition for eight-layer SVC, MDC, MDC-SDP and ALD with 10% datagram loss.

less variation. More importantly these variations are limited to the higher quality layers. MDC-SDP has the least amount of variation, with a minimum reduction in stream quality to layer five. ALD does consist of a slight increase in the level of variation, but with a predominantly minimum reduction in stream quality also to layer five. The impact of these fluctuations are reflected in the PSNR values as shown in Figure 8.

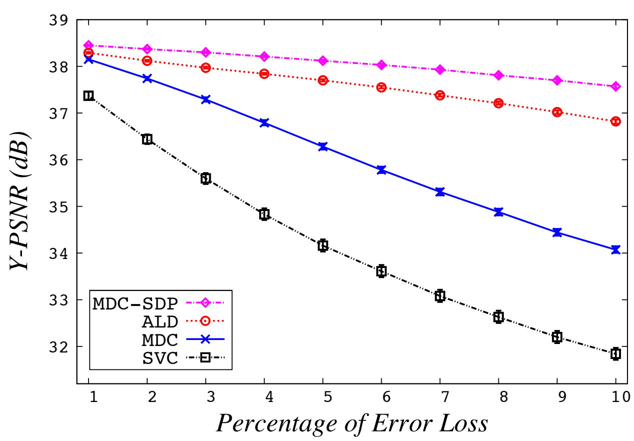


Figure 8: Mean Y-PSNR values and 95% Confidence Interval error bars with incremental datagram loss

Figure 8 plots the Y-PSNR values for the simulated streaming schemes versus the percentage of datagram loss. The values are consistent with Figure 7, as MDC-SDP performs best, ALD produces a slight reduction in quality, with MDC and SVC having the worst quality. It is important to note that while the PSNR quality of the ALD stream is slightly worse than MDC-SDP, approximately 0.75dB or 2%, the transmission overhead of ALD is approximately 36% less than MDC.

4.2.3 Improved Error Resilience

In this section, we investigate the impact of IER on the performance of ALD, as was suggested in section 3.2. Generally, the top most layer in a description based scheme is distributed over the total number of descriptions, thus providing no FEC error correction for the highest quality layer. This can lead to degradation of the maximum steam quality, unless all descriptions for a given frame are received. Figure 7 illustrates that, similar to MDC-SDP, the viewable quality level of ALD continuously moves between the higher quality layers, which can lead to noticeable video variation for a viewer. To increase the level of consistency in the media stream, we propose to increase the level of error resilience for the higher quality layers. In this manner, we distribute the highest layer data over a reduced number of ALD descriptions, thus we marginally increase the byte-size of each higher layer section.

Let us use layer eight, as an example. Currently layer eight is distributed over each of the fourteen descriptions. To facilitate an additional section, without increasing the number of description, we take the existing byte size of layer eight and distribute it over one description less, i.e. distribute over thirteen descriptions. Thus increasing the byte-size of layer eight per description section. The remaining section, description fourteen, is then populated with layer eight FEC error resilience, which is consistent with the objective of ALD. Generically, an incremental increase in IER extracts the layer byte size currently divided by N descriptions, decrements N to N-1 descriptions and add one section of FEC error resilience to the N^{th} description.

In this manner, we increase the resilience of the higher quality layers, thus increasing stream quality consistency, with a minor increase in bandwidth transmission. Table 6 highlights the increase in transmission byte-cost, number of datagrams per ALD (IER) scheme and byte-size when compared with SVC. Notice the slight increase in transmission byte-cost for the IER schemes. This is due to the highest layers being distributed over the greatest number of descriptions and containing the least amount of error resiliency.

Table 6: Increasing transmission byte-cost for ER, with reference to Table 5

Scheme	ALD	ALD-IER-1	ALD-IER-2
Transmission (bytes)	10,760,302	10,899,756	11,017,664
Layer or Description $\overline{Size}$	2,561	2,595	2,623
# Datagrams	9,380	9,492	9,562
Value compared to SVC	≈ 127%	≈ 128 %	≈ 130%

Figure 9 presents the variation in stream quality transitions, as IER is first increased on layer eight (ALD-IER-1) and then increased on both layer eight and layer seven (ALD-IER-2). Note that while the consistency of the stream quality at ALD-IER-2 is not continuous, there is a noticeable increase in view-ability for the higher quality layers, which is compatible to the MDC-SDP stream replicated. Note that such improvement is achieved at a minor increase in transmission byte-cost as shown in Table 6.

Figure 10 shows the increase in PSNR for ALD-IER-1 and ALD-IER-2, with reference to ALD and MDC-SDP with up to 10% datagram loss. As can be seen, as the percentage of loss increases, the PSNR curve for both IER-1 and IER-2 will converge with ALD due to the reduction in benefits yielded by only one additional section at the highest layer(s). Future works will determine the optimal number of additional sections required by IER to maintain quality consistency as loss increases.

Table 7 gives an overview of the transmission byte-cost for each of the streaming schemes evaluated. It can be seen, that the SVC transmission byte-costs per layer are lowest. ALD delivers a dramatic reduction in transmission byte-cost for higher quality layers when compared with MDC, but as mentioned in Section 3.1, ALD

Table 7: Example transmission byte-cost of the crew media clip for each quality layer, for each adaptive scheme. Number of layers or descriptions required to received the respective quality is shown in brackets. Transmission costs of MDC and MDC-SDP are equal, with only MDC being shown.

Layer	SVC	MDC	ALD	ALD-IER-1	ALD-IER-2
8	8,487,766 (8)	16,681,768 (8)	10,760,302 (14)	10,899,756 (14)	11,017,664 (14)
7	6,674,789 (7)	14,596,547 (7)	9,991,709 (13)	10,121,202 (13)	10,230,688 (13)
6	5,360,940 (6)	12,511,326 (6)	9,223,116 (12)	9,342,648 (12)	9,443,712 (12)
5	3,440,447 (5)	10,426,105 (5)	8,454,523 (11)	8,564,094 (11)	8,656,736 (11)
4	2,584,949 (4)	8,340,884 (4)	7,685,930 (10)	7,785,540 (10)	7,869,760 (10)
3	1,728,843 (3)	6,255,663 (3)	6,917,337 (9)	7,006,986 (9)	7,082,784 (9)
2	1,036,455 (2)	4,170,442 (2)	6,148,744 (8)	6,228,432 (8)	6,295,808 (8)
BL	434,947 (1)	2,085,221 (1)	5,380,151 (7)	5,449,878 (7)	5,508,832 (7)

Table 8: Example of the percentage of decodable frames per quality level for each of the adaptive schemes, based on the crew media clip with 10% packet loss

Layer	SVC	MDC	MDC-SDP	ALD	ALD-IER-1	ALD-IER-2
8	10.67%	1.67%	34.00%	6.33%	23.00%	21.67%
7	6.67%	10.67%	46.33%	40.67%	26.33%	48.00%
6	4.00%	17.33%	15.00%	40.00%	39.00%	18.67%
5	14.00%	26.33%	3.67%	11.33%	10.33%	10.66%
4	9.67%	25.67%	1.00%	1.67%	1.34%	1.00%
3	12.00A%	13.00%	0.00%	0.00%	0.00%	0.00%
2	9.67%	4.33%	0.00%	0.00%	0.00%	0.00%
BL	20.00%	1.00%	0.00%	0.00%	0.00%	0.00%
Un-viewable	13.33%	0.00%	0.00%	0.00%	0.00%	0.00%

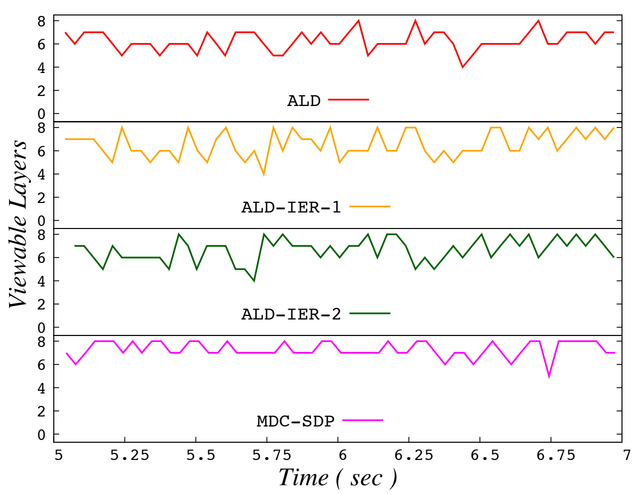


Figure 9: Two second example of viewable quality transition for ALD and MDC-SDP, from Figure 7, and ALD-IER-1 (layer eight) and ALD-IER-2 (layer eight and layer seven)

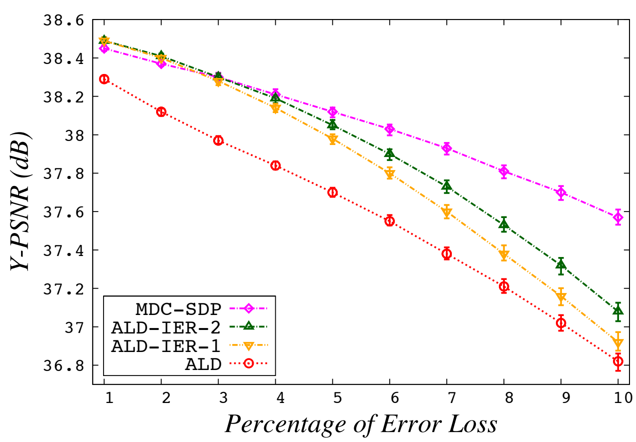
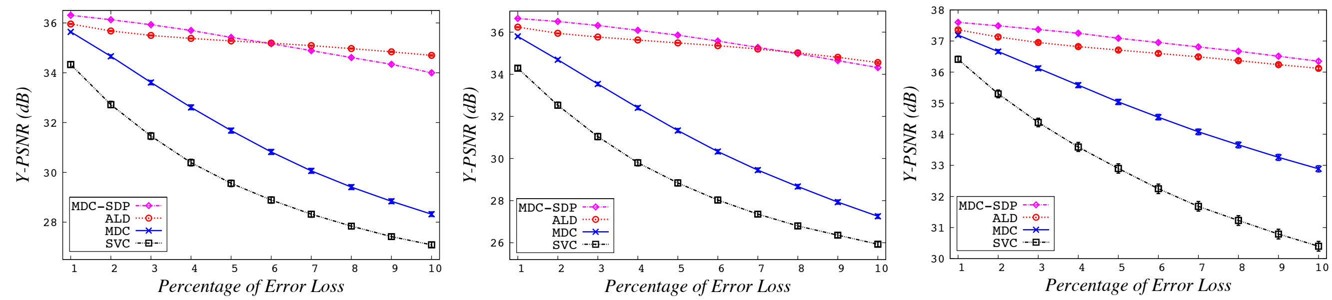


Figure 10: The PSNR increase in ALD with IER on layers eight and seven

has a higher transmission cost, than MDC, for users requiring lower layer quality. These results are consistent with Figure 5.

Table 8 outlines an example of the percentage of decodable frames per quality level for each of the adaptive schemes. It can be seen that MDC-SDP consists of a large percentage of decodable frames for quality layers six to eight, thus providing a consistent high level of viewable quality. ALD also contains large quantities of layers six and seven but the inclusion of the low bandwidth cost IER, re-allocates these high percentage levels over layers six to eight. Thus highlighting the increased benefits provided by IER.



(a) City stream, $\overline{STF}$ 2 (b) Harbour stream, $\overline{STF}$ 5 (c) Soccer stream, $\overline{STF}$ 3

Figure 11: PSNR results for additional stream evaluation. Note that for clarity, the plots use different scales on the Y-axes

4.3 Additional Evaluation

Three additional streams were evaluated using the same JSVM SVC encoding as the Crew stream and their evaluation results are given in this section. The streams are City, a slow moving aerial shoot of a city, Harbour, a stationary shot of a harbour with a number of moving boats, and Soccer, a fast moving shot of a game of soccer. All additional clips are available from the Hannover repository [20] and a single frame from all four evaluated streams is given in Figure 12. By utilising Eq (10) and based on the SVC per layer transmission byte-costs of each additional media clip as presented in Table 9, the $\overline{STF}$ values for city is 2, the $\overline{STF}$ value for harbour is 5 and the $\overline{STF}$ value for soccer is 3.

Figure 12: Single frame image from each of the four evaluated streams [20]: crew, city, harbour and soccer

Table 10 highlights the relevant maximum PSNR and transmission byte-cost per adaptive scheme for each of the additional streams; note the varying percentage increase in transmission cost for MDC and ALD.

Figure 11 outlines the achieved PSNR for each of the evaluated adaptive schemes per media clip. It is important to note that while the maximum PSNR and transmission byte-cost for each of the adaptive schemes vary, the comparability between the different levels of PSNR is consistent across all three media clips. Also observe the convergence of ALD and MDC-SDP in both the city and harbour streams. This is consistent with the initial low PSNR values and high transmission byte-cost for MDC for both streams. Future work will further evaluate this convergence.

Table 9: SVC transmission byte-cost for each quality layer, of the additional media clips

Layer	City	Harbour	Soccer
8	14,797,807	16,311,564	10,189,375
7	11,179,340	13,017,572	7,665,548
6	8,409,220	10,412,045	5,851,889
5	4,731,469	6,585,181	3,607,645
4	3,514,773	5,197,057	2,636,439
3	2,118,932	3,469,141	1,639,147
2	980,230	1,396,047	781,908
BL	358,483	577,249	319,650

Table 10: Additional stream maximum PSNR and transmission byte-cost per adaptive scheme c/w percentage value compared to SVC

Scheme	City	Harbour	Soccer
PSNR	36.47	36.78	37.72
SVC	14,797,807 (100%)	16,311,564 (100%)	10,189,375 (100%)
MDC	24,811,448 (168%)	30,466,240 (187%)	17,823,312 (175%)
ALD	20,392,250 (138%)	21,091,057 (129%)	13,637,558 (134%)

5. CONCLUSION

In this paper, Adaptive Layer Distribution (*ALD*) is proposed as a novel multifaceted approach to media streaming optimisation. ALD section thinning enables reducing the total streaming overhead while IER and section distribution improve ALD error resiliency to loss. Hence, ALD strikes a balance between stream quality and bandwidth efficiency. Ongoing work includes a testbed implementation of an ALD prototype, as well as extensive subjective testing. Future work includes investigating the optimal policy for choosing ALD parameters in different settings, while also evaluating the various schemes with larger GOF and when using retransmission mechanisms.

6. ACKNOWLEDGEMENT

The authors thank the anonymous reviewers for their valuable feedback and acknowledge the support of Science Foundation Ireland (SFI) under Research Grant 10RFP/CMS2952. The authors would like also to acknowledge the support of the National Telecommunication Regulation Authority (NTRA) of Egypt.

7. REFERENCES

[1] H. Bai, A. Wang, Y. Zhao, J.-S. Pan, and A. Abraham. *Distributed Multiple Description Coding*. Principles, Algorithms and Systems. Springer-Verlag New York Inc, Nov. 2011.

[2] T. Berkin Abanoz and A. Murat Tekalp. SVC-based scalable multiple description video coding and optimization of encoding configuration. *Signal Processing: Image Communication*, 24(9):691–701, 2009.

[3] P. Chou and H. Wang. Layered multiple description coding. *IEEE International Packet Video Workshop (PV2003)*, 2003.

[4] Cisco Visual Networking Index, 2010-2015. Available at:. http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-481360.pdf.

[5] P. Frojdh, U. Horn, M. Kampmann, A. Nohlgren, and M. Westerlund. Adaptive streaming within the 3GPP packet-switched streaming service. *IEEE Network*, 20(2):34–40, 2006.

[6] T. Gan, L. Gan, and K.-K. Ma. Reducing video-quality fluctuations for streaming scalable video using unequal error protection, retransmission, and interleaving. *IEEE Transactions on Image Processing*, 15(4):819–832, 2006.

[7] V. Goyal. Multiple description coding: Compression meets the network. *Signal Processing Magazine*, 18(5):74–93, Sept 2002.

[8] C. Hellge, D. Gomez-Barquero, T. Schierl, and T. Wiegand. Layer-Aware Forward Error Correction for Mobile Broadcast of Layered Media. *IEEE Trans. on Multimedia*, 13(3):551–562, 2011.

[9] C. H. Ke. myEvalSVC: an Integrated Simulation Framework for Evaluation of H.264/SVC Transmission. *KSII Transactions on Internet and Information Systems*, 6(1):379–394, January 2012.

[10] K. Park and W. Wang. AFEC: an adaptive forward error correction protocol for end-to-end transport of real-time traffic. In *Proceedings of the 7th International Conference on Computer Communications and Networks*, pages 196–205, 1998.

[11] M. Przybylski, B. Belter, and A. Binczewski. Shall we worry about Packet Reordering? *Computational Methods in Science and Technology*, pages 141–146, Dec. 2005.

[12] R. Puri, K. Ramchandran, and K. Lee. Forward error correction (FEC) codes based multiple description coding for Internet video streaming and multicast. *Signal Processing: Image Communication*, 16(8):745–762, May 2001.

[13] J. Reichel, H. Schwarz, and M. Wien. Joint scalable video model 11 (JSVM 11). *Joint Video Team, doc.*, Jul. 2007.

[14] D. Salomon. *Guide to Data Compression Methods.* Springer, 2002.

[15] T. Schierl, T. Stockhammer, and T. Wiegand. Mobile Video Transmission Using Scalable Video Coding. *IEEE Trans. Circuits and Systems for Video Technology*, 17(9):1204 – 1217, 2007.

[16] A. Shokrollahi. Raptor codes. *Information Theory, IEEE Trans. on*, 52(6):2551–2567, 2006.

[17] T. Stockhammer. Dynamic Adaptive Streaming over HTTP – Standards and Design Principles. In *MMSys '11 Proceedings of the second annual ACM conference on Multimedia systems*, page 133, New York, New York, USA, 2011. ACM Press.

[18] The Network Simulator - ns-2. http://nsnam.isi.edu/nsnam/index.php.

[19] I. Unanue, I. Urteaga, R. Husemann, J. Del Ser, V. Roesler, A. Rodriguez, and P. Sanchez. A Tutorial on H.264/SVC Scalable Video Coding and its Tradeoff between Quality, Coding Efficiency and Performance. In *Recent Advances on Video Coding*, pages 1–24. InTech, June 2011.

[20] Video traces for network performance evaluation. Available at:. ftp://ftp.tnt.uni-hannover.de/pub/svc/testsequences/.

[21] Z. Zhao et al. Multiple description scalable coding for video streaming. *2009 10th Workshop Image Analysis for Multimedia Interactive Services. WIAMIS '09.*, pages 21–24, 2009.