P-Frame Injection for Efficient Packet-Loss Repair in Ultra-Low-Latency Video Streaming

Hannes Mareen, Peter Lambert, and Glenn Van Wallendael
IDLab, Ghent University – imec, Ghent, Belgium, e-mail: firstname.lastname@ugent.be

***Abstract*—Applications providing ultra-low-latency video streaming to large audiences require fast and efficient packet-loss repair. Previous methods utilizing keyframe injection (such as the High Efficiency Streaming Protocol) have a low impact on the repaired stream quality, but at a cost of a significant bitrate spike during repair. In this paper, we propose injecting P-frames for packet-loss repair of ultra-low-latency streaming. We implemented (open-source) and evaluated our approach in both H.265/HEVC and H.266/VVC standards. Through extensive evaluations, we demonstrate that the proposed solution significantly reduces bitrate overhead while maintaining a similar or lower decrease in quality compared to existing packet-loss-repair techniques. Overall, the proposed approach offers a promising solution to ensure reliable packet-loss recovery, efficient resource utilization, and a high-quality streaming experience.**

***Index Terms*—Packet-loss repair, ultra-low-latency video streaming, High Efficiency Video Coding (HEVC), Versatile Video Coding (VVC), High Efficiency Streaming Protocol (HESP).**

I. INTRODUCTION

The ultra-low-latency distribution of video content to a diverse range of end-user devices with varying connectivity characteristics poses significant challenges. In an ideal streaming system, the goal is to deliver a high-quality experience to users with stable connections while also catering to those with low-fidelity connections. However, traditional video distribution systems often penalize users with stable connections by sacrificing overall performance to accommodate the packet-loss repair needs of a few users.

These drawbacks of typical packet-loss restrictions become apparent when considering the prediction structure of a video stream. On one hand, maximizing compression efficiency necessitates the use of numerous inter-predictions. On the other hand, packet-loss restrictions require the frequent insertion of inefficient intra-predicted keyframes or other forms of intra refresh. This frequency is determined by the intra-period. Striking a balance between a large intra-period for efficient compression and smaller intra-period for faster packet-loss repair is a challenging task, as it is inherently difficult to cater to the needs of both user groups simultaneously.

To ensure that end-users with high-fidelity connections are not burdened, innovative solutions such as the High Efficiency Streaming Protocol (HESP) employ for example keyframe injection (KI also known as keyframe insertion in prior research) [1]–[3]. In this approach, all users receive a (potentially multi-cast) normal stream (NS) that has infrequent keyframes, and hence is very compression efficient. To cater to users that suffer from packet loss, an additional companion stream (CS) is encoded that solely consists of keyframes (possibly at a lower frame rate). When clients suffer from packet loss, and there is no time for an automatic repeat request (ARQ), these can request a keyframe from the CS (in unicast) to quickly (re)start the stream. Although this is a promising solution and is already applied in practice in HESP, injecting a keyframe causes a short bitrate spike, as the encoded CS keyframe is much larger than the NS predicted (P) picture that it replaces.

In this paper, we propose a novel approach to repair packet losses through the injection of reference repair pictures (RRP). We define RRPs as P-frames that contain all the necessary information to be inserted and to reference temporally distant frames. These pictures reference a picture preceding the lost picture and consequently repair the stream. In contrast to the injection of keyframes [1]–[3], P-frames are much smaller in size and therefore do not cause a significant bitrate spike. In contrast to keyframe injection, injecting RRPs do not accommodate for random access. Instead, RRP-frame injection can only be used for packet-loss repair. The proposed methods targets ultra-low-latency streaming (using low-delay video compression configurations). Novelty of this paper can be found in (1) a thorough explanation on how to construct these RRPs in recent video compression standards, (2) open-source software to reproduce the proposed solution, and (3) an in-depth bitrate and quality analysis on H.265/HEVC and H.266/VVC.

II. RELATED WORK

Solutions tackling packet loss can be categorized in three types: client based, network based, and content based. Error concealment [4], [5] is an example client-based solution, but as illustrated in the results section, one with significant quality decrease. Network-based techniques require undesirable computational resources in the network, such as transcoding [6]–[8] or network-distributed video coding [9], [10].

Finally, content-based techniques are the ones where additional versions of the video stream are generated to resolve packet loss. Switching Intra (SI) frames and Switching Predicted (SP) frames from H.264/AVC can solve packet-loss issues similar to the proposed technique, but at a rate and quality impact of the normal stream [11], [12]. In contrast, our proposed technique avoids a negative impact for client devices that have a high-fidelity connection. A solution to this is keyframe injection, where an intra-predicted (I) version of

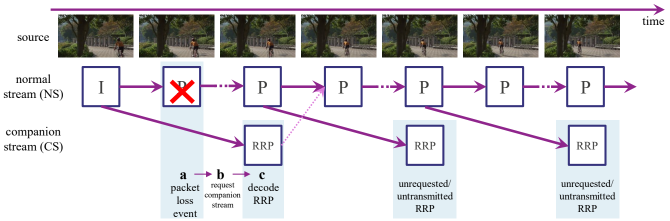

Fig. 1. Only after (a) packet-loss, (b) a companion stream RRP is requested and (c) it replaces the collocated predicted frame. Note that succeeding unrequested RRPs remain untransmitted.

a predicted (P) picture is introduced in the stream to solve a packet loss. Keyframe injection has been introduced in H.263 [13], H.264/AVC [14], [15], and HESP [1]–[3], [16].

To reduce the bitrate spike of the I picture, mixed-resolution keyframe injection has been proposed [17]. In that approach, the resolution of the CS is reduced. However, this has a significant negative impact on the quality. Additionally, changing the spatial resolution mid-stream is only possible in H.266/VVC, and not in H.264/AVC and H.265/HEVC.

In other works, you will notice that the NS [2], [3], [14] has also been named the continuation stream [1], the P-stream [13], or main stream [15]. The CS [2], [3] is known as the initialization stream [1], the I-stream [13], synchronization stream [15], the channel change stream [14].

III. REFERENCE REPAIR PICTURES (RRP)

As illustrated in Figure 1, the NS is accompanied by a CS containing RRPs which predict from pictures further in the past. Note that the companion stream remains on the server (or Content Delivery Network CDN) and interactively solves packet loss whenever the client device requests an RRP from the CS. Because the CS is synchronized with the NS, the first available RRP is injected in the NS or alternatively transmitted to the end-user device. Subsequently, packet-loss corrupted frames present at the client device will be played back followed by the RRP from the CS that replaces the collocated inter-frame. With this action, packet-loss drift errors are resolved. All while at the end user, the client device receives a single standard-compliant video stream that consists of Network Abstraction Layer Units (NALUs) combined from the NS and CS.

The proposed RRP-based method is different from earlier keyframe-injection methods [2], [3], [16], [17] since we inject P-frames rather than I-frames. As such, the bitrate of the injected P-frame of the proposed method can be lower than those of injected I-frames. Moreover, it differs from packet-loss-repair methods based on SP- and SI-frames [11], [12] since we inject P-frames that are *not* identical to the P-frame that it replaces, whereas the SP- and SI-frames are identical to the frames that they replace. As such, the bitrate of the injected P-frame can be lower, and there is no negative effect on the NS bitrate.

When injecting RRPs, the system has two options, namely identical parameter sets or parameter set replacement. Parameter sets that need to be considered are Video Parameter Set (VPS), Sequence Parameter Set (SPS) and Picture Parameter Set (PPS). When these are identical in both the NS and CS, then the RRP can replace the picture of the normal stream. If there is a difference between the sets of both streams, then the injected RRP should be preceded with the parameter sets of the CS and succeeded again with the parameter sets of the NS.

Because H.266/VVC has the concept of Adaptation Parameter Sets (APS) [18], the CS encoder or a post encoding procedure needs to perform the following stateful process. An APS provides three different types of information (aps_params_type), namely parameters of the Adaptive Loop Filter (ALF), luma mapping with chroma scaling (LMCS) parameters, and scaling list parameters. During the encoding of the NS and CS streams, a list of the APS packets with unique identification (ID) numbers (aps_adaptation_parameter_set_id) needs to be stored and intelligently concatenated to the CS P-frame to form the RRP. The solution is to keep and update a dictionary of unique APS packets during the processing of the NS and the CS. Before the CS P-frame, the APS of the CS P-frame, the one having the right APS-ID, needs to be inserted from the CS APS dictionary. Immediately after the CS P-frame, the APS having the same APS-ID from the NS APS dictionary needs to be inserted in the video stream again. This is necessary because predicted frames succeeding the injected keyframe could make use of the preceding APS information. So, the concatenation of the CS APS, the CS P-frame and the NS APS form the RRP frame. Uniqueness of the APS IDs and thus removal of earlier duplicates is crucial to guarantee standard compliance.

IV. EVALUATION

A. Experimental Setup

To evaluate the proposed method and compare it with the state of the art, we followed a similar approach as in related work [2], [3], [16], [17]. That is, we created a bitrate ladder of normal streams in four quality levels for a large and diverse set of 22 video sequences, using two coding standards. We additionally created corresponding companion streams in four quality levels. Then, we evaluated the impact on the bitrate and quality when injecting RRPs from the CS into the NS. Additionally, we compare the performance with keyframe injection [2], [3], mixed-resolution keyframe injection [17], and traditional frame-copy packet-loss repair (in which the previous frame is simply copied when a frame is lost).

We evaluated RRP injection in both the H.265/HEVC and H.266/VVC standard. More specifically, we used the HEVC reference Model (HM) [19] version 16.15 for H.265/HEVC, and the Fraunhofer Versatile Video Encoder (VVenC) [20] version 10.2 for H.266/VVC. Note that mixed-resolution KI [17] was proposed and evaluated using the VVC Test Model (VTM), whereas we implemented and evaluated our proposed method in VVenC. Keyframe injection [2], [3] was evaluated in all three encoders (HM, VVenC, and VTM), though. Even though mixed-resolution KI and RRP injection are evaluated on different encoders, we argue that we can still compare these methods to each other, because they are both compared to keyframe injection. In other words, we can compare their *relative* performance compared to KI.

The NS videos were encoded using a low-delay configuration in which the first frame is a keyframe and all other frames are predicted inter-frames that each take only the preceding frame as reference (i.e., an IPPP structure). This configuration is chosen since this paper focuses on ultra-low delay real-time video. The CS was encoded with the same configuration as the NS, where the RRP, i.e. the frame-to-be-injected, refers to a frame further in the past. We evaluated a reference frame at 4 frames prior to the encoded frame (Ref-4), as well as 8 frames prior (Ref-8). This means that these CS frames can repair packet losses that occurred in the last 3 or 7 frames, respectively. These configurations allow for ultra-low-latency delivery of 25 frames per second video with network round trip times of 120ms and 280ms respectively. In our experiments, as well as those in the state of the art, the frame of the CS is injected in the NS at frame $f = 16$, which is the 17th frame. Hence, the decoding can start from the injected frame as if any frame packet(s) between $f = 13$ and $f = 16$ had been lost for Ref-4, or between $f = 9$ and $f = 16$ for Ref-8.

The experiments were performed on 22 test sequences with resolutions between 416×240 and 2560×1600, containing between 150 and 600 frames, with framerates between 20 and 60 frames per second [21]. Each sequence is compressed using four different Quantization Parameters (QPs), namely 22, 27, 32, and 37, denoted as $\mathrm{QP}_{\mathrm{NS}}$ and $\mathrm{QP}_{\mathrm{CS}}$.

The source code of RRP injection has been made available for reproducibility[1]. Additionally, some example videos and additional images with results are shown on our website[2].

B. Impact on Frame Size

In Table I, the factor of frame size increase is presented for multiple configurations and methods. This factor indicates the ratio by which the size of the injected frame is larger than the frame of the NS that it replaces. These values represent the medians calculated across all test sequences to minimize the impact of outliers. When $\mathrm{QP}_{\mathrm{NS}} \neq \mathrm{QP}_{\mathrm{CS}}$, the results are displayed in gray instead of black, in order to aid comparison of the evaluated configurations. Additionally, note that the frame size increase factors are not reported for the frame-copy packet-loss repair approach, as they are all zero. The reported frame size increase values provide insights into the bitrate spike when comparing the proposed RRP injection with existing (multi-resolution) keyframe-injection (KI) methods.

At equal QPs ($QP_{NS} = QP_{CS}$), the injected RRP size increase factors are between 1.5 and 2.9, which is in stark contrast with the inject keyframes that are between 5.0 and 29.2 times larger than the inter-frames that they replace. In other words, RRP injection has a bitrate spike that is approximately 3 to 10 times smaller than keyframe injection.

TABLE I

MEDIAN FACTOR OF FRAME SIZE INCREASE DUE TO PACKET-LOSS REPAIR BY INJECTING AN RRP (PROPOSED) OR KEYFRAME (KI, [2], [3], [17]).

Codec	$\mathrm{QP}_{\mathrm{CS}}$	$\mathrm{QP}_{\mathrm{NS}}$ 22	27	32	37	22	27	32	37	22	27	32	37
		RRP, Ref-4				RRP, Ref-8				KI [2]			
HM	22	1.5	4.7	11.7	26.7	1.7	5.6	14.2	32.1	8.6	27.4	68.2	149.4
	27	0.6	1.8	4.7	9.7	0.7	2.4	6.0	13.1	5.3	16.1	39.8	86.4
	32	0.3	0.8	1.9	4.4	0.3	1.0	2.5	6.0	3.3	9.7	23.6	50.7
	37	0.2	0.4	1.0	2.2	0.2	0.5	1.3	2.9	2.0	5.8	13.8	29.2
		RRP, Ref-4				RRP, Ref-8				KI [3]			
VVEnc	22	1.5	3.3	7.2	16.0	1.7	3.7	8.1	18.8	5.0	11.7	25.2	62.1
	27	0.7	1.6	3.5	7.7	0.9	2.0	4.1	9.6	2.8	6.5	14.6	33.1
	32	0.4	0.9	1.8	3.9	0.5	1.0	2.1	4.6	1.7	3.5	8.2	18.5
	37	0.2	0.4	0.9	1.8	0.3	0.5	1.1	2.3	0.9	2.1	4.6	10.0
		KI, $\mathrm{Res}_{\mathrm{CS}}$ = 2.0 [17]				KI, $\mathrm{Res}_{\mathrm{CS}}$ = 1.5 [17]				KI [3], [17]			
VTM	22	2.8	8.0	20.1	44.8	4.3	13.0	32.3	70.1	8.7	28.6	63.8	142.6
	27	1.7	5.1	13.3	29.8	2.6	8.1	20.4	45.7	4.9	16.8	39.5	88.6
	32	1.0	3.1	8.2	18.4	1.5	4.7	12.5	28.0	2.7	9.3	23.4	51.3
	37	0.6	1.7	4.7	10.6	0.8	2.6	7.2	16.2	1.5	4.9	12.6	27.5

Comparing RRP Ref-4 and Ref-8, we can observe that Ref-8 RRPs are approximately 20% larger than corresponding Ref-4 RRPs. For example, for VVEnc, at $QP_{NS} = QP_{CS} = 32$, the frame size increase is 1.8 for Ref-4, and 2.1 for Ref-8. It is not a surprise that Ref-8 RRPs are coded less efficiently, as a frame that is further in the past is generally a worse choice as reference picture compared to a more-recent frame. As such, Ref-8 contains more residual errors that need to be corrected and encoded.

An alternative to reduce the frame size of injected keyframes is by lowering the resolution of those keyframes, which is possible using Mixed-Resolution KI [17]. For example, when using a resolution of half the width and height ($\mathrm{Res}_{\mathrm{CS}} = 2.0$), and at equal QPs ($QP_{NS} = QP_{CS}$), the keyframes are between 2.8 and 16.2 times larger than the P-frames in the NS that they replace. These values can be reduced by lowering the quality of the keyframes (i.e., increasing the $\mathrm{QP}_{\mathrm{CS}}$).

It should be stressed that these results are not a factor of bitrate increase of the entire video sequence, but rather only the frame size increase from the frame in the NS to the frame of the CS that is injected during packet-loss repair. The overall bitrate increase will be several magnitudes smaller, depending on how frequently a packet-loss repair has to be performed. Furthermore, if no packet-loss repair is performed in a video segment, there is no bitrate overhead at all.

C. Impact on Quality

We employed the Video Multimethod Assessment Fusion (VMAF) [22] to assess the video quality. VMAF produces a score ranging from 0 to 100, where a score of 100 indicates that the two compared videos are subjectively indistinguishable, and it has been claimed that a difference of 6 points represents a just-noticeable difference (JND) [23], [24]. The VMAF model used in this study is *vmaf_4k_v0.6.1*. In addition, we computed the Peak Signal-to-Noise Ratio (PSNR)

---

[1] Code available on https://github.com/IDLabMedia/NALUProcessing

[2] Example videos and more detailed results available on https://media.idlab.ugent.be/p-frame-injection

and the Structural Similarity Index Measure (SSIM). We omit the PSNR and SSIM results from this paper to save space, but observed similar results as using VMAF. Each quality measure was calculated between the uncompressed video and NS, as well as between the uncompressed video and RRP or KI, and this from $f = 16$ (i.e., the repaired frame) onwards. By comparing these two sets of measurements, we determined the decrease in quality from NS to RRP/KI, denoted as ΔQuality. Note that there is no CS in the frame-copy packet-loss method. However, for compactness and comparison with the other methods and configurations, we gave frame-copy ΔQuality values a separate column instead of row.

Table II presents the median VMAF decrease (ΔVMAF) for all examined configurations. To calculate these decreases, VMAF scores were averaged over all frames within each test sequence. Subsequently, the median of these average VMAF decreases for each test sequence was computed. By utilizing the median value, the impact of notable outliers is mitigated. Complementary to the mean ΔVMAF results, we present a graph with more detailed ΔVMAF results on our website^2.

We can observe that the ΔVMAF is comparable for both RRP injection configurations (Ref-4 and Ref-8), as well as for (full-resolution) keyframe injection, even though the factor of frame size increases in Table I differed significantly. One exception is in the upper-right corner of each 4×4 table, where the injected frame has a lower QP_{CS} than the NS. In this case, we notice a slight improvement of quality when injecting a keyframe, whereas the corresponding RRP injections still showcase a low-quality reduction. However, this should be placed in the context of the corresponding frame size increases of KI, which are very high (up to a factor of 149.4).

We visually inspected the RRP-injected packet-loss-repaired streams, and noted the same types of artifacts as in keyframe injection [2], [3], [17]. Some visual examples are given on our website^2. In any case, the reported median ΔVMAF values of the proposed method are significantly below 6, suggesting that the median effect on the quality is imperceptible.

When using keyframes of lower resolutions (i.e., mixed-resolution KI [17]), the impact on the quality increases significantly, as can be seen in the lowest part of Table II. For reference, when performing same-resolution KI with the VTM-encoder, at equal QPs ($QP_{NS} = QP_{CS}$), the ΔVMAF is between 0.3 and 0.8, which is similar to keyframe injection in HM and VVEnc. In contrast, injecting a lower-resolution keyframe increases the impact on the quality. For example, when using a resolution of half the width and height ($\mathrm{Res}_{\mathrm{CS}} =$ 2.0), at equal QPs ($QP_{NS} = QP_{CS}$), the ΔVMAF is between 2.8 and 3.8, which is significantly higher than same-resolution KI or RRP injection. Furthermore, the corresponding frame size increases were also much larger than for RRP injection (see Table I). As such, this demonstrates that RRP injection is a more efficient solution for (sparse) packet-loss repair compared to mixed-resolution keyframe injection.

When using a traditional frame-copy packet-loss repair solution, the impact on the quality increases even more than when using mixed-resolution KI with low-resolution keyframes. This is especially true at higher qualities (i.e., smaller QPs). For example, when utilizing a frame-copy approach with the HM encoder, we get a ΔVMAF of 3.4, on average. In contrast, the proposed RRP injection method with Ref-4 reports a 0.4 ΔVMAF, which is approximately 8 times lower. However, it must be noted that the frame-copy method has zero bitrate overhead and is less costly to implement.

TABLE II
MEDIAN DECREASE IN VMAF DUE TO PACKET-LOSS REPAIR BY INJECTING AN RRP (PROPOSED), KEYFRAME (KI, [2], [3], [17]), OR FRAME COPY.

Codec	$\mathrm{QP}_{\mathrm{CS}}$	$\mathrm{QP}_{\mathrm{NS}}$												
		22	27	32	37	22	27	32	37	22	27	32	37	
		RRP, Ref-4				RRP, Ref-8				KI [2]				Frame copy
HM	22	0.2	0.8	0.8	0.6	0.3	0.7	0.6	0.5	0.2	0.5	0.3	-0.3	4.5
	27	0.6	0.5	0.8	0.7	0.7	0.6	0.7	0.5	0.4	0.5	0.2	-0.3	4.2
	32	1.2	1.2	0.6	0.5	1.5	1.4	0.7	0.5	0.9	1.1	0.6	-0.2	3.3
	37	2.0	2.1	1.3	0.4	3.1	2.3	1.4	0.5	2.3	2.5	1.8	0.3	1.7
		RRP, Ref-4				RRP, Ref-8				KI [3]				Frame copy
VVEnc	22	0.9	1.4	1.4	1.1	0.9	1.5	1.7	1.2	0.9	1.2	1.1	0.5	6.4
	27	1.4	1.2	1.4	1.2	1.3	1.3	1.5	1.3	1.1	1.4	1.2	0.6	6.4
	32	1.6	2.1	1.2	1.1	2.1	2.3	1.4	1.2	1.7	2.5	1.5	0.6	5.2
	37	2.7	3.0	2.0	1.2	2.7	3.3	2.1	1.1	3.0	3.9	3.2	1.3	3.6
		KI, $\mathrm{Res}_{\mathrm{CS}}$ = 2.0 [17]				KI, $\mathrm{Res}_{\mathrm{CS}}$ = 1.5 [17]				KI [3], [17]				Frame copy
VTM	22	2.8	2.8	1.5	0.6	2.3	2.8	1.6	0.5	0.3	0.7	0.3	-0.1	5.3
	27	3.9	3.8	2.2	1.0	3.0	3.0	1.9	0.7	0.7	0.8	0.4	-0.1	4.7
	32	6.1	5.9	3.6	1.3	4.3	3.8	2.5	1.1	1.1	1.3	0.7	-0.1	3.0
	37	10.1	8.3	5.4	3.2	5.9	6.2	3.6	2.6	2.0	2.6	1.7	0.3	1.8

V. CONCLUSION

We proposed to repair packet losses in ultra-low-latency streaming by preparing a companion stream that consists of reference repair pictures, which are P-frames and corresponding APS packets that use frames prior to the packet loss as reference. These RRPs can be injected into a compression-efficient normal stream after a packet loss occurs. We implemented this in H.265/HEVC and H.266/VVC, made the software open-source, and list the corresponding requirements.

We evaluated the effect of RRP injection on the bitrate spike during repair, as well as the quality loss that occurs afterwards. We found that the injected RRP is up to 10 times smaller than injected keyframes from state-of-the-art packet-loss repair methods with a comparable impact on the quality. We additionally compared the proposed RRP injection method to mixed-resolution KI and traditional frame-copy packet-loss repair approaches, and found that RRP injection performs best in terms of quality loss.

In conclusion, this paper presents a promising solution that addresses the challenges of reliable packet-loss recovery, efficient resource utilization, and the delivery of a high-quality ultra-low-latency streaming experience.

ACKNOWLEDGMENT

The computational resources (STEVIN Supercomputer Infrastructure) and services used in this work were kindly provided by Ghent University, the Flemish Supercomputer Center (VSC), the Hercules Foundation and the Flemish Government department EWI.

REFERENCES

[1] P.-J. Speelmans, "HESP - High Efficiency Streaming Protocol," Internet Engineering Task Force, Internet-Draft draft-theo-hesp-01, Nov. 2021, work in Progress. [Online]. Available: https://datatracker.ietf.org/doc/html/draft-theo-hesp-01

[2] G. Van Wallendael, H. Mareen, J. Vounckx, and P. Lambert, "Keyframe insertion: Enabling low-latency random access and packet loss repair," *Electronics*, vol. 10, no. 6, 2021.

[3] H. Mareen, M. Courteaux, J. Vounckx, P. Lambert, and G. Van Wallendael, "Keyframe insertion for random access and packet-loss repair in H.264/AVC, H.265/HEVC, and H.266/VVC," in *Data Compression Conf. (DCC)*, 2022.

[4] M. Kazemi, M. Ghanbari, and S. Shirmohammadi, "A review of temporal video error concealment techniques and their suitability for HEVC and VVC," *Multimedia Tools and Applications*, 2021.

[5] W.-Y. Kung, C.-S. Kim, and C.-C. Kuo, "Spatial and temporal error concealment techniques for video transmission over noisy channels," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 16, no. 7, pp. 789–803, 2006.

[6] A. Vetro, C. Christopoulos, and H. Sun, "Video transcoding architectures and techniques: an overview," *IEEE Signal Processing Magazine*, vol. 20, no. 2, pp. 18–29, 2003.

[7] A. Erfanian, F. Tashtarian, R. Farahani, C. Timmerer, and H. Hellwagner, "On Optimizing Resource Utilization in AVC-based Real-time Video Streaming," in *2020 6th IEEE Conference on Network Softwarization (NetSoft)*, 6 2020, pp. 301–309.

[8] L. Pham Van, J. De Praeter, G. Van Wallendael, S. Van Leuven, J. De Cock, and R. Van de Walle, "Efficient Bit Rate Transcoding for High Efficiency Video Coding," *IEEE Transactions on Multimedia*, vol. 18, no. 3, pp. 364–378, 3 2016.

[9] J. De Praeter, G. Van Wallendael, J. Slowack, and P. Lambert, "Video Encoder Architecture for Low-Delay Live-Streaming Events," *IEEE Transactions on Multimedia*, vol. 19, no. 10, pp. 2252–2266, 10 2017.

[10] J. D. Praeter, C. Hollmann, R. Sjoberg, G. V. Wallendael, and P. Lambert, "Network-Distributed Video Coding," *arXiv*, 2021.

[11] M. Karczewicz and R. Kurceren, "The SP- and SI-frames design for H.264/AVC," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 7, pp. 637–644, 7 2003.

[12] W.-t. Tan and G. Cheung, "Using sp-frames for error resilience in optimized video streaming," in *2006 International Conference on Image Processing*, 2006, pp. 3049–3052.

[13] N. Farber, E. Steinbach, and B. Girod, "Robust H.263 compatible transmission for mobile video server access," in *Proceedings of First International Workshop on Wireless Image/Video Communications*, 1996, pp. 8–13.

[14] J. M. Boyce and A. M. Tourapis, "Fast efficient channel change [set-top box applications]," in *2005 Digest of Technical Papers. International Conference on Consumer Electronics, 2005. ICCE.*, 1 2005, pp. 1–2.

[15] U. Jennehag and S. Pettersson, "On synchronization frames for channel switching in a GOP-based IPTV environment," in *2008 5th IEEE Consumer Communications and Networking Conference, CCNC 2008*, 1 2008, pp. 638–642.

[16] H. Mareen, M. Courteaux, P. Lambert, P.-J. Speelmans, and G. Van Wallendael, "A study on keyframe injection in three generations of video coding standards for fast channel switching and packet-loss repair," *Multimedia Tools and Applications*, 2023, accepted in Sept. 2023.

[17] G. Van Wallendael, P. Lambert, P.-J. Speelmans, and H. Mareen, "Mixed-resolution HESP for more efficient fast channel switching and packet-loss repair," in *Picture Coding Symposium (PCS)*, 2022, pp. 319–323.

[18] B. Bross, Y.-K. Wang, Y. Ye, S. Liu, J. Chen, G. J. Sullivan, and J.-R. Ohm, "Overview of the versatile video coding (VVC) standard and its applications," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 31, no. 10, pp. 3736–3764, 2021.

[19] Fraunhofer, "High Efficiency Video Coding (H.265/HEVC) — JVT-VC." [Online]. Available: https://hevc.hhi.fraunhofer.de/

[20] A. Wieckowski, J. Brandenburg, T. Hinz, C. Bartnik, V. George, G. Hege, C. Helmrich, A. Henkel, C. Lehmann, C. Stoffers, I. Zupancic, B. Bross, and D. Marpe, "VVenC: An open and optimized VVC encoder implementation," in *Proc. IEEE Int. Conf. Multimedia Expo Workshops (ICMEW)*, 2021, pp. 1–2.

[21] F. Bossen, "Common test conditions and software reference configurations," ITU-T Joint Collaborative Team on Video Coding (JCT-VC), Tech. Rep. JCTVC-L1100, 2013.

[22] Netflix Technology Blog, "Toward A Practical Perceptual Video Quality Metric." [Online]. Available: https://netflixtechblog.com/toward-a-practical-perceptual-video-quality-metric-653f208b9652

[23] J. Ozer, "Finding the Just Noticeable Difference with Netflix VMAF." [Online]. Available: https://streaminglearningcenter.com/codecs/finding-the-just-noticeable-difference-with-netflix-vmaf.html

[24] H. Amirpour, R. Schatz, and C. Timmerer, "Between two and six? towards correct estimation of jnd step sizes for vmaf-based bitrate laddering," in *2022 14th International Conference on Quality of Multimedia Experience (QoMEX)*, 2022, pp. 1–4.