DEPENDENT RANDOM ACCESS POINT PICTURES IN HEVC

Martin Pettersson, Rickard Sjöberg, Jonatan Samuelsson

Ericsson Research

ABSTRACT

This paper describes the concept and a number of system aspects for the Dependent Random Access Point (DRAP) picture, a new feature that was recently introduced in the High-Efficiency Video Coding (HEVC) standard. A DRAP picture is an inter-coded picture that may only reference the previous Intra Random Access Point (IRAP) picture and provides a random access point in the bitstream at the DRAP picture given that the IRAP picture is available. The DRAP picture is indicated in the HEVC bitstream by a supplementary enhancement indication (SEI) message. DRAP pictures could increase coding efficiency for random access configured bitstreams for video services like adaptive video streaming, video conferencing, and may also provide faster channel switching for broadcasted TV for a certain bitrate.

Index Terms— HEVC, video coding, random access, MPEG-DASH, fast channel switch, DRAP

1. INTRODUCTION

The High-Efficiency Video Coding (HEVC) standard [1] which was finalized in January 2013 includes three types of intra-coded pictures. In addition to instantaneous decoding refresh (IDR) pictures known from H.264/MPEG-4 AVC [2], HEVC also supports clean random access (CRA) pictures and broken link access (BLA) pictures [3][4]. A CRA picture allows a leading picture to reference a picture that is preceding the CRA picture in display order and decoding order. In case the decoding starts at the CRA picture, the leading picture may need to be dropped. BLA pictures are used for indicating splicing points in the bitstream. CRA, BLA and IDR pictures are referred to as intra random access point (IRAP) pictures since they provide a tune-in point in the bitstream from where the pictures that follow in output order will be correctly decoded. In order to enable efficient random access in an inter-coded video bitstream IRAP pictures are typically inserted at a periodic interval as depicted in figure 1. Despite that leading pictures are allowed to reference across CRA pictures in decoding order, a lot of bits are still consumed by CRA pictures due to the fact that they are intra-coded.

In HEVC, as well as in H.264/MPEG-4 AVC, there is a supplemental enhancement information (SEI) message called Recovery Point SEI [4][5]. The recovery point SEI message assists a decoder in determining when the decoding process will produce acceptable pictures for display after the decoder initiates random access or after the encoder indicates a broken link in the bitstream.

2. OVERVIEW OF THE DRAP CONCEPT

Dependent Random Access Point (DRAP) pictures [6] have been introduced to improve the compression efficiency, or correspondingly to increase the random access point frequency, for random access configured bitstreams. A DRAP picture is an inter-coded picture that may only reference the previous IRAP picture in the bitstream, referred to as its associated IRAP picture. In addition, the pictures following the DRAP picture in output and decoding order may not reference any picture preceding the DRAP picture in output or decoding order, except for the IRAP picture associated with the DRAP picture. The encoder may decide to encode the IRAP and DRAP pictures at specific intervals as illustrated in figure 2. The decision of whether to encode a random access point as an IRAP or a DRAP picture could also be based on if the estimated bit-cost for the IRAP picture minus a threshold value exceeds the estimated bit-cost for the DRAP picture.

A video receiver that has access to a DRAP picture and its associated IRAP picture can decode and correctly display the DRAP picture and the pictures that follow the DRAP picture in output order. This will be the case even when the video receiver does not have access to the pictures that follow the IRAP picture but precede the DRAP picture in decoding order. Section 3 presents several examples of when this functionality is useful.

The presence of a DRAP picture can be signaled in the HEVC bitstream through the Dependent RAP indication SEI message. The SEI message guarantees that random access at the picture associated with the SEI message is possible if the IRAP picture associated with the DRAP picture is available. When random access is performed at a DRAP picture the associated IRAP picture should be decoded and used for reference but not displayed due to that it corresponds to an earlier time instant relative to where the random access is performed.

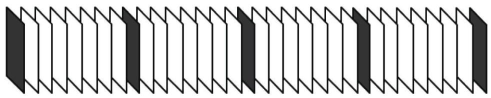

Fig. 1. Example of random access configuration using IRAP pictures (dark gray).

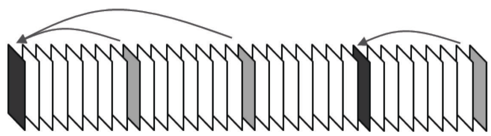

Fig. 2. Example of random access configuration with combined IRAP (dark gray) and DRAP (light gray) pictures.

3. DRAP SYSTEM ASPECTS

The advantage of using DRAP pictures compared to IRAP pictures is improved compression efficiency at the cost of having to decode the associated IRAP picture. In this section we will look further into a number of use cases where DRAP pictures are particularly useful.

3.1 Fast channel switching for broadcasted TV

In broadcasting services, IRAP pictures are typically sent at frequent intervals, e.g. once per second, to enable reasonable channel switching times. Having even more frequent IRAP pictures would be desirable to decrease the channel switching time even further, but it is not feasible since the IRAP pictures would then consume a too large portion of the available bitrate and reduce the overall quality.

One way to solve the problem of channel switching delay is to have additional tuners that receive and store data from channels other than the one that is currently being watched. However, buffering all data from several other channels might not be feasible due to limitations in receiver and/or buffer capacity. The problem can be solved through the use of DRAP pictures. Then only the latest IRAP pictures from the other channels need to be buffered and the channel switch can be performed at the first DRAP picture that occurs after the user has selected to switch channel. DRAP pictures can be sent much more frequently than IRAP pictures (e.g. every 160 ms) without significantly affecting the overall quality.

An example of channel switching using IRAP pictures every 32^{nd} picture and DRAP pictures every 8^{th} picture is shown in figure 3. A user is watching channel A. Additional tuners receive and buffer the latest IRAP picture for channel B, C and D. When the user is requesting a switch to channel B, the channel B tuner waits for the next DRAP picture before the decoder can start to decode channel B with help from the buffered IRAP picture. Similarly, when the user is switching to channel C, the decoder waits for the next DRAP picture before it can start to decode channel C. If the streams would not contain any DRAP pictures, the decoder would need to wait for the next IRAP picture before decoding the new channel after a channel switch.

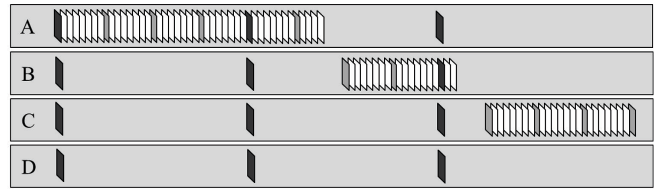

Fig. 3. Fast channel switching using DRAP pictures in broadcasted TV services.

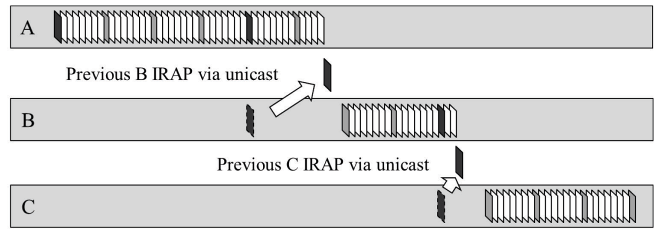

Fig. 4. Fast channel switching for IPTV broadcasted services.

Fast channel switching may also be achieved for IPTV broadcasted services using frequent DRAP pictures as illustrated in figure 4. When a channel switch is performed and the stream is tuned in at a DRAP picture, the previous IRAP picture associated with the DRAP picture is sent to the receiver via unicast. The received IRAP picture is decoded and used to decode the next DRAP picture followed by the other pictures in the stream.

3.2 Searching in stored static video content

The advantage of using DRAPs over IRAPs for random access becomes larger the more static the video scene is. This is due to that the DRAP picture then may predict more information from the associated IRAP picture.

Some video applications where the video often tends to be static include video surveillance and screen sharing applications, either for communication services or for monitoring other computers, such as servers. For these video applications it is often useful to store video logs to later be able to go back and watch certain events. Frequent random access points in the stored video stream are needed to be able to search in a good way. By using DRAP pictures instead of IRAP pictures for most of the random access points the number of bits needed to store the video can be reduced to a great extent. When there is low activity in the video, DRAP pictures could for instance be inserted every second and IRAP pictures every 60 seconds.

3.3 Massive video conference

In real-time video communication, IRAP pictures are due to their size preferably avoided in order to enable high picture quality together with a low end-to-end latency.

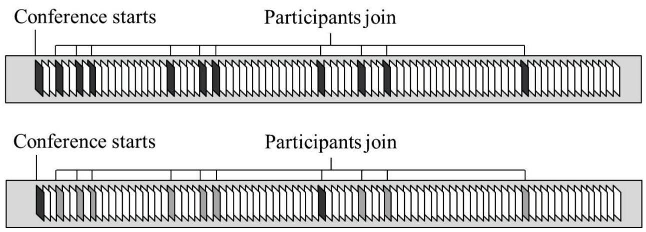

Fig. 5. Illustration of using IRAP (top) or DRAP (bottom) pictures as tune-in points in massive video conferences.

In a video conference scenario with multiple participants, the stream of the active talker is commonly sent to all other participants. As illustrated in figure 5, whenever a new participant enters the video conference a new random access point picture needs to be inserted in the active stream to allow the new participant to tune in to the stream. Using IRAP pictures for providing the random access point will result in reduced video quality and/or peaks in the bitrate. In massive video conferences, with hundreds of participants joining and leaving the call, this can severely reduce the performance and quality of experience for all the participants in the conference. DRAP pictures are encoded more efficiently than IRAP pictures, especially in typical videoconferencing scenarios where the temporal prediction from the IRAP picture works very well since the camera is static and the amount of motion is low. To be able to tune into the stream at the DRAP, the new participant also needs to receive the IRAP picture associated with the DRAP picture. The IRAP picture is preferably stored in the conference node and sent to a participant that is joining the conference before the active stream is forwarded to the participant.

3.4 Server controlled streaming

In this use case, depicted in figure 6, a Real Time Streaming Protocol (RTSP) [7] client requests a specific position in a stream to be delivered from the RTSP server, for instance by using the PLAY request with a range value in RTSP. The server will construct the stream to be delivered on-the-fly by concatenating the previous IRAP picture with the DRAP picture that corresponds to the start of the requested range.

3.5 MPEG-DASH

In Dynamic Adaptive Streaming over HTTP (DASH) [8], streams can for example be divided into segments of around ten seconds with IRAP pictures about every second to enable quick access to positions within the segment. By replacing all IRAP pictures, except the first one in the segment, with DRAP pictures, the overall compression efficiency may be increased. By including DRAP information in the Segment Index information a DASH client that recognizes the DRAP information may request to access bytes inside a segment starting from the position of a DRAP while also requesting a byte range corresponding to the IRAP picture at the start of the segment as in the example of figure 7.

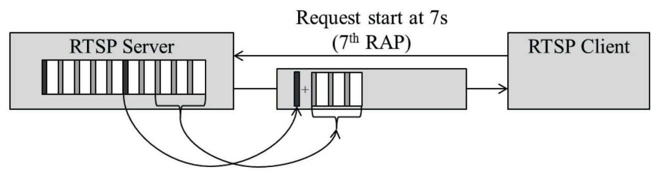

Fig. 6. In server controlled streaming the server concatenates the associated IRAP picture with the requested DRAP picture.

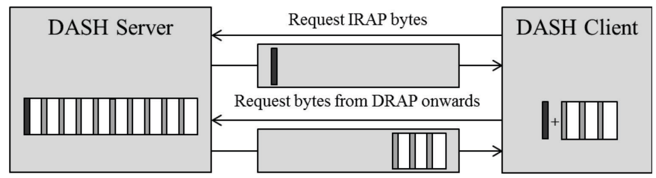

Fig. 7. Example of using DRAP pictures in MPEG-DASH.

Class	Sequence name	Nbr of frames	IRAP-period	Fps
Class A (4K)	Traffic	150	32	30
	PeopleOnStreet	150	32	30
	Nebuta	300	64	60
	SteamLocomotive	300	64	60
Class B (1080p)	Kimono	240	24	24
	ParkScene	240	24	24
	Cactus	500	48	50
	BasketballDrive	500	48	50
	BQTerrace	600	64	60
Class C (WVGA)	BasketballDrill	500	48	50
	BQMall	600	64	60
	PartyScene	500	48	50
	RaceHorses	300	32	30
Class D (WQVGA)	BasketballPass	500	48	50
	BQSquare	600	64	60
	BlowingBubbles	500	48	50
	RaceHorses	300	32	30
Class F	BasketBallDrillText	500	48	50
	ChinaSpeed	500	32	30
	SlideEditing	300	32	30
	SlideShow	500	16	20

Table 1. JCT-VC HEVC common test condition sequences

4. EXPERIMENTAL RESULTS

This section provides experimental results comparing compression efficiency for IRAP and DRAP pictures. The experiments are based on the HEVC reference software HM16.2 [9] and have been run on the RA Main test configuration of the JCT-VC HEVC common test conditions (CTC) [10]. The sequences used in the CTC are listed in table 1. Note that class E was not used for the random access configuration in the CTC. Four fixed QP points were used for each sequence, 22, 27, 32, and 37. The Bjontegaard distortion (BD) rate [11][12] was calculated for the four QP points. A negative BD rate means that the test is better than the anchor. The simulations were run using a linux LSF cluster on different machines and reliable encoding and decoding times could therefore not be obtained.

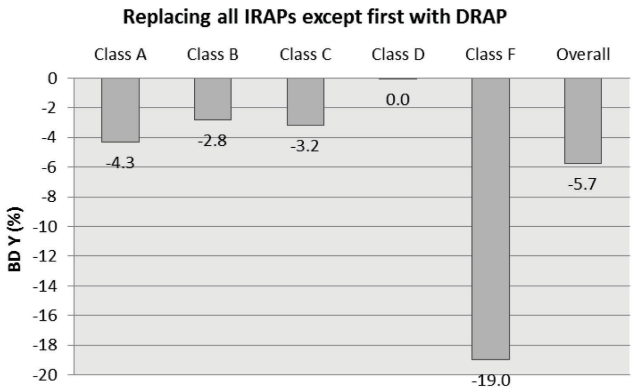

Fig. 8. Replacing all IRAP pictures but the first one with DRAP pictures. The anchor for comparison is the compressed CTC sequences as shown in table 1.

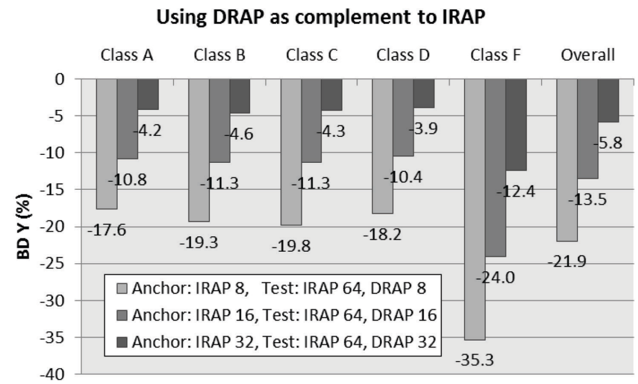

Fig. 9. Bitrate differences by using periodic DRAP pictures in combination with periodic IRAP pictures when comparing to using only periodic IRAP pictures.

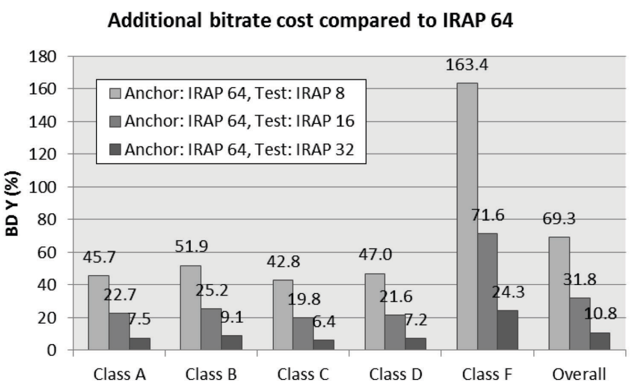

Fig. 10. Additional bitrate cost of decreasing IRAP period compared to IRAP period 64.

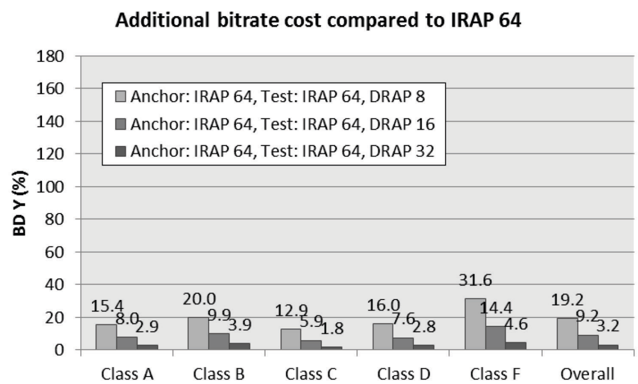

Fig. 11. Additional bitrate cost of decreasing random access point period by inserting DRAP pictures at intervals 8, 16 and 32, compared to IRAP period 64.

Modifications were made to the reference software for the DRAP simulations. Long-term pictures were implemented in the encoder to enable referencing associated IRAP pictures outside the current GOP. It was also made sure that the reference picture set (RPS) for each DRAP picture did not contain any pictures in RefPicSetStCurrBefore, RefPicSetStCurrAfter or RefPicSetLtCurr except the associated IRAP picture. For the pictures following a DRAP picture in output order it was made sure that their RPS did not contain any pictures preceding the DRAP picture except the IRAP picture associated with the DRAP picture.

In a first experiment all IRAP pictures except the first one in each CTC sequence are replaced by DRAP pictures. The results are shown in figure 8 where CTC is the anchor. Class F that comprises sequences with low motion benefits the most from using DRAP pictures. In a second experiment the bitrate savings of using DRAP pictures in combination with IRAP pictures are investigated. These results, shown in figure 9, indicate the potential for the fast channel switching use cases. Experimental data from the second experiment is also used to measure the bitrate impact from increasing the random access point periodicity with and without using DRAP pictures. Figure 10 shows the additional bitrate cost for decreasing the IRAP period compared to using IRAP period 64 and figure 11 shows the additional bitrate cost for decreasing the random access point period by introducing DRAP pictures at intervals 8, 16 and 32 in a bitstream with IRAP period 64. In a random access operation, the average extra overhead for accessing a single random access point in an IRAP/DRAP configuration compared to an IRAP only configuration were 54.7%, 47.1% and 31.6% for random access point periods of 8, 16 and 32 respectively.

5. CONCLUSION

The DRAP picture described in this paper provides increased compression efficiency for random access configured bitstreams. Alternatively, random access points may be inserted more often given a certain bitrate. Some use cases where DRAP pictures may be particular useful includes MPEG-DASH, massive video conferences, video logging of surveillance and screen sharing applications and faster channel switching in broadcasted TV. Experimental results show that DRAP pictures in combination with IRAP pictures can decrease the BD Y rate by 21.9%, 13.5% and 5.8% for random access point periods of 8, 16 and 32 respectively.

6. REFERENCES

[1] *High efficiency video coding,* ISO/IEC 23008-2 and ITU-T Rec. H.265, Oct. 2014

[2] *Advanced Video Coding for Generic Audiovisual Services*, ISO/IEC 14496-10 (MPEG-4 AVC) and ITU-T Rec. H.264, 2003

[3] V. Sze, M. Budagavi, G. J. Sullivan et al, *High Efficiency Video Coding (HEVC) Algorithms and Architectures,* Springer International Publishing, Switzerland, 2014

[4] R. Sjöberg, Y. Chen, A. Fujibayashi, M.M. Hannuksela, J. Samuelsson, T. K. Tan, Y-K, Wang and S. Wenger, "Overview of the HEVC High-Level Syntax and Reference Picture Management", *IEEE transactions on circuits and systems for video technology*, vol 22, no 12, Dec. 2012

[5] T. Schierl, M.M. Hannuksela, Y-K. Wang, S. Wenger, "System Layer Integration of High Efficiency Video Coding", *IEEE transactions on circuits and systems for video technology*, vol 22, no 12, Dec. 2012

[6] M. Pettersson, J. Samuelsson and R. Sjöberg, *HLS: Dependent RAP indication SEI message*, JCTVC-S0095, Strasbourg, France, Oct. 2014

[7] H. Sculzrinne, A. Rao and R. Lanphier, Real Time Streaming Protocol (RTSP), IETF RFC 2326, April 1998

[8] Information technology – Dynamic adaptive streaming over HTTP (DASH) – Part1: Media presentation description and segment formats, ISO/IEC 23009-1:2014

[9] F. Bossen, D. Flynn, K. Sharmann, K. Suehring and T. Suzuki, *HEVC Refernce Software for Format Range Extensions Profiles,* JCTVC-S1011, Strasbourg, France, Oct. 2014

[10] F. Bossen, *Common test conditions and softwarereference configurations*, JCTVC-J1100, Stockholm, Sweden, July 2012

[11] G. Bjøntegaard, *Calculation of Average PSNR Differences Between RDCurves*, ITU-T SG16 Q.6, Document VCEG-M33, Austin, TX, Apr. 2001.

[12] G. Bjøntegaard, *Improvements of the BD-PSNR Model*, ITU-T SG16 Q.6, Document VCEG-AI11, Berlin, Germany, Jul. 2008.