|  |  |
| --- | --- |
| **Joint Collaborative Team on Video Coding (JCT-VC)**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11**  32nd Meeting: Ljubljana, SI, 12–18 July 2018 | Document: JCTVC-AF0026-v1 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | Faster HEVC Encoder for Video Surveillance and Screen Content Applications | | |
| *Status:* | Input document to JCT-VC | | |
| *Purpose:* | Non-normative proposal | | |
| *Author(s) or Contact(s):* | Mohammed S. Sayed Egypt-Japan University of Science and Technology Egypt | Tel: Email: | Tel: +201005171734  E-mail: [mohammed.sayed@ejust.edu.eg](mailto:mohammed.sayed@ejust.edu.eg) |
| *Source:* | Egypt-Japan University of Science and Technology ; National Telecommunication Regulatory Authority (NTRA) (Egypt) | | |

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ITU Logo | INTERNATIONAL TELECOMMUNICATION UNION  **TELECOMMUNICATION STANDARDIZATION SECTOR**  STUDY PERIOD 2017-2020 | | SG16-C257-R2 | | |
|  |  | | **STUDY GROUP 16** | | |
|  |  | | **Original: English** | | |
| **Question(s):** | | Q6/16 | | Ljubljana, 9-20 July 2018 | |
| **CONTRIBUTION** | | | | | |
| **Source:** | | Egypt-Japan University of Science and Technology ; National Telecommunication Regulatory Authority (NTRA) (Egypt) | | | |
| **Title:** | | Faster HEVC Encoder for Video Surveillance and Screen Content Applications | | | |
| **Purpose:** | | Proposal | | | |
| **Contact:** | | Mohammed S. Sayed Egypt-Japan University of Science and Technology Egypt | | | Tel: +201005171734 E-mail: [mohammed.sayed@ejust.edu.eg](mailto:mohammed.sayed@ejust.edu.eg) |

|  |  |
| --- | --- |
| **Keywords:** | screen content coding, video surveillance, early termination algorithms, acceleration |
| **Abstract:** | The exhaustive search performed by the encoder makes the encoding process very slow. Screen Content Coding and Video Surveillance videos are highly used nowadays and this kind of videos require very fast encoding for real-time transmission. In this proposal we propose three early termination algorithms that accelerate the encoding process with a negligible effect on the compression efficiency. The first algorithm makes use of stable back ground and view that exist for a long period of time. The second algorithm make use of identical blocks that exist frequently in the mentioned videos. The last algorithm makes use of wide spaces of one colour that happen in screen content videos. The results show that our algorithms help in accelerating the encoding of video surveillance and screen content applications by an average of 29.06% with a negligible effect on the compression efficiency. |

1. **Introduction**

The evolution of video compression standard HEVC towards higher compression efficiency comes at the cost of excessive increasing of the encoding complexity. The encoding complexity exists because of the exhaustive search that the encoder performs to find the suitable partitioning for the code tree unit (CTU) and the optimum prediction for each block to show the best compression efficiency. Accordingly, the encoding time becomes the bottleneck in transferring the standard to market specially for real time transmission applications.

Researchers work on reducing the complexity of the encoder in order to accelerate the encoding process. Most of the literature work succeeded in making the encoder faster but, on the other hand, they affected the compression efficiency badly.

Video surveillance and screen content applications are highly used nowadays. Specially, with the increase of video communication applications like video calls, video conferences, e-learning, and remotely used computers. These kind of applications share the high demand of real-time transmission which, in turn, means that they need fast encoding/decoding devices.

In this work, we propose three algorithms aiming to accelerate the encoding process for video surveillance and screen content applications. We make use of the features of these kinds of videos such as wide areas of one colour and stable backgrounds for long periods of time. The results show that our algorithms help in accelerating the encoding of video surveillance and screen content applications by an average of 29.06% with a negligible effect on the compression efficiency.

1. **Background Based Block Encoding Termination**

Videos however they are camera content or screen content are mostly having areas of the frames as background areas [1]. Most of these areas of background are stationary objects or views. Blocks in these areas are encoded by inter-prediction with motion vector (MV) with value of 0.

For many video applications, such as video surveillance and video conference the back ground areas still for a lot of frames with no change [2,3].

A good indicator for the optimum prediction mode is a flag called coded block flag (CBF). When a block in predicted and the transformed error coefficients between the original block and the predicted block are zeros, the CBF is set to 0 which means that the current prediction mode is sufficient. CBF is proposed in [4] and included in HM5.0. It is checked for the luma and two chroma blocks. If CBF is zero for the three blocks, the rest of prediction options in current depth are skipped.

Based on the above two properties, we propose a new algorithm for early termination the block encoding process. As shown in the flowchart at Fig.1, we check the block at 10 sequential frames if it is background (MV=0) and if that encoding as background is optimum (CBF=0). If that strict condition is met, we set a flag for that block as a background block (BG=0). At next frames, before going in the normal search, we check the CBF for the block at (MV=0). If CBF is still 0, the encoding search is terminated because the inter-prediction with (MV=0) is optimum for that block. Otherwise, BG flag is raised to 1 and the normal coding search is applied to the block.

1. **Memory based Block encoding termination**

Intra-prediction uses the block and its neighbour pixels to generate intra-mode that relates the block to its neighbour blocks. When a block is encoded by intra-prediction, it is checked for 35 modes by rough mode decision then a set of modes are candidate to be checked by rate distortion optimization (RDO). At RDO process, if the CBF flag is 0 the encoding process is terminated.

Actually, if we have two identical blocks with identical neighbour pixels, they will have the same intra-mode. We use this fact in proposing a new termination algorithm for the encoding process.

Screen videos which presents window or educational presentations contain a large areas of the same colour. Intra-Block Copy (IBC) is adopted in the HEVC-SCC standard for that issue. Therefore, it is expected that there is a lot of identical blocks. We target this kind of applications by our algorithm.

We focus on the blocks that when encoded by an intra-mode their CBF is 0 which means that it has the optimum intra-prediction mode. If the CBF is 0 for a block, we save it with its neighbour pixels and the encoded mode in a memory. Later, we compare each new block and its neighbour pixels with those that are saved in the memory. If they are identical, we choose the optimum saved mode for that block and terminate the search. Otherwise, the block is processed normally searching for the optimum mode of it. The flowchart of the proposed algorithm is shown in Fig.2.

Increasing the number of blocks in the memory will lead to increasing the encoding time. Therefore, we make block sampling and store the samples as an indicator for the block. For the top and left reference samples, we take a sample every four pixels. For the block, we take a sample every four pixels in x direction and every eight pixels in y direction. The sampling for 8x8 block is shown in Fig. 3. Further, we make the memory to reset every 2 frames for two reasons; Firstly, saving memory and keeping the benefit of decreasing the encoding time that is achieved by our algorithm. Secondly, the algorithm would be inefficient if the scene changed. Furthermore, we don’t apply our algorithm on 4x4 blocks because the very high opportunity that their predicted blocks will have CBF=0 so that the algorithm would be inefficient due the large number of performed comparisons and the large amount of used area.

In HEVC, the search for the best prediction starts by inter-prediction then intra-prediction because of the fast time of encoding by inter-prediction. In intra-prediction, at the case of finding an identical block at the memory list, the proposed algorithm makes intra-prediction is sufficient to encode the block with no need to make inter-prediction search. Therefore, we modify our algorithm by making this comparison before inter-prediction search. If an identical block is found, the search is terminated before both inter- and intra-prediction search as shown in the flowchart at Fig.4.

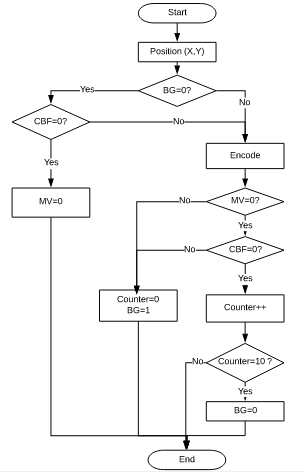


Fig. 1 Background based Block encoding termination

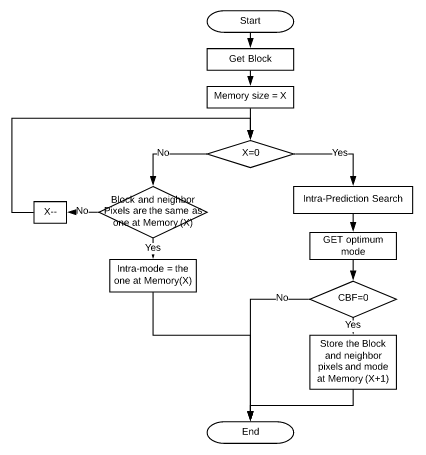


Fig. 2 Memory based Block encoding termination

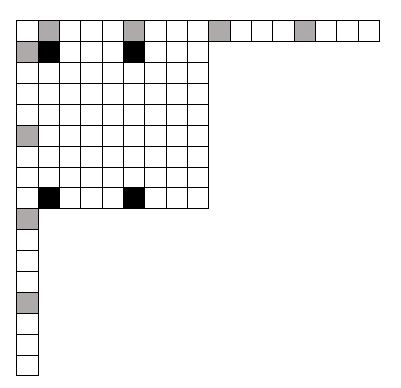


Fig. 3 Samples stored for 8x8 block

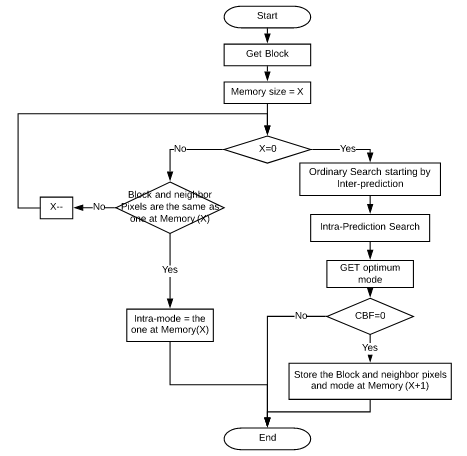


Fig. 4 Modified Memory based Block encoding termination

1. **Feature Based Block Partitioning Termination**

In HEVC, portioning decision depends on the cost of encoding a block and the cost of encoding its sub-blocks in terms of bitrate and distortion. If the cost of encoding the parent block is higher, it is partitioned and if it is lower it is not partitioned.

It is noticed that blocks having a complex texture mostly produces higher costs than the cost of their sub-blocks and they are partitioned [5]. On the other hand, blocks having a simple texture tends to be encoded as a prediction block and not partitioned.

Many researches used the block texture for fast coding unit partitioning. Most of the literature work uses variance to indicate the block texture complexity [6]. Others uses edge detection [7].

The main drawback of using block texture complexity is that sometimes blocks have large variance or a lot of edges and are assumed as complex blocks and have to be partitioned but these blocks can be inter-predicted or intra-block copied from other blocks. Therefore, hard decision of partitioning is not suitable in such cases. Researches that use that technique show a loss in BD-rate. For example, in [6] the loss in BD-rate is 5%.

In our work, we avoid the hard decision of partitioning and keep the hard decision of not portioning. Our proposed idea is based on that the simple block that has no details at all is mostly a part surface area so that it is predicted as a copy of another block either in the same frame by intra-block copy or in another frame by inter-prediction. However, if there is no block with the same size of it and can’t be copied, it is predicted as DC or planner intra-mode. This simple blocks with no details usually exist in screen content videos especially in remote presentations videos. Our algorithm state that once the block variance is smaller than a threshold the search in smaller sizes is terminated. We use a training sequence to test different values of thresholds. Our target is to make the termination of search which if it continued, the rest of search would be useless. In other words, we want to keep the BD-rate with no change by keeping the partitioning after using our algorithm the same as the standard partitioning and gaining time saving from the termination. The best values of thresholds for different block sizes are listed in table 1. As shown, we use a very small thresholds to make sure that the block is a plane one colour.

Table 1 Variance threshold for partitioning decision

|  |  |
| --- | --- |
| Size | Threshold |
| 8x8 | 0.005 |
| 16x16 | 0.005 |
| 32x32 | 0.0007 |
| 64x64 | 0 |

1. **Experimental results**

We integrate the three proposed algorithms with HEVC standard as shown at the flowchart in Fig. 5. The test model HM16.18 [8] is used as an anchor for all experiments. To test the performance of the proposed algorithms, the common test conditions Low Delay, which are specified in [9], are applied. The time saving is calculated by Eq. 1 while the cost of both bit-rate and PSNR will be calculated in terms of BD-rate and BD-PSNR [10]. The BD-rate and BD-PSNR results, shown at Table 2, are calculated at four quantization parameter (QP) 22, 27, 32, and 37.

(1)

We encoded 3 categories of videos by the original HM source and the integrated HM source with our proposed algorithms. The first category is the screen content videos (SlideShow and SlideEditing) that shows slide presentations. The second category is video surveillance videos (FourPeople, Johnny, and KristenAndSara) that have a fixed background view. The third category is normal camera captured videos.

As shown in Table 2, for screen content and video surveillance videos, our algorithms show an average acceleration in the encoding time by 29.06% with a negligible effect on either the BD-rate (increased by 0.21%) or the BD-PSNR (decreased by 0.003 dB). On the other hand, our algorithms could not achieve that high performance for other camera captured videos as shown in Table 3. This is because the conditions that our algorithms are based on are rarely happening at this kind of videos. Note that, if the encoder uses only one reference frame instead of three reference frames, the acceleration in the encoding time, for screen content and video surveillance videos, will be 21.2% while the degradation in the BD-rate will be 0.09% and the BD-PSNR will be 0.0001 dB.

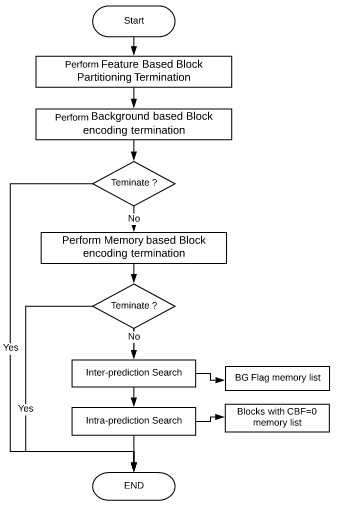


Fig. 5 Encoding process of HEVC integrated with our proposed algorithms

Table 2 Results of the proposed algorithm with screen content and video surveillance videos

|  |  |  |  |
| --- | --- | --- | --- |
|  | **(%)** | **BD-Rate (%)** | **BD-PSNR (dB)** |
| FourPeople | 19.42 | 0.126 | -0.005 |
| Johnny\_Original | 31.06 | 0.454 | -0.008 |
| KristenAndSara | 22.26 | 0.405 | -0.011 |
| SlideEditing | 32.68 | -0.176 | 0.024 |
| SlideShow | 39.91 | 0.241 | -0.017 |
| **Avg.** | 29.06 | **0.210** | **-0.003** |

Table 3 Results of the proposed algorithm with camera captured videos

|  |  |  |  |
| --- | --- | --- | --- |
|  | **(%)** | **BD-Rate (%)** | **BD-PSNR (dB)** |
| BQTerrace | 0.40 | 0.065 | -0.001 |
| BasketballDrive | -2.60 | 0.042 | -0.001 |
| Cactus | 3.51 | 0.085 | -0.003 |
| Kimono | 0.68 | 0.048 | -0.002 |
| ParkScene | -0.12 | 0.002 | 0.000 |
| BQMall | 1.08 | 0.084 | -0.003 |
| BasketballDrill | -0.11 | 0.002 | 0.000 |
| PartyScene | 0.73 | -0.006 | 0.000 |
| RaceHorsesC | 1.14 | 0.074 | -0.003 |
| BQSquare | -7.65 | 0.014 | -0.001 |
| BasketballPass | 3.26 | -0.046 | 0.002 |
| BlowingBubbles | -1.23 | 0.033 | -0.001 |
| RaceHorses | -2.16 | -0.062 | 0.002 |
| BasketballDrillText | -0.44 | 0.028 | -0.001 |
| ChinaSpeed | 3.27 | 0.077 | -0.004 |
| **Avg.** | **-0.02** | **0.029** | **-0.001** |

1. **References**

[1] Fangdong Chen, Houqiang Li,  Li Li,  Dong Liu, and Feng Wu, “Block-Composed Background Reference for High Efficiency Video Coding”, IEEE Transactions on Circuits and Systems for Video Technology, v[ol. 27, no 12](https://ieeexplore.ieee.org/xpl/tocresult.jsp?isnumber=8186326),

[2] Li Ma, Honggang Qi, Siyu Zhu, and Siwei Ma, “A fast background model based surveillance video coding in HEVC”, Visual Communications and Image Processing (VCIP), 2014

[3] Fangdong Chen, Dong Liu, Houqiang Li, and Feng Wu, “Fast encoding of surveillance videos based on HEVC”, Visual Communications and Image Processing (VCIP) 2017

[4] R. H. Gweon et al., “Early termination of CU encoding to reduce HEVC complexity,” JCT-VC document, JCTVC-F045, Jul. 2011.

[5] Maher Abdelrasoul, Mohammed S. Sayed, and  Victor Goulart, “Diagonal-based fast intra-mode decision algorithm for HEVC”, IET Image Processing, [vol. 11 no. 10](https://ieeexplore.ieee.org/xpl/tocresult.jsp?isnumber=8068786).

[6] Nicolas Dhollande,  Xavier Ducloux, Olivier Le Meur, and [Christine Guillemot](https://ieeexplore.ieee.org/search/searchresult.jsp?searchWithin=%22Authors%22:.QT.Christine%20Guillemot.QT.&newsearch=true), “Fast block partitioning method in HEVC Intra coding for UHD video”, IEEE 5th International Conference on Consumer Electronics Berlin (ICCE-Berlin), 2015

[7] Yih-ChuanLin, and Jian-ChengLai, “Feature-based fast coding unit partition algorithm for high efficiency video coding”, Journal of Applied Research and Technology, [vol. 13, no. 2](https://www.sciencedirect.com/science/journal/16656423/13/2), 2015, Pages 205-219

[8] https://hevc.hhi.fraunhofer.de/HM-doc/

[9] Frank Bossen, “Common test conditions and software reference configurations”, JCTVC-L1100, 12th Meeting of Joint Collaborative Team on Video Coding (JCT-VC), Geneva, CH 14-23 January 2013.

[10] G. Bjøntegaard, “Calculation of average PSNR differences between RD curves”. VCEG-M33, April 2001.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_