FAST INTRA MODE DECISION AND BLOCK MATCHING FOR HEVC SCREEN CONTENT COMPRESSION

Hao Zhang¹, Qiao Zhou¹, Ningning Shi², Feng Yang³, Xin Feng³, Zhan Ma²

¹Central South University, ²Nanjing University, ³Chongqing University of Technology

ABSTRACT

Screen content coding (SCC) is the latest extension of the High-Efficiency Video Coding (HEVC) aiming to improve the compression efficiency of screen content video. With newly developed tools such as intra block copy (IntraBC) and palette (PLT) mode, SCC has been able to compress the desk-top screens more efficiently but with significant complexity increase. In this paper, we improve the intra prediction from two aspects. Firstly, by leveraging the temporal correlation among coding units (CU), we develop a fast CU depth prediction scheme. Furthermore, adaptive search step is employed for further speed up of the time-consuming block matching in IntraBC. The overall encoding time is reduced by about 39% and 35% for the All Intra (AI) lossy and lossless encoding scenarios with negligible quality loss under the SCC common test condition.

Index Terms— Intra mode, intra block copy, screen content coding, HEVC

1. INTRODUCTION

Screen content video, often containing mixed content consisting of natural video, text, graphics, image, video, etc., has been widely embedded in various applications in recent years. Among them, wireless displays, shared screen collaboration, virtual desktop interface etc., are generating high volume screen content video in their applications. Therefore, compressing screen content video efficiently is important to save bandwidth and improve communication quality. As we all know, conventional video coding standards, such as the H.264/AVC [1], HEVC [2], have shown the impressive performance on camera captured natural video content compression. However, screen content is different from the natural video and it is inevitable to develop the new tools to efficiently exploit the correlation in screen images exemplified in Figure 1 [3]. Towards this goal, besides the efforts devoted from the academia, Joint Collaborative Team on Video Coding (JCT-VC) has launched screen content coding standardization activities as the extension of the well-designed HEVC on April 2014 [4], after evaluating multiple technical responses of SCC Call-for-Proposal (CfP).



Fig. 1. Illustration of typical screen content image in daily practice.

Compared with the traditional nature video, screen content has different characteristics as it is often temporally static with sharp edges, few distinct colors, etc. Hence, HEVC based SCC introduces the intra block copy (IntraBC) [5], color palette (PLT) mode [6], adaptive motion vector resolution [7], color transforms [8, 9], etc, to massively leverage the content correlation and remove the image redundancy. Among them, IntraBC and PLT modes are two of them offering significant coding performance but with noticeable complexity overhead. IntraBC is a block matching technique, which enables inter-alike block estimation and compensation technology using fixed block size for better coding efficiency [5, 10]. Instead of searching the reference in previously (temporally) reconstructed frame, it searches the reconstructed region in current frame and carries the block vector and compensation residual to the decoder [11]. Color palette mode was found to be another attractive scheme for screen content coding, which applied the color mapping using derived table or palette to represent the pixels in screen content with fewer values [6, 12, 13]. Blocks with high dynamic range (i.e., 0 - 255 for 8-bit pixel) are translated to the index map with limited dynamic range that is easier to be compressed.

As we know, HEVC provides 50% bit rate reduction in comparison to the H.264/AVC at same quality, with more than 5x complexity increase. Moreover, SCC increases the encoder complexity even more by adopting these new tools. Therefore, it is inevitably to study the fast algorithms to speed

up the SCC encoder for potential applications in practices. In this paper, we proposed a novel and fast intra mode decision method based on the background detection to skip the unnecessary calculation so as to reduce the encoder complexity. Meanwhile, we adaptively adjust the searching step size for block matching to boost IntraBC block estimation.

The rest of this paper is organized as follows: Section 2 presents a brief literature review of the fast mode decision algorithms for screen content coding. The proposed fast intra mode decision and block matching scheme are discussed in Section 3 followed by the extensive experiments in Section 4 to demonstrate the effectiveness of the proposed solution. Section 5 concludes the paper with a brief summary in this work.

2. FAST INTRA MODE DECISION AND BLOCK MATCHING FOR SCC: A BRIEF REVIEW

Fast mode decision for screen content coding has drawn more and more attention in the past two years. A number of mode decision algorithms for SCC have been proposed, such as fast coding tree unit (CTU) partition decision etc. These algorithms are briefly summarized in the following sections.

Zhang *et. al.* proposed an early CU split mode scheme based on entropy estimation to reduce the complexity of the encode process [14]. Furthermore, to avoid introduced quality loss for some sequences, the number coding bits is combined with entropy estimation to improve the efficiency. The presented algorithm can achieve 32% encoding time reduction on average with 0.8% bitrate increase.

Kwon *et. al.* presented a technique to early skip IntraBC block matching based on regular intra mode cost, horizontal and vertical activities for 8x8 CU and whether 16x16 CU is coded as IntraBC or not [15]. It can save about 21% - 24% encoding time for All Intra coding. The techniques proposed in this work have been adopted and implemented in SCC reference software SCM.

In the version of SCM 3.0, other than the techniques proposed in [15], a few fast block matching algorithms are implemented and enabled by default [16, 17, 18]. Firstly, a few predictors (e.g., the motion vectors of the collocated and the spatial neighbors) are checked. Then, a so called 1-D search is performed, which only search vertically and horizontally directions. For the 8x8 CU, additional 2-D search is performed. For the 8x8 partition, a hash based full frame search is used. In the hash based search, a hash value is calculated for both reconstructed 8x8 blocks and the current 8x8 CU. Only those positions where their hash values are identical to that of the current CU are searched. Some early termination mechanisms are also implemented to further speed up the searching process.

Even with the fast algorithms aforementioned, the coding complexity of SCC is still too high. To further improve coding speed and facilitate realtime application of SCC, we propose two techniques: Firstly, sum of absolute differences (SAD) measurement and simple thresholding technique are applied to perform the background estimation. For CUs marked as background block, certain criteria are set to determinine whether to skip the whole intra mode decision process of current depth (except color palette mode). In addition, adaptive step size increment is applied to further speed up the 2-D block matching process of IntraBC mode. In this way, more encoder complexity reduction is achieved on top of the exisiting techniques.

3. PROPOSED FAST INTRA MODE DECISION AND BLOCK MATCHING SCHEMES

Besides the normal spatial intra modes [2, 5], SCC also includes the additional IntraBC and color palette modes. We present a novel approach to reduce the computational burden required by the intra mode decision. It consists of background estimation based early termination.

3.1. Fast Intra Mode Decision Based on Background Detection

CU depth prediction is a widely used technique in fast inter or intra mode decision for HEVC [19]. For example, the CU depth of the current CU is found to be related to the depths of spatial and collocated CUs [20]. This has been exploited to predict the possible range of the current CU depth, i.e., the CU depths outside of the predicted range would not be checked at all and so encoding complexity is saved. However, we are not aware of any work exploiting the temporal correlation to speed up intra prediction for HEVC, nor SCC.

In this paper, we apply the background detection based CU depth prediction. Based on our analysis, we find that when two collocated CUs are very similar, e.g., they are in the background (or stationary) region, their CU depths are closely related. It is well known that image can be simply divided into foreground region and background region from the structural characteristics of the image. Usually, background or stationary block is used to reflected the still regions without movement, where the motion vectors are zero. For typical sequences in SCC applications (e.g., SlidesShow in Figure 1), a large portion of the input image is the background. Therefore, the background detection should be used effectively for encoding time saving. For All Intra coding, the decoder is unable to obtain the temporal information if the previous frames are IDRs. In the encoder side, however, the temporal information is still available for fast mode decision. As revealed later in simulations, with such technique, considerable encoding time could be saved by skipping CU mode decision at certain depths.

To estimate the background region, as the previous work [21, 22] has done, we calculate the sum of absolute differences (SAD) between current CU and the collocated CU in

previous encoded frame as below:

$$SAD = \sum_{i} \sum_{j} |P_c(i,j) - P_p(i,j)|, \qquad (1)$$

where P_c and P_p are the pixel values of (i, j)-th position in current CU and collocated CU in previous encoded frame. |.| indicates the absolute value.

Moreover, a threshold TH = $\gamma * \alpha$ is defined. If SAD < TH, the current CU is treated as a background or stationary block (i.e., IsBkg = TRUE); oppositely, the current CU is foreground block (i.e., IsBkg = FALSE). Here α is number of pixels in the current CU; γ is a user defined controlling factor, which is set as 2 as suggested by [21] (i.e., $\gamma = 2$).

Our contribution here is to speed up the mode decision process if a CU block is decided as a stationary background block, where its partition is closely related to the collocated CU. For this purpose, we compare the depth of current CU (D_c) and the depth of collocated CU in previous encoded frame (D_p) . We use IsBkg(d) to represent whether the CU at depth d is a stationary background block or not. To avoid the accumulation of prediction error, we switch off the fast algorithm every ten frames. The intra mode decision process in the current depth would be early skipped (except for the PLT mode) if all the conditions in the following are satisfied:

- 1. The frame number is not divisible by 10, i.e., FrameNum % $10 \neq 0$.
- 2. The collocated CU has chosen either normal intra mode or IntraBC mode as its best mode.
- 3. If the current CU is determined to be a stationary background block and the collocated CU has a larger depth, i.e., IsBkg (D_c) = TRUE and $D_p > D_c$.
- 4. If the collocated CU has a smaller depth $D_p < D_c$ and IsBkg (D_p) = TRUE.

3.2. Fast Block Matching Based on Adaptive Searching Step Size Adjustment

In some screen content sequences, e.g., videos with embedded texts, conventional fast motion estimation algorithms are not as effectively as before in natural videos. This is because in screen content videos, *true* motion might not located at an unique position (i.e., same "word" or "characters" in the text). Furthermore, it is not necessarily true that the further from the true position, the higher the match errors. Therefore, new algorithms should be re-designed for screen content videos instead of applying the existing fast motion search inherited for natural camera captured videos.

In this paper, we propose to adjust the step sizes adaptively for the 2D search process, where all the positions are divided into three non-overlapping and each part is searched individually. The search position x = x + dx, where x represents the horizontal search position and dx denotes the original searching step size. The original value of dx is set to one or two, depending on the search positions. After each step, the SAD value is compared with some preset threshold to early terminate the search process if the intermediate result is considered good enough. Readers could referred to the reference software for more details about the original 2D block matching algorithm. In this paper, we would improve the above 2D search process when encountering "bad" enough SAD results. Formally, if SAD > α SAD₀, the searching step size dx would be adjusted accordingly. Here SAD₀ represents the Mth smallest SAD value at the end of the SAD queue, which is kept for chroma IntraBC search refinement. The main idea of the proposed fast search is to dynamically adjust dx to dx'depending on the value of α by the rules detailed in the following:

In the case of dx = 1:

If SAD > 2 SAD₀, dx' = 2.
 If SAD > 4 SAD₀, dx' = 4.
 In the case of dx = 2:
 If SAD > 2 SAD₀, dx' = 4.
 If SAD > 4 SAD₀, dx' = 6.

The intuition behind is that when the SAD value is sufficient large, the search step sizes can be set to a even larger value to skip unnecessary search positions. This technique is quite useful when the current CU belongs to the high texture region (e.g., text) and the search area is a smooth region (in the reference buffer).

Table 1. Performance Evaluation on All Intra Encoding Configuration for Lossy Screen Content Coding of the proposed fast mode decision against SCM 3.0

All Intra							
		BD-Rate Increase					
		G-Y	B-U	R-V			
Lossy	TGM-G	1.3%	1.5%	1.5%			
	MIX-G	1.0%	1.1%	1.2%			
	AMT-G	0.1%	0.2%	0.2%			
	CAC-G	0.0%	0.0%	0.0%			
	TGM-Y	1.4%	1.4%	1.4%			
	MIX-Y	1.1%	1.1%	1.1%			
	AMT-Y	0.3%	0.4%	0.3%			
	CAC-Y	0.0%	0.0%	0.0%			
	Overall	0.7%	0.7%	0.7%			
	Enc Time[%]		61%				

4. EXPERIMENTAL RESULTS

To evaluate the proposed fast intra mode decision methods for SCC, HEVC reference software SCM 3.0 is modified to incorporate the proposed algorithm. Both the lossy and lossless scenarios are performed using the "All Intra(AI)" configurations (due to the cause that IntraBC and PLT are primarily developed for intra coding). The test sequences are the screen content sequences selected by the experts in Joint Collaborative Team on Video Coding group [23]. These sequences represent popular and typical screen content application scenarios consisting of four categories:

- TGM: text and graphics with motion;
- AMT: animation content;
- MIX: mixed content;
- CAC: camera captured.

Moreover, we note RGB format sequence of TGM category as TGM-G while YCbCr format as TGM-Y, which also applied to other content categories. Experimental results of BD-Rate performance and encoder time reduction for AI lossless and lossy encoding are shown in Table 1 and Table 2.

 Table 2. Performance Evaluation on AI Encoding Configuration for Lossless Screen Content Coding of the proposed fast mode decision against SCM 3.0

All Intra							
		Bit-Rate Increase					
		Total	Average	Min	Max		
Lossless	TGM-G	0.7%	0.9%	0.1%	2.9%		
	MIX-G	0.4%	0.4%	0.2%	0.5%		
	AMT-G	0.0%	0.0%	0.0%	0.0%		
	CAC-G	0.0%	0.0%	0.0%	0.0%		
	TGM-Y	0.4%	0.4%	0.1%	0.8%		
	MIX-Y	0.1%	0.1%	0.1%	0.1%		
	AMT-Y	0.0%	0.0%	0.0%	0.0%		
	CAC-Y	0.0%	0.0%	0.0%	0.0%		
	Overall	0.2%	0.2%	0.1%	0.5%		
	Enc Time[%]	65%					

In total, the proposed solution achieves 39% encoding time reduction and 0.7% BD-rate increase for AI lossy case, and 35% encoding time reduction and 0.2% bitrate increase for AI lossless case, respectively. Please note that, the performance enhancement is achieved by comparing with SCM 3.0 where most of the fast algorithms mentioned in the review are implemented. For the random access (RA) case, the encoding time saving is 4% and 2% encoding time reduction respectively. The quality loss is negligible (less than 0.1% BDBR increase). The low delay B (LD-B) case has shown similar performance. Given that IntraBC and palette mode are mainly

developed to improve the intra coding performance, RA and LD-B results are not detailed here.

5. CONCLUSION

In this work, we propose two novel fast algorithms for screen content encoding. The first algorithm exploits the temporal correlation to speed up the intra mode decision process. The second algorithm adopts a fast step size adjustment strategy to improve block matching speed for IntraBC. Compared with the reference software SCM 3.0, the proposed algorithms provide the averaged 39% and 35% encoding time reduction with only 0.7% BD-Rate increase and 0.2% bit-rate increase for AI lossy and lossless cases, respectively. In this work, thresholds are almost all fixed. As for our future work, we would study the adaptive thresholds through online training via available data. Furthermore, only 2D block matching is speed up in this paper. How to reduce the complexity of other time consuming searching modules such as hash based full frame searching is also marked in our to-do-list.

6. ACKNOWLEDGEMENT

Dr. X. Feng's work described in this paper was partially supported by National Natural Science Foundation (NSF) of China (61202348). Dr. Z. Ma is partially supported by the NSF for Young Scholar of Jiangsu Province, China (Grant No. BK20140610). Dr. Z. Ma is the corresponding author of this paper.

7. REFERENCES

- T. Wiegand, G.-J. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 13, no. 7, pp. 560 – 576, July 2003.
- [2] G. J. Sullivan, J. Ohm, W. J. Han, and T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) Standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1649– 1668, 2012.
- [3] Z. Ma, W. Wang, M. Xu, et al., "Advanced Screen Content Coding Using Color Table and Index Map," *IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society*, vol. 23, no. 10, pp. 4399 – 4412, 2014.
- [4] H. Yu et al., "Screen content coding (AHG8)," in *Joint Collaborative Team on Video Coding(JCTVC)*, Incheon, April. 2013.
- [5] J. Sole and S. Liu, "HEVC Screen Content Coding Core Experiment 1 (SCCE1):Intra Block Copying Extensions," in *Joint Collaborative Team on Video Coding(JCTVC), JCTVC-Q1121*, Valencia, ES, March. 2014.
- [6] L. Guo, W. Pu, F. Zou, et al., "Color palette for screen content coding," in 2014 IEEE International Conference on Image Processing (ICIP), Oct. 2014, pp. 5556–5560.
- [7] B. Li, J. Xu, G. J. Sullivan, Y. Zhou, and B. Lin, "Adaptive motion vector resolution for screen content," in *Joint Collaborative Team on Video Coding(JCTVC), JCTVC-S0085*, Strasbourg, FR, Oct. 2014.
- [8] L. Zhang, J. Chen, J. Sole, M. Karczewicz, X. Xiu, Y. He, and Y. Ye, "SCCE5 Test 3.2.1 : In-loop colorspace transform," in *Joint Collaborative Team on Video Coding(JCTVC), JCTVC-R0147*, Jun. 2014.
- [9] L. Zhang, J. Chen, J. Sole, and M. Karczewicz, "AHG8: In-loop color-space transform," in *Joint Collaborative Team on Video Coding(JCTVC), JCTVC-Q0112*, Valencia, ES, Mar. 2014.
- [10] D. K. Kwon and M. Budagavi, "RCE3: Results of test 3.3 on Intra motion compensation," in *Joint Collaborative Team on Video Coding(JCTVC)*, Vienna, AT, July. 2013.
- [11] C. Pang et al., "Non-RCE3:Intra Motion Compensation with 2-D MVs," in *Joint Collaborative Team* on Video Coding(JCTVC), JCTVC-N0256, Vienna, AT, July. 2013.

- [12] X. Guo, "AHG8:Major-color-based screen content coding," in *Joint Collaborative Team on Video Coding(JCTVC), JCTVC-00182*, Geneva, Oct. 2013.
- [13] L. W. Guo, "Palette Mode for Screen Content Coding," in *Joint Collaborative Team on Video Coding(JCTVC),JCTVC-M0323*, Incheon, Apr. 2013.
- [14] M. M. Zhang, Y. H. Guo, and H. H. Bai, "Fast Intra Partition Algorithm for HEVC Screen Content Coding," in 2014 IEEE Visual Communications and Image Processing Conference, Valletta, Dec. 2014, pp. 390 – 393.
- [15] D. K. Kwon et al., "Fast intra block copy (IntraBC) search for HEVC screen content coding," in 2014 IEEE International Symposium on Circuits and Systems (IS-CAS), Melbourne VIC, June. 2014, pp. 9 – 12.
- [16] M. Budagavi and D.-K. Kwon, "AHG8: Video coding using Intra motion compensation," in *Joint Collaborative Team on Video Coding(JCTVC)*, Incheon, KR, April 2013.
- [17] C. Pang, J. Sole, L. Guo, et al., "Non-RCE3: Intra motion compensation with 2-D MVs," in *Joint Collaborative Team on Video Coding(JCTVC)*, Vienna, AT, August 2013.
- [18] W.J. Zhu, W.P. Ding, J.Z. Xu, et al., "Hash-Based Block Matching for Screen Content Coding," *IEEE Transactions on Multimedia*, vol. 17, no. 7, pp. 935–944, July 2015.
- [19] L.Q. Shen et al., "An effective CU size decision method for HEVC encoders," *IEEE Transactions on Multimedia*, vol. 15, no. 2, pp. 465–470, February 2013.
- [20] H. Zhang and Z. Ma, "Fast Intra Mode Decision for High Efficiency Video Coding (HEVC)," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 24, no. 4, pp. 660–668, April 2014.
- [21] P.I. Hosur and K.K. Ma, "Motion vector field adaptive fast motion estimation," in *the 2nd International Conference on Information, Communications and Signal Processing*, Singapore,, December 1999.
- [22] D. Wu, F. Pan, K.P. Lim, S. Wu, et al., "Fast Inter mode Decision in H.264/AVC Video Coding," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 15, no. 6, pp. 953–958, July 2005.
- [23] MPEG N14715, "Joint Call for Proposals for Coding of Screen Content," Jan. 2014.