Fast HEVC Screen Content Coding By Skipping Unnecessary Checking of Intra Block Copy Mode Based on CU Activity and Gradient

Sik-Ho Tsang[1], Wei Kuang[2], Yui-Lam Chan[3], and Wan-Chi Siu[4]
Centre for Signal Processing, Department of Electronic and Information Engineering,
The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong.
E-mail: {[1]sik-ho.tsang, [3]enylchan, [4]enwcsiu}@polyu.edu.hk, [2]wei.kuang@connect.polyu.hk, [3]Tel: +852-27666213

Abstract— **The Intra Block Copy (IntraBC) mode is a very efficient coding tool for the screen content coding (SCC) extension in High Efficiency Video Coding (HEVC) by finding the repeating patterns within the same frame. Yet, it also brings along impractically high computational complexity for SCC, which can be a double of the conventional HEVC, as exhaustive block matching is done within the same frame even though there are already some constraints applied to the IntraBC mode. To reduce the complexity, we propose to skip the unnecessary IntraBC mode checking based on the activity and gradient within the coding unit (CU). With our proposed methods, the increased encoding time compared with the conventional HEVC is reduced from 90.0% to 62.2% on average while the coding efficiency can still be maintained with only negligible bitrate increased.**

I. INTRODUCTION

With the popularity of thin-client devices as well as cloud technology, computer screen sharing applications such as virtual desktop and slideshows sharing have become widespread nowadays [1]. As a result, new challenges on screen content coding (SCC) for limited network bandwidth has emerged as one of the hot research topics in the development of High Efficiency Video Coding (HEVC) [2]. Call for proposal (CfP) [3] was also launched by the Joint Collaborative Team on Video Coding (JCT-VC) of ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q6 in January 2014. SCC has then become an extension of HEVC [4-5]. Related SCC test model [6] has also been established as well.

Screen contents which are mainly generated by computers have different image characteristics from camera captured contents. Camera captured contents have the continuous-tone characteristics, and usually contain camera noise even in smooth regions. On the contrary, screen contents consist of the discontinuous-tone characteristics which have been revealed by [7-8]. Screen contents such as text contain a complex structure including sharp edges with high contrast. And usually they also contain only limited number of colors and sometimes with high contrast between colors. Thereby, the conventional HEVC intra mode [9-10], which uses the neighboring boundary pixels for prediction, cannot handle these kinds of contents well. A palette mode [11-13] was suggested to encode color information and structural information separately by base colors and palette index map.

On the other hand, screen contents have the characteristics of noiselessness that all pixels within a coding unit (CU) can be exactly equal along the horizontal or vertical directions. Sometimes all pixels within the CU are exactly the same. The conventional HEVC intra mode can perfectly predict these CUs without any distortion. In our previous work [14], we proposed to have simple intra prediction (SIP) which is a simplified version of the conventional intra prediction [9-10]. By using SIP, rough mode decision (RMD) [10] and rate distortion optimization (RDO), which include the checking of numerous intra prediction candidates, are skipped for reducing the complexity of the intra mode. However, it only considers the case of all pixels within the CU are exactly the same, but not consider the case of all pixels within a block being exactly equal along the horizontal or vertical directions.

Last but not least, screen contents have the characteristics of repetitiveness. Repeating patterns can be found commonly within the same frame. An intra block copy (IntraBC) mode [15-16] was proposed to perform intra motion estimation (ME) and motion compensation (MC) wherein the exhaustive blocking matching is done within a certain area around the current CU. Full frame hash search were also proposed in [17-18] where block matching is only done for those block candidates which have the same hash value as the current block to be encoded. It is proven that the IntraBC mode can significantly improve the coding efficiency for SCC but with extremely high computational complexity. We also proposed in [19] that IntraBC mode block matching is performed only if the hash values of current block and reference block are the same.

In this paper, we propose to skip the unnecessary checking of IntraBC mode if the CU can be efficiently encoded by the conventional HEVC intra mode. By analyzing the image characteristics within the CU, the CU is encoded by the intra mode only such that the time-consuming IntraBC mode can be skipped. We start by introducing the conventional intra and IntraBC modes. We then proceed to present how we skip the checking of the IntraBC mode based on the activity and gradient within the CU. Finally, experimental results for our proposed approaches are shown followed by conclusions.

The work described in this paper is partially supported by the Centre for Signal Processing, Department of Electronic and Information Engineering, The Hong Kong Polytechnic University, and a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Grant No. PolyU 152016/14E).

II. Intra Mode in HEVC

In an HEVC intra mode [9-10], there are 1 DC, 1 planar, and 33 directional predictions which form 35 intra predictions in total. In addition, 2N×2N and N×N predictions are supported for each CU size in order for more flexible intra prediction. For each pixel in the current CU, it is predicted by boundary pixels from the neighboring CUs.

To reduce the computational complexity of checking all 35 intra prediction candidates by rate distortion optimization (RDO), rough mode decision (RMD) [10] is applied to select a subset of prediction candidates with lower costs by estimating the sum of absolute differences of Hadamard transformed coefficients. Then RDO is performed for that subset of prediction candidates to choose the optimal one.

III. Intra Block Copy Mode in HEVC

For the intra block copy (IntraBC) mode, block matching in inter-frame [20-24] is done within the same frame during intra ME and MC. If IntraBC is used by one particular CU, each prediction unit (PU) within the CU is encoded with a block vector (BV), as well as the residual signal of that CU. To perform block matching for one searching point, sum of absolute difference (SAD) between the CU candidate and the current CU, $SAD(x,y)$, as well as the cost of BV, $BVCost(x,y)$, are estimated to be the IntraBC cost, $IBCCost(x,y)$ as a total cost in the following:

$$IBCCost(x,y) = SAD(x,y) + BVCost(x,y) \tag{1}$$

where (x,y) is the position of a CU candidate. So SAD and the BV cost are estimated once for each searching point. The one with minimum cost in (1) is selected as the optimal solution within the IntraBC mode. For the BV cost estimation, it occupies very small amount of time compared with the SAD estimation as it is just simply estimated by table look-up plus few operations. The complexity of block matching is mainly come from the SAD estimation. The drawback is that the computational complexity for the IntraBC mode is impractically high due to the exhaustive block matching. Thus several constraints for balancing the coding efficiency and computational complexity are suggested in [15] which has been implemented in SCM-4.1. First, BVs have to be integer-pel. There is no fractional-pel ME and MC but only integer-pel ME and MC. With this restriction, an interpolation process can then be skipped. Second, local search is recommended that only pre-defined area is searched instead of full frame search. Last, the IntraBC mode is only applied for small CU sizes of 16×16 and 8×8 as repeating patterns usually appeared for small CU size rather than large CU size.

For 16×16 CU, only $2N\times2N$ PU with full vertical and horizontal searches are performed. It is due to the fact that large CU size tends to have fewer repeating patterns found within the same frame.

For 8×8 CU, if it is $N\times2N$ PU, only full vertical and full horizontal searches are performed. If it is $2N\times2N$ or $2N\times N$, local vertical, local horizontal and 2D searches within the left and current CTUs are performed. The 2D search for $2N\times2N$ is not performed if the CU activity, Act, is smaller than a pre-defined threshold TH_1 [16] while the 2D search for $2N\times N$ is not performed if CU activity and the best RD cost so far are smaller than the corresponding pre-defined thresholds TH_1 and TH_2 where the CU activity, Act, is estimated as follows:

$$\begin{gathered} Act_H = \sum_{j=0}^{7}\sum_{i=1}^{7} p(i,j) - p(i-1,j) \\ Act_V = \sum_{j=1}^{7}\sum_{i=0}^{7} p(i,j) - p(i,j-1) \\ Act = min(Act_H, Act_V) \end{gathered} \tag{2}$$

And the thresholds TH_1 and TH_2 are set in the following:

$$\begin{gathered} TH_1 = 168 \times (1 \ll BitDepth - 8) \\ TH_2 = max(66 \times \lambda, 800) \end{gathered} \tag{3}$$

where $p(i,j)$ is the luminance component of the pixel at the corresponding position (i,j) within the CU, $min(a,b)$ is the operation to choose the minimum value of a and b, $\ll$ represents the left shift operation, $BitDepth$ is the bit depth of the pixel, $max(a,b)$ is the operation to choose the maximum value of a and b, and λ is the Lagrange multiplier which is quantization parameter (QP) adaptive. Fig. 1 illustrate the full vertical and horizontal searches, local vertical and horizontal searches, and 2D search.

Besides, there is a full-frame hash search for 8×8 CU, block matching is only done for those CU candidates which have the same hash value as the current CU to be encoded wherein the hash value is composed of four DC values of 4×4 blocks as well as the gradient value of whole 8×8 CU [17]. It is only performed for 2N×2N PU with the consideration of the tradeoff between computational complexity and coding performance.

IV. Proposed Approaches

The constraints and fast approach proposed in [15-16] successfully reduce the computational complexity brought by the IntraBC mode. [17] offered a fast full-frame hash search in which block matching is only done for those CU candidates which have the same hash value as the current CU. Nevertheless, there are still rooms to improve the encoding speed. For instance, CUs with all pixels being exactly equal along the horizontal or vertical directions, or even all pixels being exactly the same, can skip the checking of the IntraBC mode as the intra mode with either horizontal, vertical or DC prediction has been already capable of predicting these kinds of CUs without any distortion. In addition, if a CU is a low-gradient CU which contains no shape edges, the intra mode can efficiently encode it. With these motives, we propose to skip the unnecessary the IntraBC mode checking based on the activity and the gradient of the CU.

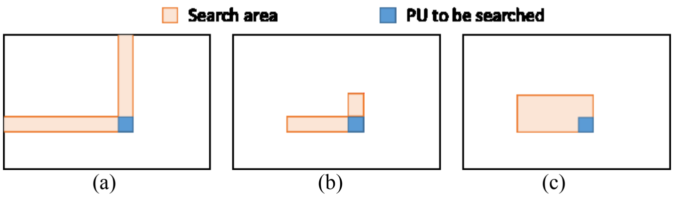

Fig. 1 Illustrations of (a) full vertical and horizontal searches, (b) local vertical and horizontal searches, and (c) 2D search.

A. Skipping IntraBC for CUs with Zero Activity

As aforementioned, the CU activity, estimated in (2), decides whether the 2D search (Fig. 1(c)) is performed based on the thresholds in (3). We further make use of the CU activity in (2) to skip the entire IntraBC mode checking including the full vertical and horizontal searches, local vertical and horizontal searches, 2D search, as well as the hash search. Fig. 2 illustrates an example of a screen content sequence *sc_programming* with areas highlighted by red color. In these areas, each row or each column of pixels are exactly the same, or all pixels are exactly the same. To analyze the CU activity of various CU sizes, we conduct an analysis to collect the statistics regarding to the number of CUs with either the horizontal activity, Act_H, and vertical activity, Act_V equals to zero using the screen content sequences recommended in the Common Test Conditions (CTC) for SCC [25], and it is tabulated in TABLE I. It is noted that we only have the results for the CUs with sizes 16×16 and 8×8 because block matching is only done for 16×16 and 8×8 CU as mentioned in previous section. It is observed that there are about 33.7% and 48.1% of CUs which have zero CU activity which are very high portions. If either Act_H or Act_V is equal to zero, it means that the all pixels within the CU are exactly equal along the horizontal or vertical directions correspondingly. To simplify the operation in (2), we propose to check whether Act_H is zero as follows:

$$p(i,j) = p(0,j) \quad \forall p(i,j) \in P \tag{4}$$

where P is the set of pixels within the CU. For Act_V:

$$p(i,j) = p(i,0) \quad \forall p(i,j) \in P \tag{5}$$

If both (4) and (5) are satisfied, it means that the all pixels within the CU are exactly the same. In these cases, an intra mode with horizontal, vertical or DC prediction can efficiently encode these CUs. Hence, the entire IntraBC mode checking would be skipped. Besides, the RMD and RDO processes [9-10] mentioned in Section II can also be skipped as like in [14]. For instance, if a CU satisfies the condition in (4), horizontal activity, Act_H, is zero and all pixels are exactly equal along the horizontal direction, only most probable modes (MPMs) prediction, plus horizontal prediction if it is not included in MPMs, are checked within the conventional intra mode. RMD and RDO in the intra mode, as well as the entire IntraBC mode checking are skipped.

B. Skipping IntraBC Mode for CUs with Low Gradient

If the CU is low-gradient that there is no shape edges, we can consider that this particular CU exhibits the continuous-tone characteristics. To classify the CU as a low-gradient CU, we check the number of high-gradient pixels within the CU. Thereby, a pixel is defined as a high-gradient pixel if the luminance difference between itself and one of the neighboring pixels is larger than a pre-defined threshold, TH_3. If the number of high-gradient pixels within the CU is lower than a pre-defined threshold TH_4, the CU is classified as a low-gradient CU. Otherwise, it is classified as a high-gradient CU. In this paper, with empirical testing, TH_3 is set to 32 and TH_4 is set as follows which depends on the CU size:

$$TH_4 = W_{CU} \times H_{CU}/64 \tag{6}$$

where W_{CU} and H_{CU} are the width and height of the CU respectively. This classification technique is based on [8] with modification that we would not analyze the number of base-color pixels to classify CU as screen content for further enhanced coding. Instead, we only classify CU as a low-gradient CU such that the IntraBC mode checking can be skipped due to its low misclassification probability. Hence, if it is a low-gradient CU, only MPMs, plus planar prediction if it is not included in MPMs, are checked within the conventional intra mode. RMD and RDO in intra mode, as well as the entire IntraBC mode checking are skipped.

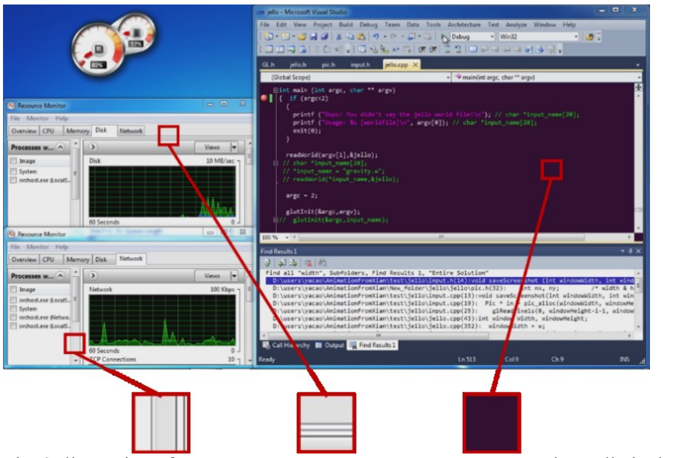

Fig. 2 Illustration of *sc_programming* screen content sequence where all pixels within the areas highlighted by red color are exactly equal along the horizontal direction or vertical direction.

TABLE I
Number of CUs (%) with zero CU activity.

Sequences	16×16	8×8
Basketball_Screen	38.83	51.12
MissionControlClip2	25.29	35.83
sc_desktop	36.86	58.11
MissionControlClip3	20.88	33.96
sc_web_browsing	47.63	63.96
sc_programming	32.56	45.59
Average	**33.68**	**48.10**

V. EXPERIMENTAL RESULTS

To evaluate the performance of the proposed algorithm, simulations were carried out by using the HEVC reference software SCM-4.1 [26]. The experiments were performed on the computer Dell Precision T1700 with Intel i7-4770 3.40GHz processor and 16GB memory. All experiments were conducted with all-intra (AI) configuration using QP {22, 27, 32 and 37} and YUV 4:4:4 sequences according to CTC [25]. First 100 frames were encoded. For the sake of simplicity, HEVC+SCC means HEVC SCC extension, whereas ACT and GRAD are denoted as our proposed fast approaches based on CU activity and gradient in Section IV A and Section IV B respectively. It is noted that the fast approaches in [15-17] have been implemented in HEVC+SCC.

TABLE II and TABLE III tabulate the Bjontegaard delta bitrate (BD-rate) [27] and encoding time in percentage for various approaches against the conventional HEVC. For HEVC+SCC, there is 62.2% BD-rate reduction on average but with very high computational complexity of about 90.0% increase in encoding time on average compared with the conventional HEVC. With our proposed HEVC+SCC+ACT, the increased encoding time is reduced to about 79.6% compared to the conventional HEVC with only about 0.2% drop in BD-rate compared to HEVC+SCC. With our proposed HEVC+SCC+ACT+GRAD, the increased encoding time is reduced to about 62.2% compared to the conventional HEVC. And there is only about 1.1% drop in BD-rate compared with HEVC+SCC which is negligible while there is still about 61.1% BD-rate reduction compared to the conventional HEVC. It is proved that CUs with zero activity can be efficiently predicted by the intra mode and CUs with low-gradient would seldom use the IntraBC mode. Fig. 3 and 4 show the RD performance of *MissionControlClip3* and *sc_desktop* for various approaches respectively. TABLE IV shows the number of searching points in percentage for the IntraBC mode of various approaches against HEVC+SCC. With our proposed approaches, the number of searching points are reduced by about 8.2% and 19.8% for HEVC+SCC+ACT and HEVC+SCC+ACT+GRAD, respectively. This means that there is redundant IntraBC mode checking and our approaches successfully skip it.

VI. CONCLUSIONS

The Intra Block Copy (IntraBC) mode helps to increase the coding efficiency of HEVC screen content coding by finding the repeating patterns within the same frame but encoding time is largely increased. By using our proposed approaches, unnecessary IntraBC mode checking can be skipped, the increased encoding time is largely reduced from 90.0% to 62.2% on average while bitrate is increased negligibly.

TABLE II
BD-rate (%) against the conventional HEVC.

Sequences	HEVC +SCC	HEVC +SCC+ACT	HEVC+SCC +ACT+GRAD
Basketball_Screen	-49.546	-49.357	-47.686
MissionControlClip2	-45.973	-45.785	-44.226
sc_desktop	-82.889	-82.833	-82.707
MissionControlClip3	-64.959	-64.804	-64.047
sc_web_browsing	-79.175	-78.985	-78.501
sc_programming	-50.711	-50.260	-49.305
Average	**-62.209**	**-62.004**	**-61.079**

TABLE III
Encoding time (%) against the conventional HEVC.

Sequences	HEVC +SCC	HEVC +SCC+ACT	HEVC+SCC +ACT+GRAD
Basketball_Screen	97.099	84.668	58.716
MissionControlClip2	104.376	91.724	49.876
sc_desktop	81.726	70.345	70.870
MissionControlClip3	94.482	86.605	68.175
sc_web_browsing	80.453	71.974	63.661
sc_programming	81.774	72.528	62.108
Average	**89.985**	**79.641**	**62.234**

TABLE IV
Number of searching points (%) against HEVC+SCC.

Sequences	HEVC +SCC+ACT	HEVC+SCC +ACT+GRAD
Basketball_Screen	-9.987	-27.788
MissionControlClip2	-8.657	-27.979
sc_desktop	-11.383	-11.846
MissionControlClip3	-5.718	-15.739
sc_web_browsing	-5.405	-22.493
sc_programming	-7.768	-12.899
Average	**-8.153**	**-19.791**

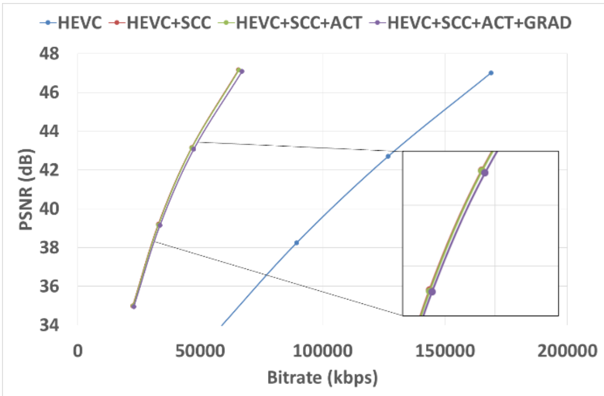

Fig. 3 RD performance of *MissionControlClip3* for various approaches.

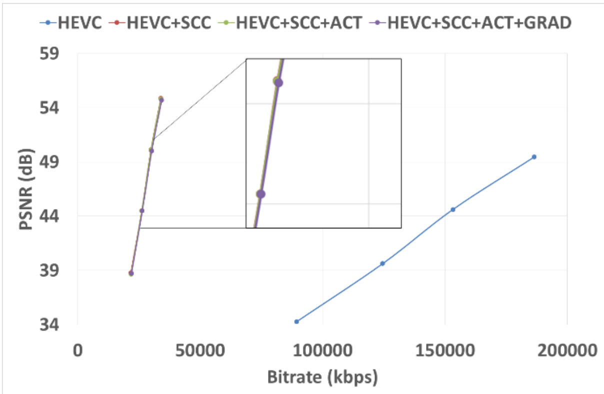

Fig. 4 RD performance of *sc_desktop* for various approaches.

References

[1] Y. Lu, S. Li, and H. Shen, "Virtualized screen: A third element for cloud-mobile convergence," *IEEE Multimedia*, vol. 18, no. 2, pp. 4–11, February 2011.

[2] G. J. Sullvian, J. Ohm, W.-J. Han, and T. Wiegand. "Overview of the High Efficiency Video Coding (HEVC) standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp.1649-1668, December 2012.

[3] "Joint Call for Proposals for Coding of Screen Content," *ISO/IEC JTC1/SC29/WG11*, document N14175, San Jose, California, U.S.A., January 2014.

[4] G. J. Sullivan, J. M. Boyce, Y. Chen; J.-R. Ohm, C. A. Segall, and A. Vetro, "Standardized Extensions of High Efficiency Video Coding (HEVC)," *IEEE J. Sel. Topics. Signal Process*, vol. 7, no. 6, pp.1001-1016, December 2013.

[5] J. Xu, R. Joshi, and R. A. Cohen, "Overview of the Emerging HEVC Screen Content Coding Extension," *IEEE Trans. Circuits Syst. Video Technol.*, Early Access Articles, September 2015.

[6] "Screen content coding test model 4 (SCM 4)," JCT-VC, JCTVC-T1014, Geneva, Switzerland, February 2015.

[7] T. Lin, Peijun Zhang, Shuhui Wang, Kailun Zhou, and Xianyi Chen, "Mixed Chroma Sampling-Rate High Efficiency Video Coding for Full-Chroma Screen Content," *IEEE Trans. Circuits Syst. Video Technol.*, vol.23, no.1, pp.173-185, January 2013.

[8] Z. Pan, H. Shen, Y. Lu, S. Li, and N. Yu, "A Low Complexity Screen Compression Scheme for Interactive Screen Sharing," *IEEE Trans. Circuits Syst. Video Technol.*, vol.23, no.6, pp.949-960, June 2013.

[9] J. Lainema, F. Bossen, W.-J. Han, J. Min, K. Ugur, "Intra Coding of the HEVC Standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol.22, no.12, pp.1792-1801, December 2012.

[10] Y. Piao, J.H. Min, and J. Chen, "Encoder Improvement of Unified Intra Prediction," *JCT-VC*, JCTVC-C207, Guangzhou, China, October, 2010.

[11] Weijia Zhu, Wenpeng Ding, Jizheng Xu, Yunhui Shi, and Baocai Yin, "Screen Content Coding Based on HEVC Framework," *IEEE Trans. Multimedia*, vol.16, no.5, pp.1316-1326, August 2014.

[12] S.-H. Tsang, Y.-L. Chan, and W.-C. Siu, "Exploiting Inter-layer Correlations in Scalable HEVC for the Support of Screen Content Videos," in *Proc. of Int. Conf. on Digital Signal Process. (DSP)*, pp.888-892, Hong Kong, August 2014.

[13] Z. Ma, W. Wang, M. Xu, and H. Yu, "Advanced Screen Content Coding Using Color Table and Index Map," *IEEE Trans. Image Process.*, vol.23, no.10, pp.4399-4412, October 2014.

[14] S.-H. Tsang, Y.-L. Chan, and W.-C. Siu, "Fast and Efficient Intra Coding Technique for Smooth Regions in Screen Content Coding Based on Boundary Prediction Samples," in *Proc. of IEEE Int. Conf. on Acoustics, Speech and Signal Process. (ICASSP)*, pp.1409-1413, Brisbane, Australia, April 2015.

[15] M. Budagavi, and D.-K. Kwon, "Intra Motion Compensation and Entropy Coding Improvements for HEVC Screen Content Coding," in *Proc. of Picture Coding Symp. (PCS)*, pp.365-368, San Jose, U.S.A., December 2013.

[16] M. Budagavi, and D.-K. Kwon, "Fast Intra Block Copy (IntraBC) Search for HEVC Screen Content Coding," in *Proc. of IEEE Int. Symp. on Circuits Syst. (ISCAS)*, pp.9-12, Melbourne, Australia, June 2014.

[17] J. Chen, Y. Chen, T. Hsieh, R. Joshi, M. Karczewicz, and W.-S. Kim et al., "Description of Screen Content Coding Technology Proposal by Qualcomm," *JCT-VC*, JCTVC-Q0031, Valencia, Spain, April-March 2014.

[18] W. Zhu, W. Ding, J. Xu, Y. Shi, and B. Yin, "Hash Based Block Matching for Screen Content Coding," *IEEE Trans. Multimedia,* vol. 17, no. 7, pp.935-944, July 2015.

[19] S.-H. Tsang, Y.-L. Chan, and W.-C. Siu, "Hash Based Fast Local Search for Intra Block Copy (IntraBC) Mode in HEVC Screen Content Coding," in *Proc. of APSIPA Annual Summit and Conf. (APSIPA ASC)*, pp.396-400, Hong Kong, December 2015.

[20] Y.-L. Chan, and Wan-Chi Siu, "An Efficient Search Strategy for Block Motion Estimation using Image Features," *IEEE Trans. on Image Process.*, vol. 10, no. 8, pp. 1223-1238, August 2001.

[21] K.-C. Hui, W.-C. Siu, and Y.-L. Chan, "Fast Motion Estimation of Arbitrarily Shaped Video Objects in MPEG-4," *Signal Process.: Image Comm.*, vol. 18, pp.33-50, January 2003.

[22] Y.-L. Chan, and W.-C. Siu, "A New Block Motion Vector Estimation using Adaptive Pixel Decimation," in *Proc. of the IEEE Int. Conf. on Acoustics, Speech and Signal Process. (ICASSP)*, vol. 4, pp. 2257-2260., Detroit, Michigan, U.S.A., May 1995.

[23] Y.-L. Chan, K.-C. Hui, and W.-C. Siu, "Adaptive Partial Distortion Search Algorithm for Block Motion Estimation," *J. of Visual Comm. and Image Representation*, vol. 15, no. 4, pp.489-506, December 2004.

[24] S.-H. Tsang, Y.-L. Chan, and W.-C. Siu, "Region-based Weighted Prediction for Coding Video with Local Brightness Variations", *IEEE Trans. on Circuits Syst. Video Technol.*, vol. 23, no. 3, pp. 549-561, March 2013.

[25] "Common Test Conditions for Screen Content Coding," *JCT-VC*, JCTVC-T1015, Geneva, Switzerland, February 2015.

[26] HM-16.6+SCM-4.1, HEVC Test Model Version 16.6 Screen Content Model Version 4.1, [Online], available at: https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/HM-16.6+SCM-4.1/.

[27] G. Bjontegaard, "Calculation of Average PSNR Differences Between RD Curves," *Video Coding Experts Group (VCEG)*, VCEG-M33, Austin, Texas, U.S.A., April, 2001.