Fast Coding Strategy for HEVC by Motion Features and Saliency Applied on Difference Between Successive Image Blocks

Pallab Kanti Podder[1][(✉)], Manoranjan Paul[1], and Manzur Murshed[2]

[1] School of Computing & Mathematics, Charles Sturt University,
Bathurst, NSW 2795, Australia
`{ppodder,mpaul}@csu.edu.au`
[2] School of Information Technology, Federation University,
Churchill, VIC 3842, Australia
`manzur.murshed@federation.edu.au`

Abstract. Introducing a number of innovative and powerful coding tools, the High Efficiency Video Coding (HEVC) standard promises double compression efficiency, compared to its predecessor H.264, with similar perceptual quality. The increased computational time complexity is an important issue for the video coding research community as well. An attempt to reduce this complexity of HEVC is adopted in this paper, by efficient selection of appropriate block-partitioning modes based on motion features and the saliency applied to the difference between successive image blocks. As this difference gives us the explicit visible motion and salient information, we develop a cost function by combining the motion features and image difference salient feature. The combined features are then converted into area of interest (AOI) based binary pattern for the current block. This pattern is then compared with a previously defined codebook of binary pattern templates for a subset of mode selection. Motion estimation (ME) and motion compensation (MC) are performed only on the selected subset of modes, without exhaustive exploration of all modes available in HEVC. The experimental results reveal a reduction of 42 % encoding time complexity of HEVC encoder with similar subjective and objective image quality.

Keywords: HEVC · Motion features · Saliency feature · Area of Interest (AOI) · Intermode selection

1 Introduction

The prime goal of the latest HEVC [1] standard is to provide the same perceptual quality compared to its antecedent H.264 [2], at approximately 50 % bitrate reduction for efficient transmission and storage of high volume video data [3]. HEVC introduces inventive approaches, including the coding unit (CU) size extension from 16×16 up to 64×64-pixels, variable size transform unit (TU), and symmetric/asymmetric block partitioning phenomenon that have achieved

on improved performance gain, but at a cost of more than 4 times algorithmic complexity for a particular implementation [4,5]. For this reason, any electronic devices with limited processing capacity could not fully exploit HEVC encoding and decoding features. This motivated us to reduce the computational time of HEVC encoder by appropriate selection of inter-prediction modes. For this to happen, only the AOI in a video is taken into account that comprise with phase correlation based motion features and the saliency feature after applying it on the difference between successive image blocks. Unlike *HEVC test model* (HM12.1) [6], the selected CUs with, and without, AOI are motion estimated and motion compensated with modes in the higher and lower depth levels, respectively. Thus, exhaustive exploration of all modes in each coding depth level can be avoided (level 64×64, 32×32, 16×16 and 8×8 are denoted as depth level 0, 1, 2, 3 respectively) by proposed method. This results in computational time reduction. To select a particular motion prediction mode, HM exhaustively checks the *Lagrangian cost function* (LCF) [7]using all modes in each coding depth level. The LCF, $\Theta(k_n)$ for mode selection (kn is the nth mode) is defined by:

$$\Theta(k_n) = D(k_n) + \lambda \times R(k_n) \tag{1}$$

where λ is the Lagrangian multiplier, $D(k_n)$ is the distortion, and $R(k_n)$ is the resultant bit, which are determined by modes for each CU. In order to select the best partitioning mode in a coding depth level, HM checks at least 8, and at most 24, inter-prediction modes with lowest LCF. This process is time consuming due to the exploration of all modes in one or more coding depth levels. Moreover, only the LCF-based mode decision could not always provide the best *rate distortion*(RD) performance in different operational coding points due to advanced parameter settings in HEVC. Therefore, instead of merely depending on LCF, some consecutive pre-processing stages are executed by the proposed technique that makes mode decision process more appropriate and less time consuming (shown in Fig. 1).

In literature, some researchers have contributed to reduce time complexity [8,9]. In order to terminate the exploring modes in lower level, Hou *et al.* [10] recommend a RD cost based threshold to explore modes only in the higher level that results in 30 % time savings with 0.5 % quality loss. Vanne *et al.* [11] propose an efficient inter-mode decision scheme by finding the candidate PU modes of symmetric and asymmetric motion partition. The tested results reveal the reduction of HEVC encoder complexity by 31 %–51 % at the cost of 0.2 %–1.3 % bit-rate increment. Pan *et al.* [12] introduce an early MERGE mode decision algorithm to reduce computational complexity of HEVC encoder. They achieve 35 % time savings with the bit-rate increment of 0.32 %, and quality loss of 0.11 dB peak signal to noise ratio (PSNR).

The *energy concentration ratio* (ECR) from phase correlation was extracted and employed for mode selection in HEVC by Podder *et al.* [13](used in [14] for H.264 standard). They save computational time by sacrificing 0.24 dB PSNR on average compared to the full search approach of HM. As ECR stipulates only the residual error between current block and motion-compensated reference block,

it would not provide expected compression results. For more accurate decision on ME and MC modes, in this paper, we extract three motion features from phase correlation. Other than motion, visual attentive areas are also extracted by exploiting the saliency feature. Compared to the current image, as the difference between two successive images provides the actual displacement and salient information, we apply the saliency feature on image difference to capture motion and saliency dominated precise areas perceived by human visual system. Moreover, as there is no difference of color or contrast in static areas, the image difference produced motion and salient information would be the premier option to determine visual observant areas. The features are incorporated by developing a weighted cost function to actuate AOI-based binary pattern for the current block to select a subset of inter-modes. From the selected subset, the final mode is determined based on their lowest LCF. The proposed method not only reduces the computational time by appropriate selection of AOI based ME and MC modes but also demonstrates similar subjective and objective image quality.

The remainder of this paper is structured as follows: Sect. 2 describes all the key steps of the proposed method; experimental results and discussions are detailed in Sect. 3, while Sect. 4 is the conclusion of the paper.

2   Proposed Mode Selection Technique

The phase-correlation renders relative displacement information between current block and the reference block by Fast Fourier Transformation (FFT). We first extract phase-correlation related three motion features including (i) ECR (α), (ii) predicted motion vector (dx, dy) and (iii) phase correlation peak (β) and combine them with saliency feature (γ) that is applied on the difference between successive image blocks. We evolve a unified AOI based cost function using the normalized motion features and weighted average of the saliency values to determine a unified AOI feature for the current block. The binary AOI pattern from the unified AOI features is then configured by applying threshold and to select a sub-set of inter-prediction modes. The final mode from the selected subset is determined by their lowest LCF. The whole process is shown as a block diagram in Fig. 1.

2.1   Motion Features Extraction

The phase correlation is calculated by applying the FFT and then *inverse* FFT (IFFT) of the current and reference blocks and finally applying the FFTSHIFT function as follows:

$$\Omega = fftshift \left| ifft \left(e^{j(\angle F_r - \angle F_c)} \right) \right| \tag{2}$$

where Fc and Fr are the Fast Fourier transformed blocks of the current C and reference R blocks respectively and $\angle$ is the phase of the corresponding transformed block. Note that Ω is a two dimensional matrix. We evaluate the
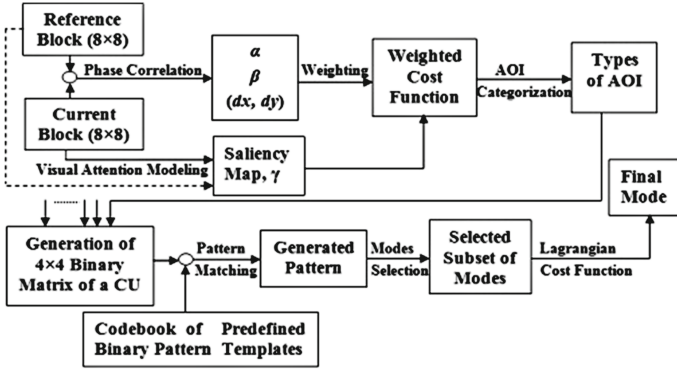
Fig. 1. Block diagram of the proposed mode selection process.

phase correlation peak (β) from the position of $(dx + blocksize/2 + 1, dy + blocksize/2 + 1)$ as follows:

$$\beta = \Omega\left(dx + blocksize/2 + 1, dy + blocksize/2 + 1\right) \tag{3}$$

where the blocksize is 8 if 8×8-pixel block is used for phase correlation. Then we compute the predicted motion vector (dx, dy) by subtracting blocksize-1 from the (x, y) position of Ω where we find the maximum value of Ω. We use the phase of the current block and magnitude of the motion-compensated block in the reference frame and finally calculate the matched reference block (Ψ) for the current block by:

$$\Psi = \left| ifft\left(|F_r|\, e^{j(\angle F_c)}\right) \right| \tag{4}$$

Now the displacement error (§) is enumerated by: §=C-Ψ. We then apply the *discrete cosine transform* (DCT) to error §and calculate the ECR (i.e., α) as the ratio of low frequency component and the total energy of the error block by:

$$\alpha = \left(D_{error\_low}/D_{error\_total}\right) \tag{5}$$

where D_{error_low} and D_{error_total} represent the top-left triangle energy and the whole area energy of a particular block.

2.2   Saliency Feature Extraction

The saliency map is operated as a tool (based on [15,16]) and incorporated into our coding architecture. The exploited graph-based visual saliency (GBVS) technique gives us the variance map of AOI based human visual features for an 8×8-pixel block consisting of the values ranging from 0 to 1. We extract the saliency feature, γ, by averaging all values of a given 8×8 block. Our focus is to encode the AOI based salient portions with more bits to achieve better compression and improve coding performance.

2.3   Cost Function Based AOI Categorization

After evaluating the phase correlation extracted motion features (i.e., α, β and (dx, dy) and saliency extracted variance map (i.e., γ), we finally determine a cost function $\eta(i,j)$ from these four features for (i,j)th block by-

$$\eta(i,j) = \omega_1\alpha(i,j) + \omega_2(1-\beta) + \omega_3((\frac{|dx|}{\delta}) + (\frac{|dy|}{\delta})) + \omega_4(\gamma) \qquad (6)$$

where δ is the maximum block size, and $\omega1$ to $\omega4$ are the weights with $\sum_{i=1}^{4}\omega_i = 1$. We innovatively derive weights for each feature and only consider 0.50, 0.25, 0.125, and 0.125 weights based on the relative texture deviation of the current block against that of the whole frame. First we sort four features based on their values and if the value of the *Standard Deviation*(STD) of the block is smaller than the value of the current frame then the highest weight (i.e., 0.50) is applied to the feature 1 (i.e., sorted) and the lowest weight (i.e., 0.125) is applied to the feature 4 (according to sorted list); otherwise, inverse weighted order is applied. If the resultant value of the cost function (i.e., η) is greater than a predefined threshold the block is tagged by '1' otherwise tagged by '0' where binary '1' and '0' corresponds to AOI and non-AOI respectively. The rationality of the proposed weight selection strategy is that if the current block has higher texture variation compared to the current frame, the current block should be encoded with more bits compared to the rest of the blocks to achieve similar/improved RD performance. The relationship of the quantitative motion and salience features with the human visual features are depicted in Fig. 2. Figure 2 (b-d) shows the categories of motion-peak (β) and their corresponding values provided by ECR (Fig. 2 (e)) and saliency feature (Fig. 2 (f)) for *Tennis* video. Figure 3 illustrates

(a) Difference between 11th and 12th frame on *Tennis* video

(b) No Motion   (c) Simple motion   (d) Complex motion

Energy Concentration Ratio

Salience Feature Applied On Difference Between Successive Image Blocks

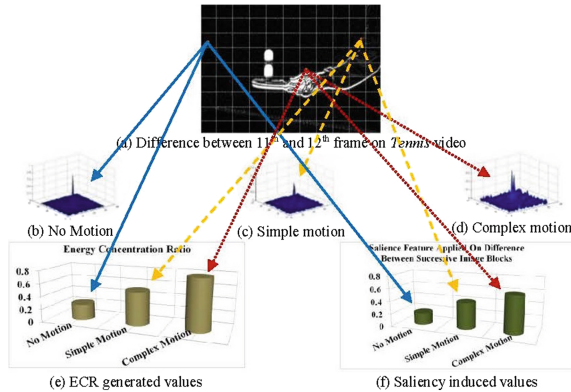(e) ECR generated values   (f) Saliency induced values

Fig. 2. Illustration of motion features and saliency feature generated at different CUs of 12th frame on Tennis video; (b-d) are the phase shifted plots for no motion (0.4), simple motion (0.7) and complex motion (0.8); (e-f) corresponds to the respective values generated by ECR and saliency map for CUs at positions (3, 1), (3, 10) and (5, 7) respectively. For clear visualization we use 32×32 block size.

the impact of saliency map and from Fig. 3 (a), we clearly depict the ellipse like red marked area as the table tennis court edge (obviously a visually attentive area) that is precisely identified by the saliency feature (red marked ellipse like area in Fig. 3 (c)) although three motion features of phase correlation could not grasp the court edge as it does not have any motion (white marked ellipse like area in Fig. 3 (b)). Thus, the proposed combined strategy improves the RD performance by recognizing not only the AOI-based motion features of phase correlation but also the AOI-based visually attentive areas inside the videos.
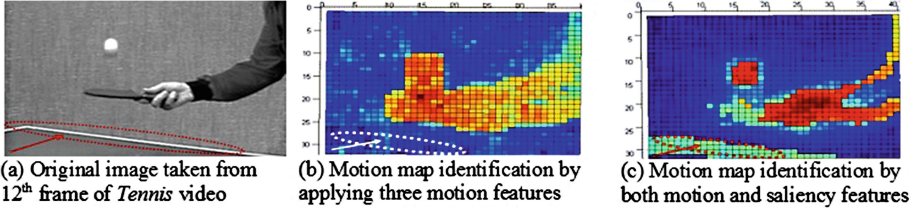

(a) Original image taken from 12th frame of *Tennis* video

(b) Motion map identification by applying three motion features

(c) Motion map identification by both motion and saliency features

Fig. 3. Identification of motion and salient areas with and without saliency feature based cost function (Color figure online).

2.4   Intermode Selection

For the generation of binary matrix, we exploit each of the 8×8 pixel blocks from the 32×32 pixel blocks (i.e., CU) and produce a matrix of 4×4 binary values for each CU (applying threshold). The cost function generated 4×4 binary matrix is then compared with a codebook of predefined binary pattern templates (BPT) to select a subset of modes. Each of the templates is constructed with a pattern of AOI and non-AOI block focusing on the rectangular and regular object shapes at 32×32 block level as shown in Fig. 4. Both in Fig. 4 (for 32×32 level) and Table 1(b) (for 16×16 level), the cells with black squares present the AOI (i.e., binary 1) and the rest are non-AOI (i.e., binary 0). We use a simple similarity metric using the Hamming Distance (DH) between the binary matrix of a CU generated by phase correlation and the BPTs in Fig. 4. We select the best-matched BPT that provides minimum sum of the absolute values of their differences for a CU.

The DH, D_h is determined as follows where S is the binary motion prediction matrix of a CU comprising 4×4 '1' or '0' combinations and Tk is the k-th BPT:

$$D_h(x,y) = \sum_{x=0}^{4} \sum_{y=0}^{4} |S(x,y) - T_k(x,y)| \tag{7}$$

The best-matched j-th BPT is selected from all BPTs as follows:

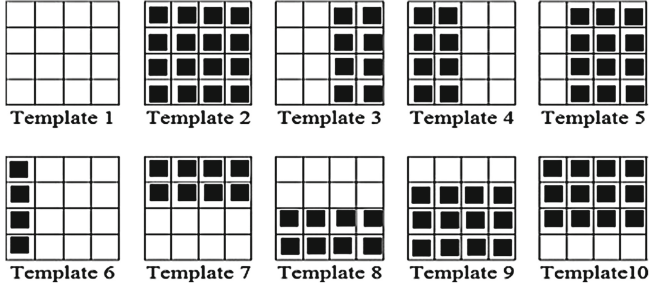$$B_j = arg_{T_k \forall BPT} \min(D_h) \tag{8}$$

Fig. 4. Codebook of the proposed binary pattern templates for subset of inter-mode selection where template cells with black squares present AOI (i.e., binary 1) and the rest are non-AOI (i.e., binary 0).

The mode selection process from the BPTs at 32×32 and 16×16 coding depth levels is illustrated in Table 1 (a) and (b) respectively. Once a particular template selects a subset of candidate modes at 32×32 level, the final mode is decided by their lowest Lagrangian cost function. Again, at 32×32 level, if any of the 16×16 coding depth level modes is selected, we further explore smaller modes at 16×16 level using the AOI pattern (i.e., presence or absence of binary 1 and 0 as shown in Table 1 (b)).

Table 1. Selection technique of inter-modes at 32×32 and 16×16 coding depth levels.

Templates Based on AOI at 32×32 Block Level	Selection of Modes at 32×32 Block Level
Template 1	Skip or 32×32
Template 2	Intra 16×16 or Inter 16×16
Template 3 & 4	32×16 or Inter 16×16
Template 5	32×8 or Inter 16×16
Template 6	32×24 or Inter 16×16
Template 7 & 8	16×32 or Inter 16×16
Template 9	8×32 or Inter 16×16
Template 10	24×32 or Inter 16×16

(a) Mode selection at 32×32 block level based on predefined binary pattern templates

(b) Subset of mode selection at 16×16 block level based on AOI pattern

Then the equation for the final mode (ξ) selection is given by:

$$\xi = arg_{\forall k_n} \min(\Theta(k_n)) \tag{9}$$

where $\Theta(k_n)$ is the Lagrangian cost function for mode selection.

2.5   Threshold Determination

Due to the imbalanced distribution of ECR values Podder *et al.* [13] (also mentioned different thresholds in [17,18]) use a range of thresholds from 0.37 to 0.52 for different bit-rates. Those thresholds could not perform well in the proposed method as the distribution of cost function values is more compact and almost in all cases it exceeds the value of 0.15 in all types of sequences for the blocks with motion and dominant salience. Therefore, in the proposed technique, we use the fixed value of threshold (i.e., 0.15) for a wide range of bit-rates.

3   Experimental Results and Analysis

To verify the performance of the proposed algorithm, experimental results are presented with six *standard definition* (SD) videos- *Tennis, Tempete, Waterfall, Silent, Paris, Bridgeclose*, four *high definition* (HD) videos- *Pedestrian, Bluesky, Rushhour, Parkrun* and two *multiview* (MV) videos- *Exit and Ballroom*. Each of the video sequences are encoded with 25 frame rate and search length ±64. We compare the proposed method results with HM of HEVC standard as HM outperforms the existing mode selection techniques in the literature. We use IPPP format with Group of picture (GOP) 32 for both techniques and two reference frames.

Table 2. Average time savings (%) by the proposed method (against HM) in terms of mode selection for each type of sequences-a theoretical analysis.

Sequence types	Average no. of inter-modes selected per CU by HM	Average no. of inter-modes selected per CU by proposed method	Average time savings (%) in terms of mode selection
SD	16.89	7.24	57.14
HD	18.92	9.63	49.11
MV	19.58	11.16	43.01
Average percentage (%) of time savings			49.75

3.1   Experimental Setup

In this paper, the experiments are conducted by a dedicated desktop machine (with Intel core i7 3770 CPU @ 3.4 GHz, 16 GB RAM and 1TB HDD) running 64 bit Windows operating system. The proposed scheme and HEVC with exhaustive mode selection scheme are developed based on the reference software HM (version 12.1) [6]. RD performance of both schemes are compared considering the maximum CU size of 32×32 by enabling both symmetric and asymmetric partitioning block size of 32×32 to 8×8 depth levels for a wide range of bit-rates (using QP=20, 24, 28, 32 and 36). The calculations of BD-PSNR and BD-Bit Rate were performed according to the procedures described in [19].

3.2 Results and Discussions

For the theoretical justification of computational time of all type of sequences, we first compare the average number of modes selected in each CU by HM12.1 and the proposed method. The results in Table 2 shows that HM checks more options in all cases and normally requires more computational time. From Table 2, the overall average percentage of encoding time savings by the proposed method is 49.75 and the reason of this acquisition is the efficient subset of intermode selection with simple criteria. However, we cannot ignore the pre-processing stages of the proposed method, and by calculation we notice that over twelve sequences on average 6.71 % extra encoding time is required to execute phase correlation and saliency related pre-processing overheads (see Fig. 1). Thus, theoretically we anticipate to acquire 43.04 % computational time savings on average. The experimental evaluation reveals that over twelve different sequences, and for a wide range of bit-rates, the proposed method reduces on average 42 % (range: 37 %–45 %) computational time as shown in Fig. 5 (a). The equation for the time savings (TS) is defined as:

$$TS = \frac{(T_o - T_p)}{T_o} \times 100\,\%\qquad(10)$$

where T_0 and T_p denote the total encoding time consumed by HM and the proposed method respectively. For comprehensive performance test, we execute the computational time analysis of both techniques based on video categories and find that the proposed method achieves on average 41 % encoding time savings compared to HM12.1 as shown in Fig. 5 (b). The figure also reveals that the proposed technique saves more computational time for SD video type (48 %).
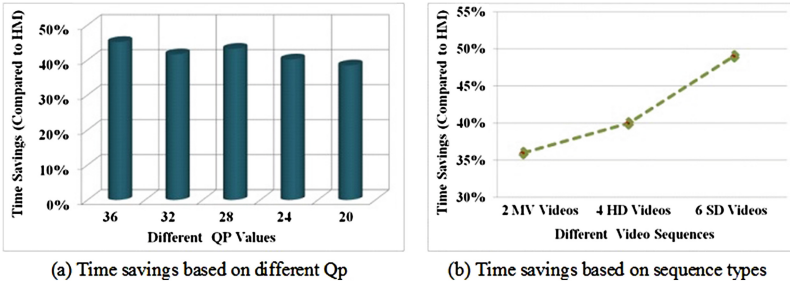

(a) Time savings based on different Qp   (b) Time savings based on sequence types

Fig. 5. Illustration of time savings by the proposed method against HM.

To test the performance of the proposed method objectively, we first compare the RD performance against HM using three different sequence types (one SD, HD and MV) for a wide range of bit-rates as demonstrated in Fig. 6. The figure shows that the proposed method achieves similar RD performance with HM12.1 especially caring about the AOI based CUs and partitioning them by efficient
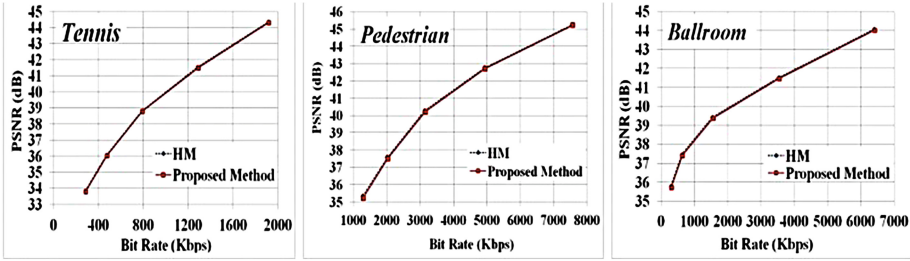
Fig. 6. Comparative study on RD performance by HM12.1 and the proposed method for a wide range of bit-rates.

selection of appropriate block partitioning modes. Table 3 represents the performance comparison of the proposed method for twelve divergent video sequences. The results reveal that compared to the mode selection approach in HM, the proposed technique achieves an almost similar RD performance (small average reduction of 0.021 dB PSNR) with a negligible bit-rate increment of 0.14 %. Figure 7(a) shows the original image of *Tennis* video taken for subjective quality test and Fig. 7(b) and Fig. 7(c) illustrate the reproduced images by HM and the proposed method respectively. To present the comparison in image quality let us concentrate on the cuff and sleeve sections of the shirt in the three images which are marked by the red, yellow, and white ellipses respectively. It can be perceived that the three ellipse marked sections have almost similar image quality. It was presented earlier in this manuscript that the proposed method requires less encoding time. Hence it can be concluded that the proposed technique shows significant computational time savings compared to HM12.1 with similar image quality for a wide range of bit-rates. Due to this phenomenon, the proposed implementation is expected to become more suitable for all real

Table 3. Performance comparison of proposed technique compared to HM using BD-Bit Rate and BD-PSNR.

Video sequences	BD-PSNR (dB)	BD-Bit Rate (%)	Video sequences	BD-PSNR (dB)	BD-Bit Rate (%)
Tennis	−0.010	0.06	*Pedestrian*	−0.025	0.14
Tempete	−0.021	0.14	*Bluesky*	−0.039	0.22
Waterfall	−0.018	0.11	*Rushhour*	−0.031	0.18
Silent	−0.024	0.16	*Parkrun*	−0.023	0.15
Paris	−0.013	0.10	*Ballroom*	−0.026	0.20
Bridgeclose	−0.012	0.13	*Exit*	−0.013	0.19
Average BD-PSNR and BD-Bit Rate over twelve sequences				−0.021	0.14

(a) Original image of *Tennis* sequence   (b) HM reproduced image   (c) Proposed method reproduced image
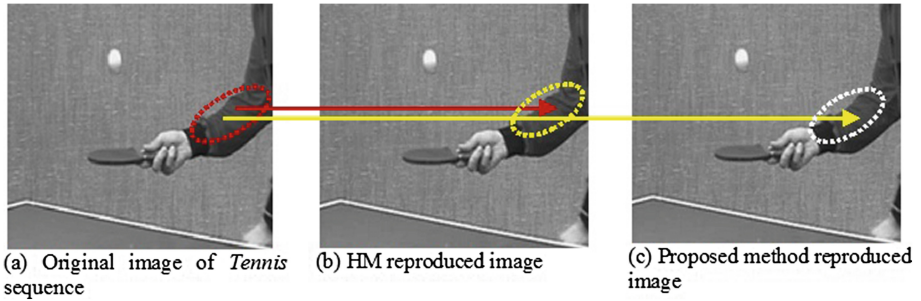
Fig. 7. Subjective quality assessment for HM12.1 and the proposed method for *Tennis* video sequence. The figures are achieved from the 20*th* frame of the *Tennis* video at the same bit-rate (Color figure online).

time video coding applications especially for a number of electronic devices with limited processing power and battery capacity.

4 Conclusion

In this work, a novel coding framework for HM performance improvement is presented. This is implemented by exploiting the AOI based mode selection technique comprising three different motion features of phase correlation and saliency features of human visual attention. The motion features focus on three different aspects of motions in each CU and the saliency feature captures the visual attentive areas. An adaptive cost function is formulated to determine a subset of inter-modes using predefined AOI based binary pattern templates. The Lagrangian optimization criterion is employed in the selected subset of modes to fix the final mode. Compared to HM with exhaustive mode selection strategy, the proposed scheme reduces on average 42 % computational time (range: 37 %–45 %) while providing similar image quality for a wide range of bit-rates.

References

1. High Efficiency Video Coding, document ITU-T Rec. H.265 and ISO/IEC 23008–2 (HEVC), ITU-T and ISO/IEC, April 2013
2. Wiegand, T., Sullivan, G.J., Bjontegaard, G., Luthra, A.: Overview of the H.264/AVC video coding standard. IEEE Trans. Circ. Syst. Video Technol. **13**(7), 560–576 (2003)
3. Bross, B., Han, W.J., Ohm, J.R., Sullivan, G.J., Wiegand, T.: High efficiency video coding text specification draft 8. JTCVC- J1003, Sweden (2012)

4. Bossen, F., Bross, B., Suhring, K., Flynn, D.: HEVC complexity and implementation analysis. IEEE Trans. Circ. Syst. Video Technol. **22**(12), 1684–1695 (2012)

5. Lu, Y.: Real-Time CPU Based H.265/HEVC Encoding Solution with Intel Platform Technology. Intel Corporation, Shanghai, PRC (2013)

6. Joint Collaborative Team on Video Coding (JCT-VC), HM Software Manual, CVS server (2013) (http://hevc.kw.bbc.co.uk/svn/jctvc-hm/)

7. Paul, M., Murshed, M.: Video coding focusing on block partitioning and occlusions. IEEE Trans. Image Process. **19**(3), 691–701 (2010)

8. Lee, A., Jun, D.S., Kim, J., Seok, J.: An efficient interprediction mode decision method for fast motion estimation in HEVC. In: International Conference on ICT Convergence (ICTC), pp. 502–505, October 2013

9. Tan, H.L., Liu, F., Tan, Y.H., Yeo, C.: On fast coding tree block and mode decision for high efficiency video coding. In: International Conference on Acoustic and Speech Signal Processing (ICASSP), pp. 825–828, March 2012

10. Hou, X., Xue, Y.: Fast coding unit partitioning algorithm for HEVC. In: IEEE International Conference on Consumer Electronics (ICCE), pp. 7–10 (2014)

11. Vanne, J., Viitanen, M., Hamalainen, T.: Efficient mode decision schemes for HEVC inter prediction. IEEE Trans. Circ. Syst. Video Technol. **24**(9), 1579–1593 (2014)

12. Pan, Z., Kwong, S., Sun, M.T., Lei, J.: Early MERGE mode decision based on motion estimation and hierarchical depth correlation for HEVC. IEEE Trans. Broadcast. **60**(2), 405–412 (2014)

13. Podder, P., Paul, M., Murshed, M., Chakrabarty, S.: Fast intermode selection for HEVC video coding using phase correlation. In: IEEE International Conference on Digital Image Computing: Techniques and Applications (DICTA), pp. 1–8 (2014)

14. Paul, M., Lin, W., Lau, C.T., Lee, B.S.: Direct inter-mode selection for H.264 video coding using phase correlation. IEEE Trans. Image Process. **20**(2), 461–473 (2011)

15. Harel, J., Koch, C., Perona, P.: Graph-Based Visual Saliency. California Institute of Technology, Pasadena, (2006)

16. January 2015. http://www.vision.caltech.edu/~harel/share/gbvs.php and http://libra.msra.cn/Publication/4113493/graph-based-visual-saliency

17. Paul, M., Frater, M., Arnold, J.: An efficient mode selection prior to the actual encoding for H.264/AVC encoder. IEEE Trans. Multimedia **11**(4), 581–588 (2009)

18. Podder, P., Paul, M., Murshed, M.: A novel motion classification based intermode selection strategy for HEVC performance improvement. Elsevier J. Neurocomputing **173**(3), 1211–1220 (2016). doi:10.1016/j.neucom.2015.08.079

19. Bjontegaard, G.: Calculation of average PSNR differences between RD curves. ITU-T SC16/Q6, VCEG-M33, Austin, USA (2001)