Int. J. Electron. Commun. (AEÜ) 95 (2018) 73–81

Contents lists available at ScienceDirect

International Journal of Electronics and Communications (AEÜ)

journal homepage: www.elsevier.com/locate/aeue

Gradient-oriented directional predictor for HEVC planar and angular intra prediction modes to enhance lossless compression

S. Shilpa Kamath [a,*], P. Aparna [a], Abhilash Antony [b]

[a] National Institute of Technology Karnataka, Karnataka, India
[b] Muthoot Institute of Technology and Science, Kerala, India

ARTICLE INFO

Article history:
Received 4 June 2018
Accepted 30 July 2018

Keywords:
HEVC
Planar
Angular
Gradient
Lossless

ABSTRACT

Recent advancements in the capture and display technologies motivated the ITU-T Video Coding Experts Group and ISO/IEC Moving Picture Experts Group to jointly develop the High-Efficiency Video Coding (HEVC), a state-of-the-art video coding standard for efficient compression. The compression applications that essentially require lossless compression scenarios include medical imaging, video analytics, video surveillance, video streaming etc., where the content reconstruction should be flawless. In the proposed work, we present a gradient-oriented directional prediction (GDP) strategy at the pixel level to enhance the compression efficiency of the conventional block-based planar and angular intra prediction in the HEVC lossless mode. The detailed experimental analysis demonstrates that the proposed method outperforms the lossless mode of HEVC anchor in terms of bit-rate savings by 8.29%, 1.65%, 1.94% and 2.21% for Main-AI, LD, LDP and RA configurations respectively, without impairing the computational complexity. The experimental results also illustrates that the proposed predictor performs superior to the existing state-of-the-art techniques in the literature.

1. Introduction

The global network is presently burdened due to the plethora of multimedia services available to the end user. Demand for superior quality digital videos has become a global necessity, thereby requiring a video coding standard with superior compression capabilities. These standards play an imperative role in setting up the technology, by ensuring interoperability amongst the products developed by the various manufacturers. Although the current most popular standard H.264/AVC provides gain in coding efficiency up to 50% over wide range of video resolutions as compared to its predecessors, the infrastructure is incapable in transmitting the high definition (HD) and ultra high definition (UHD) video content to the customer due to the available bandwidth and the memory constraints. With the development of enhanced coding tools that improves the coding efficiency, HEVC thus aims to provide a robust solution to aid the existing multimedia applications supported by the current standards.

HEVC [1–3] is a video coding standardization project developed by Joint Collaborative Team on Video Coding (JCT-VC), a collaboration between ITU-T Video Coding Experts Group (VCEG) and ISO/IEC Moving Pictures Expert Group (MPEG). It possesses a generic syntax and structure design similar to its predecessor H.264/AVC [4,5], but with new enhancement and improvement. For the same visual quality, the standard intends to achieve 50% reduction in bit-rate compared to its predecessor. This leads to a reduction in video storage space and transmission cost, thus mitigating the burden on the global networks and paving way for faster streaming of the HD/UHD content.

In general, lossless compression technique deals with representing an image/video content with fewer bits in a manner that the original content can be reconstructed without any error or distortion [6–8], meanwhile retaining the archival quality. It is desirable in storage applications such as medical imaging, video analytics, image archiving, remote sensing, high precision image analysis and mainstream applications like video surveillance, video conferencing, video streaming. Due to the enormous growth in the demand for these applications currently, there has been a lot of interest in the lossless compression techniques. Lossless coding is achieved by bypassing the transformation, quantization, their inverse operations and in-loop filtering operations which includes a deblocking filter, sample adaptive offset (SAO) and adaptive loop filter (ALF) at the encoder and decoder [9]. The prediction residuals can hence be generated at the pixel level instead of block level and be directly fed to the entropy coding engine, thus inducing a need to enhance the intra prediction accuracy in the lossless mode.

* Corresponding author.
E-mail address: shilpa.1107@yahoo.co.in (S. Shilpa Kamath).

https://doi.org/10.1016/j.aeue.2018.07.037
1434-8411/

The proposed work focuses on assigning the best prediction mode to every prediction block, out of the 35 intra prediction modes based on the gradient analysis. It is accomplished by considering the edges and the directional structures within the current block. This implicitly results in the enhancement of the prediction accuracy and overall coding efficiency of the HEVC intra prediction in lossless mode.

Remaining paper is organized as follows. Section 2 provides an overview of the HEVC intra prediction modes, Section 3 gives a brief insight into the associated literature study and Section 4 expounds the proposed method. A discussion on the generated simulation results is dealt in Section 5, followed by conclusion in Section 6.

2. Overview of HEVC block-based intra prediction modes

Quadtree-based block partitioning structure is embedded within the HEVC to facilitate a flexible usage of the large and small prediction and transform block sizes [10–12]. It possesses 35 intra prediction modes which include planar (M0), DC (M1) and 33 angular directional modes (M2-M34) to exploit the spatial correlations existing within the given frame. Although the actual prediction process is instigated in the prediction unit (PU), the selection of the best-suited mode for the prediction process is performed at the coding unit (CU) level [13–15]. The prediction process for an $N \times N$ block is simplified by the formation of a 1-d array of $(4N + 1)$ reference samples using the previously reconstructed adjacent blocks on the top and left of the current block as depicted in Fig. 1. The HEVC modes utilize these reference samples to generate the prediction values for the block under consideration.

The planar prediction mode is well suited to provide smooth prediction signals with gradual changes. This mode works by predicting the block using vertical and horizontal slopes derived from the neighboring reference samples. It utilizes the four reference samples for averaging the two linear predictions to eliminate the discontinuities along the block boundaries as shown in Fig. 2. In this mode, the prediction value of the target pixel $C_{x,y}$ is generated using the reference samples {a, b, c, d} that are chosen from the array of reference samples on the left (P_{rleft}) and above (P_{rabove}).

For the regions with homogeneous texture, DC prediction mode is preferred since it suitably adapts to the region. In such a scenario, the prediction value is generated by averaging the adjacently

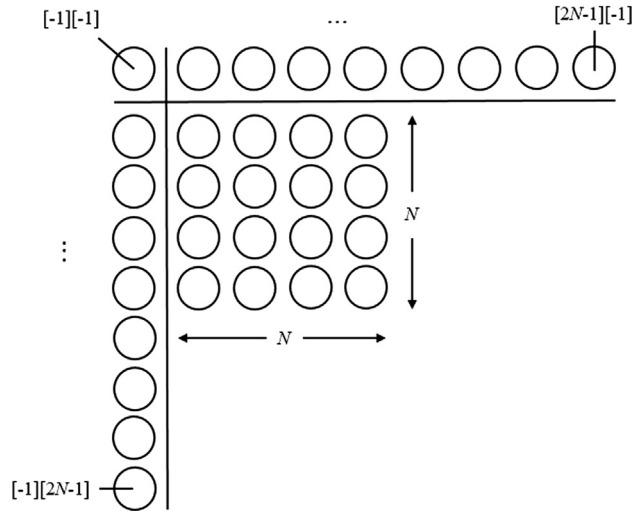


Fig. 1. Reference samples formation in HEVC intra prediction for an $N \times N$ block.

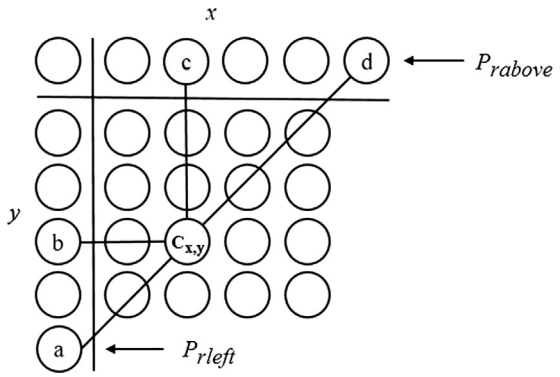


Fig. 2. Pixel prediction illustration in planar mode of HEVC.

available horizontal and vertical reference samples as specified in (1). The computation of the average DC value (DC_{val}) essentially requires $2N$ reference samples.

$$DC_{val} = \frac{1}{2N}\left(\sum_{x=0}^{N-1} P_{rabove}(x,-1) + \sum_{y=0}^{N-1} P_{rleft}(-1,y)\right) \quad (1)$$

The angular intra prediction in HEVC intends to efficiently model the various directional structures existing in the image/video content. The 33 angular prediction modes are categorized as vertical and horizontal modes. Modes 2–17 represent the horizontal modes, while modes 18–34 represent the vertical modes. The angularity of these modes is defined using the displacement parameter (θ_{disp}) specified as +/−{0, 2, 5, 9, 13, 17, 21, 26, 32} from the vertical/horizontal axis. Here, the numeric part indicates sample positional displacements in (1/32) fractions of sample grid positions. The parameter θ_{disp} is smaller when tending towards the horizontal/vertical directions to provide a more precise prediction for nearly horizontal/vertical patterns. Meanwhile, it becomes larger when tending towards the diagonal direction to minimize the prediction modes for less frequently occurring directions.

3. Related work

Similar to H.264/AVC, HEVC incorporates the conventional block-based hybrid framework for removal of the spatiotemporal redundancies within and across the frames [16–21]. Some of the existing sample-based intra prediction methods in literature are discussed below.

Sample-based angular prediction (SAP) technique [16,17] enhances the coding efficiency of the angular prediction within the HEVC in lossless mode by replacing the block-wise angular prediction process with the pixel-wise prediction strategy. The prediction value is generated from the row/column of pixels immediately above the target pixel based on whether the prediction mode is vertical/horizontal. Two reference samples i.e. R_i and R_{i+1} are used to generate the prediction value for the target pixel by means of linear interpolation. However, in case of the boundary pixel, the reconstructed pixel will be simply extended to serve as the R_{i+1} sample since it is unavailable for the prediction process, thereby making the effective angle of prediction to be zero. Also, this algorithm does not require any modification in syntax and semantics as all the samples within the PU share the same prediction angle and there is no change in the definition and signaling of the prediction angles.

The basic idea behind the sample-based weighted prediction (SWP) strategy [18] in producing a considerable gain in terms of coding efficiency is through computation of the prediction value based on the weighted average of the pixels encapsulated in the patch around the target pixel. Sum of Absolute Difference (SAD),

a measure of similarity is used to determine the weighted factor for the relative causal pixel in the patch. An exponentially decaying basis function is used to derive the weights. These results achieve a noticeable enhancement in the coding efficiency for the natural images and video sequences. The directional template matching (DTM) predictor is similar to the SWP predictor with the exception that the weighted factors need not be computed. Instead, the pixel with the estimated closest similarity measured using the SAD parameter is utilized for prediction [18]. However, this method provides significant improvement only for screen content video sequences.

The improved sample-based angular prediction (ISAP) [19] is beneficial for the boundary samples in terms of elimination of the zero angle prediction error that is otherwise prevalent in SAP. In this work, the R_{i+1} sample for prediction of the boundary pixel is determined using the previously generated neighboring sample in the same row/column for vertical and horizontal prediction respectively. A weighted prediction process is adopted for the generation of the prediction value using the reference samples R_i and R_{i+1}.

In combined intra prediction (CIP) [20], the coding efficiency is enhanced by combining the closed-loop prediction with the open-loop prediction component. In addition to prediction using the neighboring pixels, CIP utilizes the directional prediction from the adjacent blocks. The weighting factors for the open and closed-loop components are computed to generate the prediction value. This tool basically addresses the enhancement in the rate-distortion performance as well as the low-complexity requirements for codecs aiding high-resolution content.

The piece-wise mapping in HEVC for lossless mode [21] aims to reduce the energy of the intra predicted residual blocks which is imperative for the reconstruction process. This is achieved using the piece-wise mapping (PWM) function, wherein the range of values in the residual blocks are analyzed to apply a suitable PWM function to map the specific residual values to lower levels. All the associated PWM parameters are encoded and signaled to the decoder side, so as to reconstruct the mapped residual blocks error-free.

The pixel-by-pixel differential pulse code modulation (DPCM) referred as RDPCM in [22] intends to enhance the prediction accuracy by extracting the gradient information from the surrounding pixels. In the proposed algorithm, additional DPCM is applied based on the residuals obtained post applying the pixel-by-pixel DPCM. This method is applicable only for horizontal and vertical intra prediction modes. An extension of this work is proposed in [23], wherein additional prediction strategy is imposed on the residuals obtained after applying RDPCM.

HEVC based lossless compression of Whole Slide pathology Images (WSIs) is proposed in [24]. These images are highly textured containing several edges and multidirectional patterns, due to the existence of cellular structures and tissues. This serves as a robust motive for enhancement of the prediction performance, where the edge information is dominant. It is accomplished by embedding the sample-by-sample differential pulse code modulation (SbS-DPCM) and edge prediction strategy into the intra coding process.

In context-based predictive techniques, the context of each pixel is determined separately to obtain the optimal sub-predictors. Median edge detection (MED) predictor and gradient adaptive predictor (GAP) are amongst the primary predictors for edge detection.

The MED predictor suitably adapts in the presence of the local edges [25]. The horizontal/vertical edges are detected by examining the North *(N)*, West *(W)* and North-West *(NW)* directional neighbors of the pixel under consideration as specified in (2). It is obvious that pixel *N* and *W* serves as the prediction value in case of the vertical edge and horizontal edge detection respectively. Under the remaining scenarios, pixels *N*, *W* and *NW* are used to generate the prediction value.

$$P_{pred} = \begin{cases} min(N,W) & if \quad NW \geqslant max(N,W) \\ max(N,W) & if \quad NW \leqslant min(N,W) \\ N+W-NW & else \end{cases} \tag{2}$$

The GAP predictor is a simple, non-linear one that renders robust performance. The prediction process adapts itself based on the local gradients in the horizontal and vertical directions [26] using the seven directional neighborhood pixels *N*, *W*, *NW*, *NE*, *WW*, *NN* and *NNE*. Intensity gradient around the target pixel in horizontal and vertical direction is computed using (3) and (4) respectively. Based on the difference between the computed gradient values and predefined threshold, edges within the block can be categorized as sharp, normal or weak horizontal/vertical edges.

$$G_H = |W - WW| + |N - NW| + |N - NE| \tag{3}$$

$$G_V = |W - NW| + |N - NN| + |NE - NNE| \tag{4}$$

As the resolution of the video sequences increase, larger regions with smooth variations as well as regions with several intensity variations become quite prominent. For efficient compression, the chosen prediction technique must be capable to tune the entire prediction process based on the local region characteristics. Sample-based prediction methods such as SWP and DTM predictors accomplish the same, but does not cater to the extraction of the edge details. Hence, context-based predictive techniques are designed to identify the existence of vertical/horizontal edges and carry out the prediction process based on the gathered gradient information around the target pixel. However, in cases where a diagonal edge exists, it could result in propagation of the prediction errors while traversing through the various stages during encoding. These shortcomings serve as motivation for Section 4.

4. Proposed method

In this section, we propose better prediction strategies for both planar and angular intra prediction modes - (1) threshold controlled gradient adaptive planar prediction (TGAPP) and (2) gradient-oriented selection and sample-based weighted angular prediction (GSSWAP). The TGAPP and GSSWAP, jointly referred as gradient-oriented directional prediction (GDP) intends to enhance the overall coding efficiency of the planar and angular intra prediction modes of HEVC. TGAPP, a planar prediction modification features a low complexity mechanism to identify the sharp edges within the frame. It primarily computes the gradient values in the vertical and horizontal directions for each of the samples within the block. Secondly, the prediction value is derived based on the threshold value (*T*). Meanwhile, GSSWAP aims to enhance the prediction quality of the directional structures within the given frame. Initially, the average gradients around the pixel under consideration are estimated in the four specified directions: horizontal (0°), vertical (90°), diagonal (45°) and anti-diagonal (135°). Then the weighted averages of the nearby causal pixels corresponding to the directional gradient are computed to generate the prediction value for the target pixel.

4.1. TGAPP predictor for planar prediction

Threshold controlled gradient adaptive planar prediction (TGAPP) is a simple blend of the gradient and median predictors. Local gradient estimation and a predefined threshold are essential to obtain an optimal predictor from the sub-predictors, i.e. the prediction depends on whether the pixel is in context of a horizontal edge, vertical edge or the smooth region. The pseudocode presented in Algorithm 1 replaces the conventional planar mode embedded in the HEVC anchor.

Algorithm 1. Pseudocode for TGAPP intra prediction

Compute vertical (G_V) and horizontal (G_H) gradients:

$$G_V = |NW - W| + |NN - N|$$

$$G_H = |WW - W| + |NW - N|$$

if (G_V - $G_H > T$) **then**
  P_{pred} = W
else if (G_V - $G_H < -T$) **then**
  P_{pred} = N
else
  P_{pred} = (N + W - NW)
end if

In the first phase, pixel level computation of the vertical (G_V) and horizontal (G_H) gradients relative to the target pixel $C_{x,y}$ is performed using the five causal pixels located at *N, W, NW, NN* and *WW* directional positions as shown in Fig. 3.

Gradient computation is highly essential to capture the level of activity in terms of smoothness, edginess etc. around the target pixel, thus governing the statistical behavior of the prediction residuals. The prediction value is then derived, based on the comparison between the gradient difference and threshold *T*. The value of *T* was chosen to be 80 after the experimental analysis on the basis of [26]. This value was later fixed for the set of HEVC test sequences chosen for algorithm evaluation.

4.2. GSSWAP predictor for angular prediction

A template consisting of ten causal pixels is utilized for gradient estimation around the target pixel ($C_{x,y}$) as elaborated in Figs. 4 and 5 for vertical and horizontal angular prediction respectively. Here, the neighboring pixels are represented using alphabets A–F and the four directions as direction 1–4, for ease of understanding.

Initially, the gradients are estimated by interpolation of the absolute difference between the causal pixels in the specified four directions using (5)–(8).

Direction 1:

$$D_1 = |A - E| + |C - B| + |C - D| + |B - F| \quad (5)$$

Direction 2:

$$D_2 = |A - B| + |C - I| + |D - J| + |B - H| \quad (6)$$

Direction 3:

$$D_3 = |A - C| + |C - J| + |E - B| + |B - I| \quad (7)$$

Direction 4:

$$D_4 = |A - F| + |C - H| + |D - I| + |B - G| \quad (8)$$

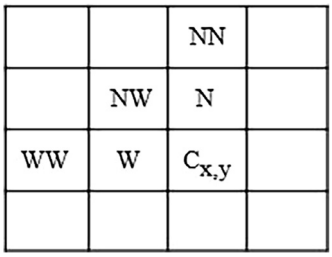
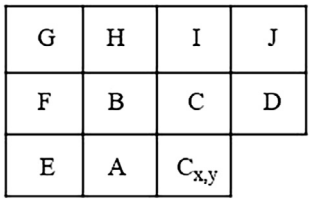


Fig. 3. Causal samples for gradient computation in TGAPP.



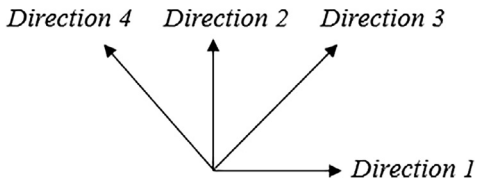


Fig. 4. Causal pixel template for vertical prediction.

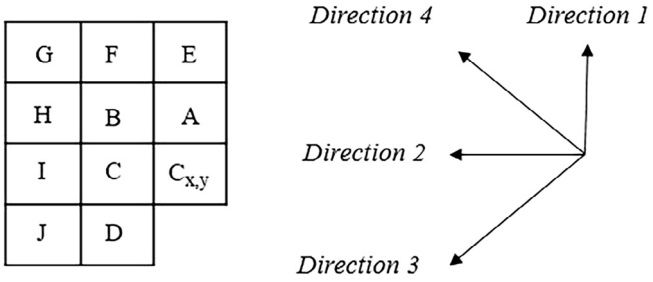


Fig. 5. Causal pixel template for horizontal prediction.

The estimated gradient values are averaged in order to derive the final gradient values using (9)–(12).

$$G_1 = \lfloor D_1/4 + 0.5 \rfloor \quad (9)$$

$$G_2 = \lfloor D_2/4 + 0.5 \rfloor \quad (10)$$

$$G_3 = \lfloor D_3/4 + 0.5 \rfloor \quad (11)$$

$$G_4 = \lfloor D_4/4 + 0.5 \rfloor \quad (12)$$

After the estimation of the gradient values, the prediction value is generated. This is accomplished by assigning a higher weight to the immediate causal pixel in the direction of the lower estimated gradient value. The pixel at the minimal pixel distance from the target pixel will be the relative pixel to the corresponding directional gradient i.e. the relative causal pixel to the gradients G_1, G_2, G_3, G_4 are A, C, D, and B respectively for vertical/horizontal angular predictions as depicted in Figs. 4 and 5.

While the displacement parameter (θ_{disp}) is zero or ±32 and only one of the gradient being non-zero, the relative causal pixel in the direction perpendicular to the current directional gradient serves as the prediction value. In case of more than one non-zero gradients, the prediction value is computed based on the generation of the sample-based weights corresponding to the relative causal neighbor in the specified four directions of the gradient. In all the remaining cases, the SAP prediction algorithm is applied to the target pixel to generate the prediction value as explained in Section 4.3.

In SWP prediction, the weighted averaging of the surrounding pixels is performed to predict the target pixel. The prediction value is realised using (13) as stated in [18].

$$P_{pred} = \left\lfloor \frac{\sum_{i \in S} wt_{int}[i,j] \cdot h[i]}{\sum_{i \in S} wt_{int}[i,j]} \right\rfloor \quad (13)$$

where P_{pred} is the generated prediction, $h[i]$ is the surrounding pixels comprising of four previously reconstructed pixels {A, B, C, D} in close proximity to the target pixel, S represents the patch of pixels considered for prediction and wt_{int} is the integer weights of the relative causal pixels $h[i]$. Eq. (14) is used to compute the integer weights.

$$wt_{int}[i,j] = \left\lfloor a \cdot b^{-SAD\left(\frac{P_h[i],P_h[j]}{h_{dist}}\right)} \right\rfloor \tag{14}$$

Here, a is a factor that typically ensures the major weights are non-zero, b is the basis of the exponentially decaying weights, P_h [i] is a patch around the pixel $h[i]$, P_h [j] is a patch around the pixel $h[j]$ and h_{dist} is a prediction modeling parameter. The value of b is chosen to be 2 and a is derived from (15). Here, β denotes the internal bit-depth of the video sequence.

$$a = 2^{31-2-\beta} \tag{15}$$

The prediction modeling parameter h_{dist} is dependent on the sub-sampling format of the chroma component. The h_{dist} value for the luma and chroma component is empirically chosen as 5.25 and 3.25 respectively. The SAD is an operator for computing the sum of absolute differences between the two patches associated with the four immediate causal pixels, which is calculated as in (16).

$$SAD(P_h[i], P_h[j]) = \sum_N |h[i] - h[j]| \tag{16}$$

The patches for the computation of SAD value in case of vertical prediction of the causal pixel C with respect to the target pixel $C_{x,y}$ is as shown in Fig. 6. For the specified causal pixel, the value is computed using the sum of absolute differences between the two co-located pixels within the patch i.e. $P_h[i]$ = {A, B, C, D} and P_h[j] = {B, H, I, J}. Here, N refers to the number of pixels in the considered patch.

In cases when the computed weights are zero, the prediction is generated using the causal pixels corresponding to the two minimum valued directional gradients. The minimum valued gradients are determined using (17) and (18).

$$G_{min1} = min(G_i), \qquad i = 1,2,3,4 \tag{17}$$
$$G_{min2} = min(G_j)|j \neq i, \qquad j = 1,2,3,4 \tag{18}$$

Once the minimum directional gradients are obtained [27], then C_{Gmin1} and C_{Gmin2} needs to be determined as they represent the relative causal pixels corresponding to the two minimal directional gradients G_{min1} and G_{min2} respectively. For example, if G_{min1} = G_1 and G_{min2} = G_3, then the corresponding causal pixel value C_{Gmin1} = A and C_{Gmin2} = D. The prediction of the target pixel is based on these values and computed using (19). From this, it is evident that the causal pixel relative to the second minimal directional gradient makes a predominant contribution to the prediction generation.

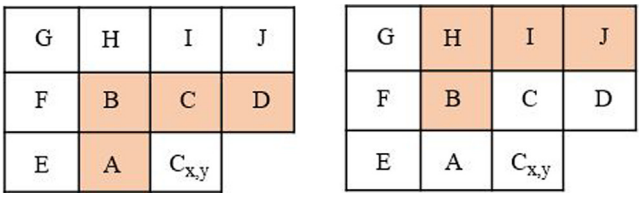


Fig. 6. Patch around the target pixel $C_{x,y}$ and casual pixel C corresponding to a directional gradient.

$$P_{pred} = \frac{G_{min1} \cdot C_{Gmin2} + G_{min2} \cdot C_{Gmin1}}{G_{min1} + G_{min2}} \tag{19}$$

4.3. Boundary analysis for GSSWAP predictor

In GDP, when one of the two rows/columns of pixels essential for prediction of the target pixel in the first row/left column of the current block is unavailable, the prediction is done as in the SAP prediction technique. This strategy is an enhancement to the conventional block-based intra prediction technique embedded within the HEVC anchor. It is observed that the reconstructed pixels provide a precise prediction of the pixel to be predicted. Hence, the current row/column of pixels are predicted from the previously reconstructed immediate row/column of pixels. Every pixel is predicted using two reconstructed samples R_i and R_{i+1}. In general, to predict a single row/column of an $N \times N$ block, $N + 2$ reference samples are essential.

In the proposed method, the boundary pixels i.e. pixels in the first row and column of the current block are predicted from the previous row/column of the already reconstructed neighboring blocks, which are on the top/left of the block under consideration. Parameter *Fract* is used to determine the two reconstructed reference samples i.e. R_i and R_{i+1} by projecting the target pixel onto the reference row/column as described in Fig. 7.

Once the reference samples are chosen, then the prediction value is computed using (20) which is a linear interpolation of R_i and R_{i+1}. While the parameter *Fract* represents the distance measured at (1/32) pixel accuracy between the generated prediction value (P_{pred}) and the reference pixel (R_i), the operator $\gg$ denotes the right bit-shift operation.

$$P_{pred} = ((32 - Fract) * R_i + Fract * R_{i+1} + 16) \gg 5 \tag{20}$$

When the boundary pixels are in the last row/column, the prediction strategy simply extends the boundary samples to compensate for the unavailable reference pixels during prediction. The pseudocode in Algorithm 2 summarizes the integer implementation of the proposed method.

Algorithm 2. Pseudocode for GSSWAP intra prediction

$Fract = \theta_{disp}$ &(32–1)
if $(x|y = 0)$ **then**
  $P_{pred} = ((32 - Fract) * R_i + Fract * R_{i+1} + 16) \gg 5$
else
  $D_1 = |A - E| + |C - B| + |C - D| + |B - F|$
  $D_2 = |A - B| + |C - I| + |D - J| + |B - H|$
  $D_3 = |A - C| + |C - J| + |E - B| + |B - I|$
  $D_4 = |A - F| + |C - H| + |D - I| + |B - G|$
  Obtain averaged gradient values: G_1, G_2, G_3 and G_4
  if $(\sum_{i=1}^{4} G_i \neq 0 \; and \; \theta_{disp} = (0 \, or \pm 32))$ **then**
    if $(G_i \neq 0|_{i=n} \, and \sum_{i \neq n}^{4} G_i = 0)$ **then**
      $P_{pred} = h[j]; where j \perp i$
    else
      if $(wt\_A + wt\_B + wt\_C + wt\_D == 0)$ **then**
        $P_{pred} = \frac{G_{min1} \cdot C_{Gmin2} + G_{min2} \cdot C_{Gmin1}}{G_{min1} + G_{min2}}$
      else
        $P_{pred} = \frac{A.wt\_A + B.wt\_B + C.wt\_C + D.wt\_D}{wt\_A + wt\_B + wt\_C + wt\_D}$
      end if
    end if
  else
    $P_{pred} = ((32 - Fract) * R_i + Fract * R_{i+1} + 16) \gg 5$
  end if
end if

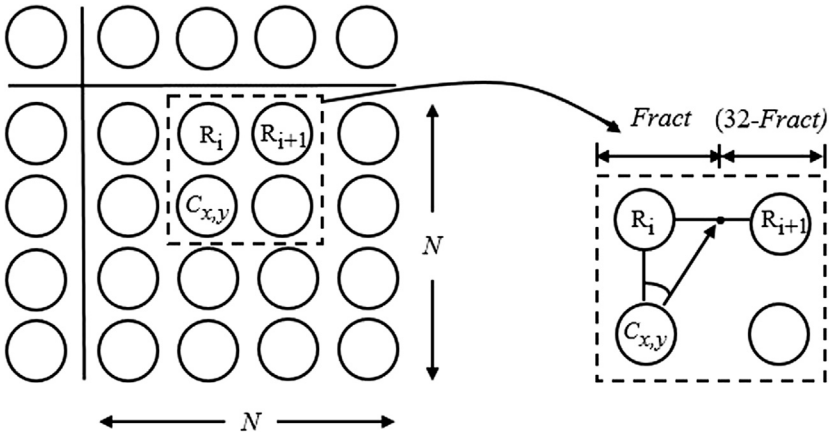


Fig. 7. Prediction strategy in sample-based angular prediction for positive vertical prediction angle.

5. Simulation results and discussion

The proposed GDP method for enhancement of coding efficiency of the planar and angular intra prediction modes is implemented using HEVC Test Model (HM) software-HM 16.12 [28]. In the paper, the unmodified HM 16.12 is referred as the HEVC anchor. The algorithmic validation is performed under the standard common test conditions [29] for lossless compression using the All-Intra (AI), Low Delay (LD), Low Delay P (LDP) and Random Access (RA) configurations. The Main (low computational complexity) profile of the HM software is the test setting considered for analysis. It defines a set of coding tools that are to be used to create a compliant bitstream.

The YUV video sequences (4:2:0 color format with 8 bits/color sample) covering a wide range of resolutions and possessing rich textural and motion properties [30] are chosen for evaluation of the compression performance of the proposed method. The attributes of the test sequences are elaborated in Table 1. To ensure a fair comparison, we have considered fourteen video sequences ranging from Class A–E.

In the HEVC lossless mode, transformation, quantization and other associated in-loop filtering operations are bypassed to ensure a perfect lossless reconstruction. Hence, the residual samples obtained using the block-based intra prediction are directly fed to the entropy coding block. Lossless mode can be chosen by setting the two relevant flags, specified in the configuration settings. The sample-based TGAPP and GSSWAP predictors replace the conventional planar and angular block-based intra prediction modes respectively, embedded within the HEVC anchor.

The simulation is performed considering the initial hundred frames from each of the video sequences. As the quantization stage is bypassed in the lossless mode, the Quantization Parameter (QP) is set to 0 for the Main (8-bit) profile. The performance validation of the proposed algorithm and the state-of-the-art prediction techniques is done in comparison with the HEVC anchor. The analysis is based on the simulation results obtained in terms of total bit-rate (BR) savings and run-time (RT) savings computed using (21) and (22) respectively. The numerical value obtained on computation is considered positive since *savings* indicate gain. The simulations are carried out on a 2.30 GHz Intel (R) Core (TM) i5-6200U processor with 8 GB RAM, Windows 10 OS.

$$BR\ Savings = \frac{(BR_{anchor} - BR_{proposed})}{BR_{anchor}} * 100\% \quad (21)$$

$$Time\ Savings = \frac{(RT_{anchor} - RT_{proposed})}{RT_{anchor}} * 100\% \quad (22)$$

The prediction strategies like SAP [16,17], SWP [18], DTM [18] and ISAP [19] are considered to evaluate the performance of the proposed algorithm. From Tables 2 and 3, it is obvious that the proposed method in comparison with the HEVC anchor provides a significant enhancement in the coding efficiency by about 8.29%, 1.65%, 1.94% and 2.21% on an average, in terms of overall bit-rate savings for Main-AI, LD, LDP and RA configurations respectively. A higher percentage of improvement in coding efficiency is observed for the Main-AI configuration compared to the Main-LD, LDP and RA. This is due to the fact that the Main-AI configuration utilizes only spatial correlation, while the Main-LD, LDP and RA configurations use spatiotemporal correlation. Also from the tabular analysis, it is evident that the proposed method outperforms the SAP, SWP, DTM and ISAP prediction strategies in terms of overall bit-rate savings derived with respect to the HEVC anchor for Main-AI, LD, LDP and RA configurations respectively as depicted in Fig. 8.

The number of blocks coded with different PU sizes, ranging between 4×4 and 32×32 implicitly signifies the enhancement in prediction accuracy for the GDP predictor. To substantiate this, we computed the number of blocks of different sizes considering the first frame of one representative video sequence from each class as presented in Table 4 for the AI-Main configuration. This tabulation remains the same for Main-LD, LDP and RA as every first frame in each of the four configurations is intra coded. From the analysis, it is evident that using the proposed GDP predictor, more number of pixels favor larger PU sizes which is graphically elaborated in Fig. 9. This implies that there is a substantial reduction in the number of blocks to be coded using the proposed algorithm, as compared to the HEVC anchor to achieve additional compression.

Table 5 provides an insight into the distribution of the intra prediction modes. The video sequences chosen for analysis possess rich textural and directional properties, resulting in nearly $(2/3)^{rd}$ of the total blocks being predicted using the available 33 angular modes (M2-M34), followed by planar (M0) and DC (M1) modes.

Table 1
Details of the test video sequences.

Class	Video resolution	Category	No. of sequences
A	2560×1600	4 K × 2 K ultra-HD at 30 fps (cropped)	2
B	1920×1080	1080p HD at 24, 60 fps	3
C	832×480	WVGA at 30, 50, 60 fps	3
D	416×240	WQVGA at 30,50, 60 fps	3
E	1280×720	720p video conferencing at 60 fps	3

Table 2
Comparison of bit-rate savings (%) using the state-of-the-art prediction strategies and proposed GDP predictor with the HM 16.12 lossless coding for Main-AI and Main-RA configurations.

Class	Video sequence	AI					RA				
		SAP	SWP	DTM	ISAP	GDP	SAP	SWP	DTM	ISAP	GDP
A	Traffic	11.06	7.79	7.01	11.18	11.44	1.82	0.78	0.93	1.86	1.97
	People on Street	10.32	8.99	6.47	10.52	11.1	4.48	3.62	2.43	4.6	4.96
B	Kimono	7.18	6.83	3.02	7.33	7.82	3.21	2.99	0.99	3.28	3.57
	Park Scene	6.65	5.26	2.78	6.76	6.94	0.82	0.63	0.22	0.83	0.85
	BQ Terrace	6.57	2.46	3.96	6.6	6.75	2.68	0.2	1.64	2.67	2.69
C	Race Horses	7.19	5.1	3.6	7.25	7.59	4.06	2.34	1.82	4.12	4.22
	BQ Mall	5.75	4.53	3.31	5.82	5.95	0.89	0.64	0.42	0.9	0.93
	Party Scene	4.58	3.89	2.74	4.64	4.86	1.07	0.95	0.52	1.1	1.16
D	Blowing Bubbles	6.39	5.61	3.19	6.48	6.84	1.19	1.09	0.6	1.22	1.3
	Race Horses	9.27	8.52	6.27	9.35	10.15	3.43	2.64	1.83	3.48	3.72
	BQ Square	3.61	2.59	2.41	3.62	3.97	0.24	0.17	0.16	0.24	0.26
E	Vidyo 1	11.53	5.56	5.49	11.6	11.83	1.88	0.64	0.6	1.92	1.98
	Vidyo 3	9.61	5.56	5.49	9.68	10.23	1.34	0.56	0.54	1.4	1.47
	Vidyo 4	10.07	5.35	5.23	10.15	10.57	1.75	0.75	0.56	1.8	1.89
Avg. Bit-rate Savings (%)		7.84	5.57	4.36	7.93	8.29	2.06	1.29	0.95	2.1	2.21

Table 3
Comparison of bit-rate savings (%) using the state-of-the-art prediction strategies and proposed GDP predictor with the HM 16.12 lossless coding for Main-LD and Main-LDP configurations.

Class	Video sequence	LD					LDP				
		SAP	SWP	DTM	ISAP	GDP	SAP	SWP	DTM	ISAP	GDP
A	Traffic	1.33	0.41	0.61	1.34	1.44	1.9	0.62	0.9	1.96	2.06
	People on Street	4.1	3.26	2.13	4.22	4.57	4.5	3.57	2.32	4.6	5.0
B	Kimono	2.9	2.63	0.76	3.02	3.22	3.8	3.35	0.86	3.92	4.17
	Park Scene	0.45	0.33	0.06	0.45	0.46	0.62	0.46	0.05	0.62	0.64
	BQ Terrace	2.34	0.09	1.42	2.35	2.35	2.67	0.16	1.61	2.67	2.68
C	Race Horses	3.9	2.15	1.69	3.94	4.04	4.18	2.36	1.8	4.22	4.34
	BQ Mall	0.52	0.39	0.24	0.54	0.58	0.66	0.46	0.25	0.7	0.75
	Party Scene	0.79	0.7	0.34	0.82	0.86	0.84	0.73	0.37	0.86	0.91
D	Blowing Bubbles	0.83	0.77	0.42	0.85	0.9	0.86	0.81	0.45	0.87	0.93
	Race Horses	3.01	2.16	1.46	3.1	3.26	3.18	2.3	1.53	3.22	3.43
	BQ Square	0.04	0.05	0.04	0.05	0.07	0.04	0.03	0.02	0.06	0.1
E	Vidyo 1	0.31	0.16	0.12	0.3	0.31	0.47	0.21	0.15	0.47	0.48
	Vidyo 3	0.28	0.14	0.14	0.28	0.29	0.34	0.17	0.19	0.34	0.36
	Vidyo 4	0.75	0.35	0.25	0.76	0.81	1.28	0.61	0.34	1.3	1.37
Avg. Bit-rate Savings (%)		1.54	0.97	0.69	1.57	1.65	1.81	1.13	0.77	1.84	1.94

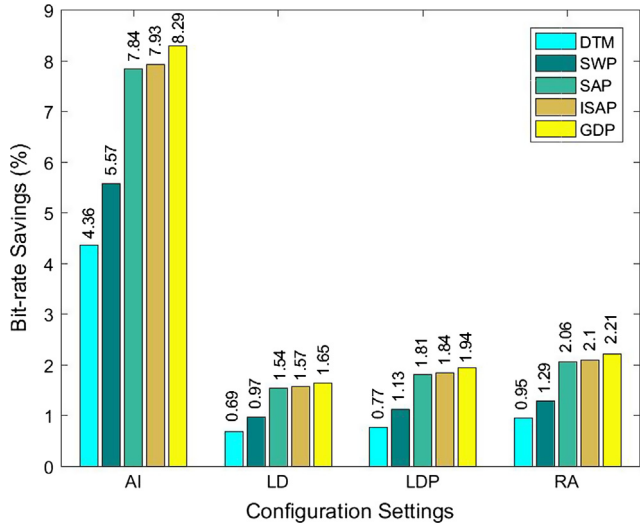


Fig. 8. Bit-rate savings (%) using the state-of-the-art prediction strategies and proposed GDP predictor in comparison with the HM 16.12 lossless coding for the Main-AI, LD, LDP and RA configurations.

Enhancement in prediction accuracies of the planar and angular intra prediction can also be observed in form of a substantial increase in the pixels favoring the GDP prediction strategy. For the AI-Main configuration, the net increase in the number of blocks predicted using the proposed algorithm is around 1.7%, thus contributing to the prominent enhancement in the bit-rate savings.

The run-time results at the encoder and decoder denoted as ET and DT respectively, are subtly less than or at par with the HEVC anchor. The run-time savings for the Main-AI, LD, LDP and RA configurations are tabulated in Table 6. It clearly indicates that the complexity of the algorithm has not been affected due to the proposed modification.

The gradient-based computation in the proposed GDP algorithm is expected to minimally increase the encoding time compared to the HEVC anchor. But the enhanced prediction accuracy results in smaller residuals values, thus lowering the bit-rate. The bit-rate reduction results in swift entropy coding for all the sample-based prediction techniques. Lesser entropy coding time thus serves to compensate the increase in the encoding time, which is due to the increased computation involved in gradient

Table 4
Block count of various block sizes coded using HM 16.12 lossless coding and proposed GDP predictor for Main-AI configuration.

Class	Video sequence	Anchor					Proposed				
		4 × 4	8 × 8	16 × 16	32 × 32	Total	4 × 4	8 × 8	16 × 16	32 × 32	Total
A	Traffic	233,240	5458	54	1	238,753	76,456	20,974	3014	741	101,185
B	Kimono	110,124	4565	72	1	114,762	10,600	9118	1858	825	22,401
C	BQMall	20,236	1025	35	1	21,297	12,496	1084	324	46	13,950
D	RaceHorses	6040	50	0	0	6090	2636	529	57	9	3231
E	Vidyo1	51,200	1440	36	1	52,677	20,912	3620	960	107	25,599

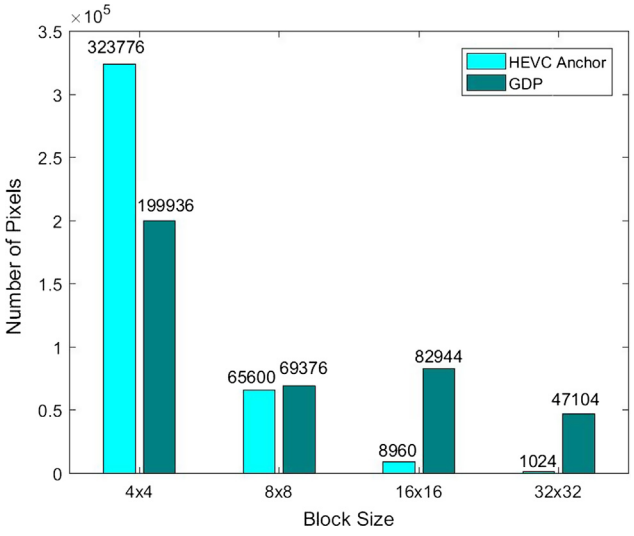


Fig. 9. Comparison of the number of pixels represented under the various block sizes in HM 16.12 lossless coding and proposed GDP predictor for the video sequence BQMall.

determination. Meanwhile, the same holds good at the decoder side as well.

6. Conclusion

In the proposed work, we present an efficient gradient-oriented directional prediction strategy to enhance the compression efficiency of the HEVC anchor in the lossless mode. The experimental results show that the proposed GDP predictor is able to minimize the bit-rate by 8.29%, 1.65%, 1.94% and 2.21% for Main-AI, LD, LDP and RA configurations respectively in comparison with the HEVC anchor. It also lucidly illustrates how the proposed GDP predictor outperforms that state-of-the-art techniques present in the literature. The reduction in the number of blocks coded using the proposed method implicitly indicates the improvement in prediction accuracy, that can trigger the further reductions in the overall bit-rate. This suitably helps to minimize the bit-rate while maintaining the computational complexity at the encoder and decoder slightly less than or at par with the HEVC anchor. The highlight of the proposed algorithm is that there is a significant enhancement in prediction accuracy and coding efficiency without any additional latency.

Table 5
Distribution (%) of intra prediction modes using proposed GDP predictor in comparison with the HM 16.12 lossless coding for Main-AI configuration.

Class	Video sequence	Anchor			Proposed		
		M0	M1	M2-M34	M0	M1	M2-M34
A	Traffic	12.8	6.1	81.1	11.1	1.7	87.2
	People on Street	12.1	7.2	80.7	12.4	2.2	85.4
B	Kimono	21.1	8.7	70.2	4.9	11.8	83.3
	Park Scene	15.7	11.1	73.2	7.7	12.9	79.4
	BQ Terrace	8.6	6.5	84.9	15.6	10.3	74.1
C	Race Horses	15.8	7.9	76.3	9.4	4.7	85.9
	BQ Mall	11.9	7.8	80.3	9.1	13.4	77.5
	Party Scene	9.7	6.7	83.6	7.7	7.0	85.3
D	Blowing Bubbles	11.4	5.8	82.8	7.2	2.9	89.9
	Race Horses	10.7	6.1	83.2	10.5	3.5	86.0
	BQ Square	7.4	5.9	86.7	8.2	5.6	86.2
E	Vidyo 1	15.1	8.7	76.2	22.0	2.6	75.4
	Vidyo 3	14.2	10.9	74.9	22.7	3.4	73.9
	Vidyo 4	16.0	9.4	74.6	20.1	3.4	76.5
Avg. (%)		13.0	7.8	79.2	12.0	6.1	81.9

Table 6
Run-time savings (%) for the proposed GDP predictor in comparison with HM 16.12 lossless coding for Main-AI, LD, LDP and RA configurations.

Class	AI		LD		LDP		RA	
	ET	DT	ET	DT	ET	DT	ET	DT
A	0.23	0.82	0.13	0.48	0.10	0.12	−0.20	0.05
B	0.22	0.16	−0.15	1.08	−0.40	0.26	0.24	0.14
C	1.38	0.45	0.90	0.24	−0.35	−0.15	0.06	0.18
D	0.44	1.34	0.92	0.14	0.22	0.06	0.18	0.20
E	0.78	1.26	1.02	0.46	0.40	0.30	−0.12	0.10
Avg. (%)	0.61	0.81	0.56	0.48	−0.01	0.12	0.03	0.13

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.aeue.2018.07.037.

References

[1] Sullivan GJ, Ohm JR, Han WJ, Wiegand T. Overview of the high efficiency video coding (HEVC) standard. IEEE Trans Circuits Syst Video Technol 2012;22(12):1648–67.
[2] Hang HM, Peng WH, Chan CH, Chen CC. Towards the next video standard: high efficiency video coding. In: Asia-Pacific signal and information processing association annual summit and conference (APSIPA); 2010. p. 609–18.
[3] Ohm JR, Sullivan GJ. High efficiency video coding: the next frontier in video compression [standards in a nutshell]. IEEE Signal Process Mag 2013;30(1):152–8.
[4] Richardson IE. The H.264 advanced video compression standard. John Wiley; 2010.
[5] Yang G, Li J, He Y, Kang Z. An information hiding algorithm based on intra-prediction modes and matrix coding for H.264/AVC video stream. AEU-Int J Electron Commun 2011;65(4):331–7.
[6] Dai W, Krishnan M, Topiwala J. AHG22: Lossless transforms for lossless coding. JCTVC-G268, Geneva; 2011.
[7] Antony A, Ganapathy S. Highly efficient near lossless video compression using selective intra prediction for HEVC lossless mode. AEU-Int J Electron Commun 2015;69(11):1650–8.
[8] Sayood K. Lossless compression handbook. Academic Press; 2002.
[9] Kamath SS, Aparna P, Antony A. Sample-based DC prediction strategy for HEVC lossless intra prediction mode. In: Fourth international conference in image information processing (ICIIP); 2017. p. 1–5.
[10] Sze V, Budagavi M, Sullivan GJ. High efficiency video coding (HEVC): algorithms and architectures. Switzerland: Springer; 2014.
[11] Wein M. High efficiency video coding: coding tools and specification. Berlin: Springer; 2015.
[12] Li X, He X, Peng X, Xiong S. An image feature-based method to efficiently determine inter-coding depth in HEVC. AEU-Int J Electron Commun 2017;71:96–104.
[13] Pourazad MT, Doutre C, Azimi M, Nasiopoulos P. HEVC: the new gold standard for video compression: how does HEVC compare with H.264/AVC? IEEE Consumer Electron Mag 2012;1(3):36–46.
[14] Lainema J, Bossen F, Han WJ, Min J, Ugur K. Intra coding of the HEVC standard. IEEE Trans Circuits Syst Video Technol 2012;22(12):1792–801.
[15] Lucas LFR, da Silva EAB, de Faria SMM, Rodrigues NMM, Pagliari CL. Prediction techniques for image and video coding. In: Efficient predictive algorithms for image compression. Springer; 2017. p. 7–33.
[16] Zhou M. Sample-based angular prediction (SAP) for HEVC lossless coding. JCTVC-G093, Geneva; 2011.
[17] Zhou M, Gao W, Jiang M, Yu H. HEVC lossless coding and improvements. IEEE Trans Circuits Syst Video Technol 2012;22(12):1839–43.
[18] Wige E, Yammine G, Amon P, Hutter A, Kaup A. Sample-based weighted prediction with directional template matching for HEVC Lossless Coding. In: Proceedings of picture coding symposium PCS; 2013. p. 305–8.
[19] Antony A, Sreelekha G. HEVC-based lossless intra coding for efficient still image compression. Multimedia Tools Appl 2017;76(2):1639–58.
[20] Gabriellini A, Flynn D, Mrak M, Davies T. Combined intra-prediction for high-efficiency video coding. IEEE J Sel Top Signal Process 2011;5(7):1282–9.
[21] Sanchez V, Auli-Llinas F, Serra-Sagrista J. Piecewise mapping in HEVC lossless intra-prediction coding. IEEE Trans Image Process 2016;25(9):4004–17.
[22] Jeon G, Kim K, Jeong J. Improved residual DPCM for HEVC lossless coding. In: SIBGRAPI conference on graphics, patterns and images; 2014. p. 95–102.
[23] Kim K, Jeon G, Jeong J. Improvement of Implicit Residual DPCM for HEVC. In: Tenth international conference on signal-image technology and internet-based systems; 2014. p. 652–58.
[24] Sanchez V, Auli-Llinas F, Bartrina-Rapesta J, Serra-Sagrista J. HEVC-based lossless compression of whole slide pathology images. In: IEEE global conference on signal and information processing; 2014. p. 297–301.
[25] Weinberger MJ, Seroussi G, Sapiro G. LOCO-I: a low complexity, context-based, lossless image compression algorithm. In: IEEE data compression conference proceedings; 1996. p. 140–9.
[26] Wu X, Memon N. Context-based, adaptive, lossless image coding. IEEE Trans Commun 1997;45(4):437–44.
[27] Knezovic J, Kovac M. Gradient based selective weighting of neighboring pixels for predictive lossless image coding. In: IEEE Proceedings of the 25th international conference on information technology interfaces; 2003. p. 483–8.
[28] HEVC Reference Software [online]. 2012. [Available at: <https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/HM-16.12/>].
[29] Bossen F. Common test conditions and software reference configurations. JCTVC-G1100, San Jose; 2012.
[30] Test Sequences. 2012. [Available at:ftp://ftp.tnt.uni-hannover.de].