

Review on Intra-prediction in High Efficiency Video Coding (HEVC) Standard

Dhruvi Patel

PG Student, Department Of
Electronics & Communication
Engineering,
C.G. Patel Institute of
Technology, Uka Tarsadia
University, Bardoli,
Gujarat

Tarun Lad

Department Of Electronics &
Communication Engineering,
C.G. Patel Institute of
Technology, Uka Tarsadia
University, Bardoli
Gujarat

Dharam Shah

Department Of Electronics &
Communication Engineering,
C.G. Patel Institute of
Technology, Uka Tarsadia
University, Bardoli
Gujarat

ABSTRACT

The objective of this paper is to provide a review of the intra prediction part in recently developed HEVC standard. HEVC (H.265) standard is the latest enhanced video coding standard which was planned to improve the rendered specifications of its preceding standard MPEG-4 (H.264). The main goal of the HEVC standardization effort is to enable significantly improved compression performance relative to existing standard H.264. For similar video quality, HEVC bit-streams consume only about half of the bitrate compare to previous standard H.264. Any video possess redundant bits in every frame. For remove this redundancy prediction process is used. Intra-picture prediction is a tool in HEVC which uses some prediction of data spatially from region to region within a specific picture, but has no dependence on other pictures in the video frames. HEVC gives higher compression comparing with previous standard H.264 because of its new features like quadtree structure, more directional intra-prediction modes. HEVC suitable for resolutions up to Ultra High Definition (UHD) video coding in the future.

General Terms

Angular modes, DC mode, Partition, Planner mode, Prediction, Reference samples.

Keywords

AVC, H.264/MPEG-4, HEVC, Intra-prediction.

1. INTRODUCTION

Demands of video applications such as video telephony, high definition television (HDTV), video conferencing, online streaming of video, streaming on mobile, TV, 24 hours surveillance system and many others are increasing exponentially, to fulfil this demand continuous evolution in video compression standard [1] is required for more compression. The video data compression is the process of reducing the amount of data required to represent a video sequence. Video coding is a process of video encoding and decoding together. New video compression standard needed to satisfy growing demand of video applications and it must be more effective than previously developed standards. There are two organization namely ITU-T (VCEG) (international telecommunication union – telecommunications (video coding experts group)) and ISO (international standardization organization)/IEC (international electro-technical committee) MPEG (moving picture experts group), they introduce many video coding standards. ITU-T (VCEG) (international telecommunication union – telecommunications (video coding experts group)) developed H.120 standard – the first ever video coding standard in 1984 and then developed H.261

standard in 1990. ISO/IEC started MPEG project in 1988 and developed MPEG-1. ITU-T (VCEG) and ISO/IEC (MPEG) jointly developed MPEG-2/H.262 standard approved in 1993. In 1995, ITU-T developed H.263 and ISO/IEC (MPEG) introduced MPEG-4 in 1999. H.264 video codec was announced by ITU-T and ISO/IEC around 2004. Finally, ITU-T (VCEG) and ISO/IEC (MPEG) collaborated and formed the Joint Collaborative Team on Video Coding (JCT-VC) in April 2010 to develop the next-generation video coding (NGVC) standard HEVC (High Efficiency Video Coding), which is approved as coding standard in 2013. This all standards follow the same block base hybrid coding structure.

Organization of this paper is organized as follow: Section 2- process the basic traditional hybrid block structure is given. Section 3- intra-prediction process is given in detail. Section 4- compression of intra prediction in HEVC and H.264/MPEG is reviewed. Section 5- conclusion is given

2. HYBRID CODEC

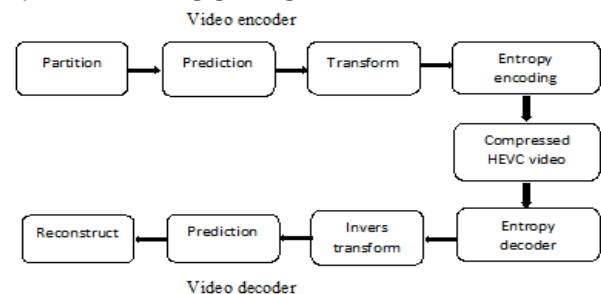


Fig 1: Block based Hybrid codec [2]

All video coding standards from H.120 and above follows one common traditional codec architecture [2] this codec can be spited in two major parts: 1) encoder, 2) decoder. In the encoding part, every frame is partitioned in to same size of picture blocks and next step is apply on each and every picture block. This picture blocks are known as macroblocks for coding standards from H.120 to H.264. In HEVC picture blocks are known as coding tree unit (CTU). In the next step, each unit is being predicted by prediction process: intra-frame prediction or inter-prediction. The result of prediction process is subtracted from the picture block. In next step, the residual is transformed mostly by Discrete Cosine Transform (DCT) and quantized. Finally, transformed output, prediction information, mode information and headers are encoded by entropy encoder. At decoder side every counter part of encoder blocks does the inverse operation to deliver the picture to the other end of the communication.

3. INTRA PREDICTION IN HEVC

In picture two distinct kinds of redundancy can be identified: 1) Spatial or temporal redundancy, 2) Psycho-visual redundancy. For remove spatial redundancy, prediction process is used. Intra-prediction is the process to predict pixels of picture frame. Intra-picture prediction uses neighbourhood pixels for predict picture block. Before intra-prediction frame must be splitted.

3.1 Picture Partition

As shown above, previous standards from H.120 to H.264 divided the pictures in block-shaped regions called Macroblocks. But HEVC splits the picture in to quadtree structure. HEVC support variable block size. Nowadays high-resolution video contents are used, so the use of larger blocks is beneficial for encoding. To support this variety of blocks size in efficient manner HEVC pictures are spited into coding tree Units (CTUs). Depending on the parameters like texture complexity, the CTUs can have the size: 64×64 , 32×32 , or 16×16 [3]. Coding Tree Unit (CTU) is therefore a coding logical unit, which is in turn encoded into an HEVC bit-stream. It consists of three blocks, namely luma (Y) two chroma components (Cb and Cr). Luma component have $L \times L$ samples and each chroma component have $L/2 \times L/2$ samples. Each block is called Coding Tree Block (CTB) [4]. Each CTB has the same size ($L \times L$) as the CTU (64×64 , 32×32 , or 16×16). Each CTB can be divide repetitively in a quad-tree structure [5], from the same size as CTB to as small as 8×8 . Each block resulting from this partitioning is called Coding Blocks (CBs) and becomes the decision making point of prediction type (inter or intra prediction) [4]. The prediction type along with other parameters is coded in Coding Unit (CU). So CU is the basic unit of prediction in HEVC, each of which is predicted from previously coded data. And the CU consists of three CBs (Y, Cb and Cr). CBs could still be too large to store motion vectors (inter-picture (temporal) prediction) or intra-picture (spatial) prediction mode. Therefore, Prediction Block (PB) was introduced. Each CB can be split into PBs differently depending on the temporal and/or spatial predictability. The CTUs can have the size: 32×32 , 16×16 , 8×8 or 4×4 [5].

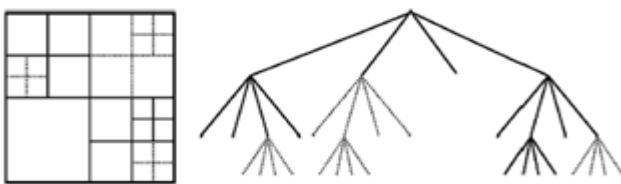


Fig 2: Quadtree structure [5]

3.2 Intra-Prediction Modes

The Intra-picture prediction uses the previously decoded boundary samples from spatially neighbouring block in order to predict a new prediction block PB. So the first frame of a video sequence and the first picture at each clean random access point into a video sequence are coded using only intra-picture prediction [6].

Several improvements have been introduced in HEVC in the intra prediction module: 1) Due to the larger size of the pictures, the range of supported coding block sizes has been increased. 2) A plane mode that guarantees continuity at block boundaries is desired. 3) The number of directional orientations has been increased. 4) For intra mode coding, efficient coding techniques to transmit the mode for each

block are needed due to the increased number of intra modes. 5) HEVC supports a large variety of block size, so it needs consistency across all block size.

HEVC employs 35 different intra modes to predict a PB: 1) 33 Angular modes, 2) Planar mode and 3) DC mode. The table shows the mode name with their corresponding intra prediction mode index.

Table 1. Intra-Prediction Mode

| Intra prediction mode name | Mode index |
|----------------------------|--------------|
| Planar mode | Mode 0 |
| DC mode | Mode 1 |
| Angular modes | Mode 2 to 34 |

3.3 Pb Partition

The CB can be split into size of $M \times M$ or $M/2 \times M/2$. The first one means that the CB is not split, so the PB has the same size as the CB. It is possible to use it in all CUs. The second partitioning means that the CB is split into four equally-sized PBs. This can only be used in the smallest 8×8 CUs. In this case, a flag is used to select which partitioning is used in the CU. Each resulting PB has its own intra prediction mode [6]. The prediction blocks size range from 4×4 to 64×64 .

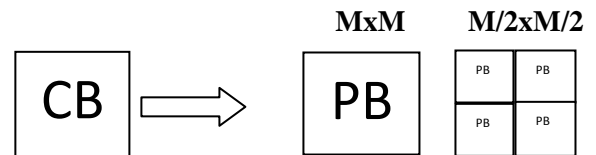


Fig 3: Prediction Block

3.4 Reference Samples

All the intra prediction modes use the same set of reference samples, which are extracted at the boundary from the upper (red samples) and left (blue samples) blocks adjacent to the current PU. For the diagonal directions the top-left corner (pink) sample may also be used. Also, it is possible to use the lower left (orange samples) and above right (green samples), if they are available from preceding decoding operations. Therefore, to predict the current PU of size $M \times M$, a total of $4M+1$ spatially neighbouring reference samples may be used. In figure, the reference samples are named as $R_{x,y}$ with the origins in the topleft corner. Also, the predicted sample value in the current PU in the position (x,y) is denoted as $P_{x,y}$. This notation is shown as follows in a block of size $M \times M$ samples. Here, in figure 4 the reference samples are named as $R_{x,y}$ with the origins in the topleft corner. Also, the predicted sample value in the current PU in the position (x,y) is denoted as $P_{x,y}$. This notation is shown as follows in a block of size $M \times M$ samples.

Reference samples are constructed by checking the availability of the neighbouring samples has to be tested. There are some cases reference samples not available. When there are some reference samples unavailable not are filled with the nearest available reference sample value. If none of the reference sample locations is available, the set of prediction samples is filled with the mid-value of the specified dynamic range of the encoded video (i.e. the value $2Bd-1$, e.g. 128 for 8-bit video).

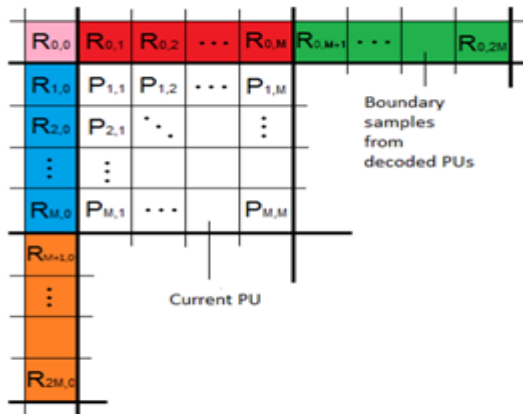


Fig 4: Reference sample

3.5 Angular Mode

The angular prediction process is operating in the spatial domain. Comparing with previous standard H.264, this method has though improved significantly because of increased number of directional orientations and the bigger block sizes, providing a good compromise between encoding complexity and coding efficiency for typical video material[8]. The angular mode in HEVC supports 33 different directional orientations. In figure, Angular modes are indexed from 2 to 34. Directions may cover angles from near-horizontal through near-diagonal to near-vertical. Each mode has associated a displacement parameter d , where the value of d indicates the numeric part which is the pixel's displacement expressed in $1/32$ pixel fractions of accuracy, and H and V indicate the horizontal and vertical directionalities [8]. The modes 10 are 26 are known as pure horizontal prediction and pure vertical prediction.

In order to calculate the value of each sample of the PB, the angular mode[8] extrapolates the samples from the reference samples, depending on the directional orientation in order to achieve lower complexity. When the direction selected is between 2 and 17, the samples located in the above row are projected as additional samples located in the left column, extending the left reference column. When the direction selected is between 18 and 34, the samples located at the left column are projected as samples located in the above row, extending the top reference row. In both cases the samples projected would have negative indexes. Once the reference samples are known, each sample are predicted by projecting its location into the reference samples and applying the selected prediction direction.

Sample prediction equation for vertical modes i.e., angular modes 18-34:

$$P_{x,y} = ((32 - wy) \cdot R_{i,0} + wy \cdot R_{i+1,0} + 16) \dots \dots \dots (1)$$

$$cy = (y \cdot d) \gg 5$$

$$wy = (y \cdot d) \& 31$$

$$i = x + cy$$

Sample prediction equation for horizontal modes i.e angular modes 2-17:

$$P_{x,y} = ((32 - wx) \cdot R_{i,0} + wx \cdot R_{i+1,0} + 16) \dots \dots \dots (2)$$

$$cx = (x \cdot d) \gg 5$$

$$wx = (x \cdot d) \& 31$$

$$i = y + cx$$

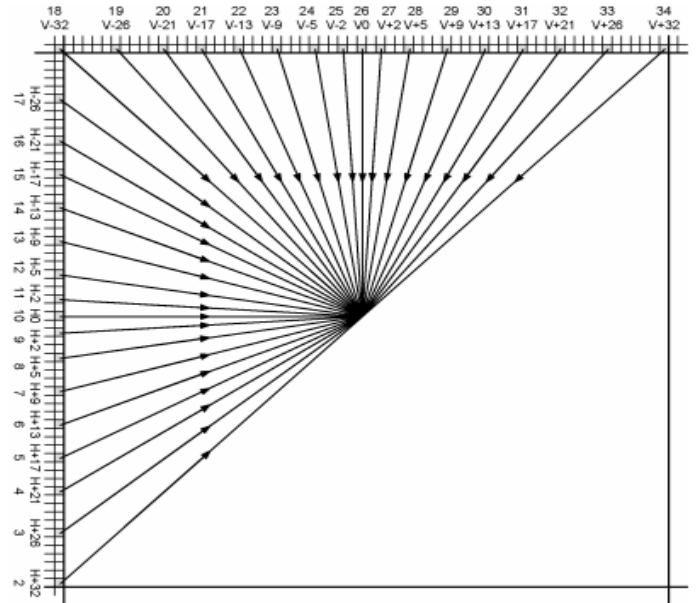


Fig 5: Angular modes [5]

In this equations, $R_{x,y}$ represents the reference samples and $P_{x,y}$ represents the sample to be predicted, where x, y are the spatial coordinates. The sub pixel location in between $R_{i,0}$ and $R_{i+1,0}$, cx and cy represents pixel parameters corresponding to x and y coordinates, wy represents the weight/weighting parameter of weighed prediction, i stands for reference sample index, ' \gg ' denotes a bit shift operation to the right and $\&$ denotes logical AND operation. The parameters cy and wy depend only on y coordinate in vertical mode and cx and wx depend only on x coordinate for horizontal modes [5].

3.6 Planner Mode

This mode in HEVC is similar to the planar mode in H.264/MPEG-4 AVC, and is known as mode 0. In H.264/MPEG-4 AVC this method is a plane prediction mode for textured images, and may introduce discontinuities along the block boundaries. Conversely, in HEVC this mode was improved in order to preserve continuities along the block edges. The value of each sample of the PB is calculated assuming an amplitude surface with a horizontal and vertical smooth gradient derived from the boundaries samples of the neighbouring blocks [9].

Sample prediction equation for planner mode

$$P_{x,y} = PV_{x,y} + PH_{x,y} + N \gg \log_2(N) + 1 \dots \dots \dots (3)$$

$$PV_{x,y} = (N - y) \cdot R_{x,0} + y \cdot R_{N,0}$$

$$PH_{x,y} = (N - x) \cdot R_{0,y} + x \cdot R_{N,y}$$

In this equation, x and y denoted location of sample. N is stand for total number of sample. H is for vertical and V is for horizontal. R stands for reference samples.

3.7 Dc Mode

This mode is also similar to the DC mode in H.264/MPEG-4 AVC and is known as mode 1. It is efficient to predict plane areas of smoothly-varying content in the image, but gives a coarse prediction on the content of higher frequency components and as such it is not efficient for finely textured areas. The value of each sample of the PB is an average of the reference samples. As explained before, for this case the reference samples will be the boundary samples of the top and

left neighbouring blocks. So, all the samples of the PB are predicted with the same value [9].

3.8 Reference Samples Smoothing

HEVC uses a smoothing filter[10] in order to reduce the discontinuities introduced by the intra-prediction modes. This is applied to the boundary samples, namely the first prediction row and column for DC mode, or the first prediction row for pure horizontal prediction, or the first prediction column for pure vertical prediction. The smoothing filter consists of a two-tap finite impulse response filter for DC prediction or a gradient based smoothing filter for horizontal (mode 10) and vertical (mode 26) prediction. Due to the fact that chroma components tend to be already smooth, this filter is not used in this case. Prediction boundary smoothing is only applied to luma component. The smoothing filter is applied to the reference samples depending on the size of the blocks and the directionalities of the prediction. For the block sizes which is larger than 4, lowpass filtering[10] is applied. The diagonal directions (modes 2, 18, and 34) are always filtered. The number of directions where filtering is applied increases with the block size. Reference samples for planar prediction of blocks with larger than 4 are always filtered. For the reference samples of 32×32 blocks, the conditions for regular or strong filtering. If regular lowpass filtering is indicated a 3-tap lowpass filter is applied.

4. COMPARISON OF INTRA-PREDICTION IN HEVC AND H.264

H.264/AVC have eight directional intra prediction modes and one DC intra prediction mode. This nine modes are specified for 4×4 and 8×8 blocks, but in case of 16×16 intra predicted blocks, only horizontal, vertical, DC, and plane prediction are available. The directions included horizontal, vertical, left and right diagonal, and half diagonal directions [11]. For 8×8 blocks, pre-filtering is applied. This process is same explained in previous section for HEVC. In HEVC, the intra prediction design has been generalized. All 35 modes are applied for all size of prediction block. The number of possible angular directions for intra prediction has increased to 33. The directional prediction is further applicable to all possible intra prediction block sizes of 4×4 to 64×64 . The construction of a uniform prediction reference helps for angular implementation of the intra prediction process. The application of the planar intra prediction mode is also available for all prediction block sizes.

Table 2. Comparison Of Intra-Prediction In H.264 And Hevc

| Functionality | H.264/AVC | HEVC |
|--------------------------------|--|---|
| Prediction block sizes | 4×4 , 8×8 and 16×16 | 4×4 , 8×8 , 16×16 and 32×32 |
| intra prediction modes | 9 for 4×4 , 8×8 and 4 for 16×16 | 35 |
| Boundary smoothing | N/A | Used for directly horizontal, vertical and DC modes |
| Reference sample smoothing | For 8×8 | For 8×8 and above |
| when reference samples missing | Use DC mode | Reference sample substitution |

5. CONCLUSION

In this paper different papers have been surveyed and presented as a review of the recently proposed for the High Efficiency Video Coding (HEVC) standard. It follows the same basic design principles as already established by the state-of-the-art video coding standard profile, with generalizations for more flexibility in terms of prediction. The new features of the proposals for HEVC intra-prediction design included the partition of prediction blocks which use quad-tree structure for splitting, the interpolation scheme, directional interpolation filters, angular prediction, planar coding. Many compression improvements were made by increasing adaptability. From this survey confirm that the goal of developing a high-efficiency video coding standard, which delivers significantly improved compression performance relative to that of the H.264/AVC.

6. REFERENCES

- [1] Jens-Rainer Ohm, Gary J. Sullivan, Heiko Schwarz, Thiow Keng Tan, Thomas Wiegand, 2012. Comparison of the Coding Efficiency of Video Coding Standards—Including High Efficiency Video Coding (HEVC). IEEE Transaction on circuits & systems for video technology, Vol.22, No. 12.
- [2] Iain Richardson. HEVC: An introduction to High Efficiency Video Coding.
- [3] <http://www.vcodex.com/images/uploaded/342512928230717.pdf>
- [4] Gary J. Sullivan, Jens-Rainer Ohm, Woo-Jin Han, Thomas Wiegand, December 2012. Overview of the High Efficiency Video Coding(HEVC) Standard. IEEE Transaction on circuits & systems for video technology, Vol.22, No. 12,
- [5] Philipp Helle, Simon Oudin, Benjamin Bross, Detlev Marpe, M. Oguz Bici, Kemal Ugur, Joel Jung, Gordon Clare, and Thomas Wiegand, December 2012. Block Merging for Quadtree-Based Partitioning in HEVC. IEEE Transactions on Circuits and System for VideoTechnology.
- [6] Jani Lainema, Frank Bossen, Woo-Jin Han, Junghye Min, Kemal Ugur, December 2012. Intra Coding of the HEVC Standard. IEEE Transaction on circuits & systems for video technology, Vol.22, No. 12.
- [7] Frank Bossen, Benjamin Bross, Karsten Suhring, David Flynn, December 2012. HEVC Complexity and Implementation Analysis. IEEE Transaction on circuits & systems for video technology, Vol.22, No.12.
- [8] Jani Lainema, Kemal Ugur, 2011. Angular Intra Prediction in High Efficiency Video Coding (HEVC). Multimedia Signal Processing (MMSp).
- [9] https://www4.rgu.ac.uk/files/h264_intrapred.pdf
- [10] Mathies Wain, 2015. High Efficient Video Coding(HEVC) – coding tool and specifications. Springer.
- [11] Advanced video coding for generic audiovisual services. ITU-T Rec. H.264 (AVC). <http://www.itu.int/rec/T-REC-H.264/en> (2014) Accessed 14 Apr 2014.