International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-8 Issue-11, September 2019

H.265 Intra-Picture Prediction Acceleration using Low Complexity Model

Chhaya Shishir Pawar, Sudhir Deoraoji Sawarkar

Abstract*: *H.265 also called High Efficiency Video Coding is the new futuristic international standard proposed by Joint collaboration Team on Video Coding and released in 2013 in the view of constantly increasing demand of video applications. This new standard reduces the bitrate to half as compared to its predecessor H.264 at the expense of huge amount of computational burden on the encoder. In the proposed work we focus on intraprediction phase of video encoding where 33 new angular modes are introduced in addition to DC and Planar mode in order to achieve high quality videos at higher resolutions. We have proposed the use of applied machine learning to HEVC intra prediction to accelerate angular mode decision process. The features used are also low complexity features with minimal computation so as to avoid any additional burden on the encoder. The Decision tree model built is simple yet efficient which is the requirement of the complexity reduction scenario. The proposed method achieves substantial average encoding time saving of 86.59%, with QP values 4,22,27,32 respectively with minimal loss of 0.033 of PSNR and 0.0023 loss in SSIM which makes it suitable for acceptance of High Efficiency Video coding in real time applications

***Index Terms*: Angular mode decision, HEVC Intraprediction, Machine learning, Video Coding.**

I. INTRODUCTION

Information in Visual form is more interesting to human as compared to any other form like oral or textual. It can be better understood, interpreted and imagined. Therefore, we find the use of video in almost every domain of our daily lives. Hence JCT-VC has introduced a new video coding standard H.265 aka HEVC which is expected to reduce the existing bitrate to half without compromising on quality of the video and being network friendly at the same time.

H.265 introduces many techniques to achieve the required bitrate like Coding unit's flexible partitioning structure, 33 angular modes for intraprediction, improved deblocking filters, parallel processing techniques etc.

Intraprediction is the process where picture content is predicted without the reference to earlier decoded picture. In order to achieve a high-quality picture HEVC introduced 33 directional modes in addition to Planar and DC modes. Due to this every prediction unit from the size of 32x32 till 4x4 of an intrapicture from various depth levels of the quadtree has to check for all 35 modes for the selection of the best suitable mode with minimum error. This put a lot of extra computational burden on the HEVC encoder.

In the proposed work we demonstrate how we can speed up the mode decision process with the help of machine learning technique. Machine learning techniques are capable of learning the function from huge amount data and use this knowledge for future predictions. In the domain of video coding these can learn from various features of the sample videos in offline training and the trained model can use the acquired knowledge about the mode prediction during the real time intraprediction phase of video compression.

II. HEVC INTRAPREDICTION

High Efficiency Video Coding has two categories for intra picture prediction. In the first category which is meant for directional edges in the picture in order to bring more clearer picture. And the second one with DC and planar prediction of smooth image area with homogeneous content. The intraprediction is carried in mainly three phases as Top and Left reference array generation, prediction of the samples and post processing of the samples.

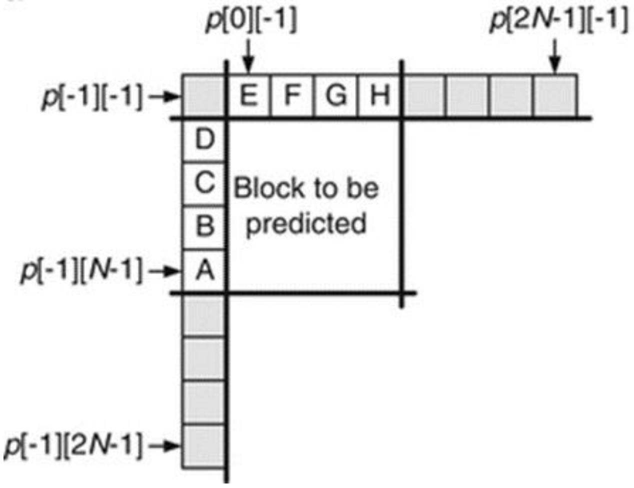

Fig 1: Reference Array for HEVC Intraprediction

As seen in Fig.1 the block to be predicted uses the reference pixels in the top and left-hand side. Unlike H.264 , HEVC allows to extrapolate and use the Reference array substitution for prediction of the current block using complete set of intraprediction modes. As opposed to H.264 the unavailable reference samples are not replaced by DC mode whereas in HEVC the reference array is substituted by the available reference samples in the clockwise direction.

Set of 33 angular modes have been designed for HEVC so as to efficiently present the directional structures in the image and video content. In natural scenes horizontal and vertical patterns are more common.

Manuscript published on 30 September 2019.
***Correspondence Author(s)**
Chhaya Shishir Pawar, Department of Computer Engineering, Datta Meghe College of Engineering, Maharashtra, India.
Dr.Sudhir Deoraoji Sawarkar, Department Of Computer Engineering, Datta Meghe College of Engineering, Maharashtra, India.
This work is supported by Minor research grant from University of Mumbai.

Hence we can observe from Fig.2 that the displacement from the vertical and horizontal direction is less as compared to modes away from these two.

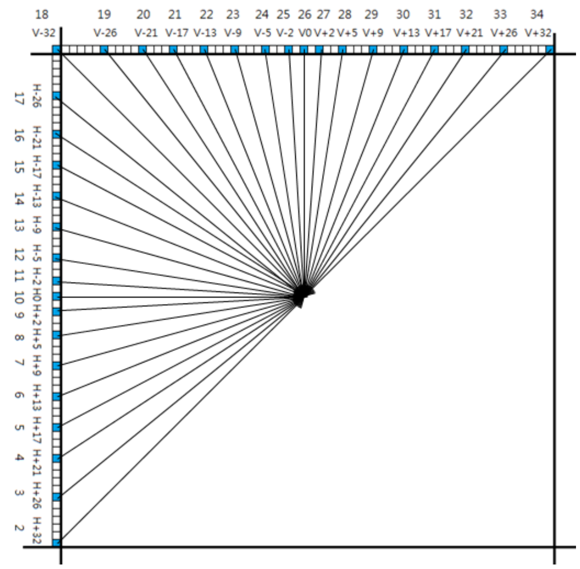

Fig 2: Angular mode Definitions in HEVC

III. LITERATURE SURVEY

Heming Sun et al (2012) [1] proposed a scheme which they proposed the calculation of SATD values based on original pixels instead of reconstructed pixels. A cost vector is defined for each prediction unit with 0 to 34 SATD values. The minimum 9 values are selected from this vector. Thereby we reduce the number of modes for selection are reduced from 35 to 9. They replaced the HAD based cost calculation to SATD based cost vector.

Hao Liu et al (2016) [2] makes use of Hybrid cost ranking that is it uses Hadamard cast along with the RD cost to decide the most promising mode. RD candidate set is obtained from the increasing order of Hadamard candidate set. It also makes use of the correlation between the rough mode and the most promising mode decision.

Jakub Siast et al (2016) [3] demonstrated a hierarchical technique where the subsets of every second, third and fourth modes are created and the most probable modes are checked against those subsets. The method checks for near optimal RD costs and reduces the time for Rough mode decision.

Anis BenHajyoussef et al (2017) [4] calculates the gradient direction and gradient magnitude for the picture and for all 33 angular modes. It supports the rough mode decision process with the gradient values to calculate the cost. The higher the cost more optimal the mode is. They reduce the number of modes by probabilistically considering the different cases where how much the RMD candidate set matches with the mode candidates suggested with the gradient based approach.

Yu Zhang et al (2014) [5] proposes the partial differential equation-based solution to intraprediction speed up problem. The author defines the function over reconstructed block and predicted block. They then apply the RDO based checking for finding the best intra mode between the Partial differential equation based inpainting mode and its two neighboring modes and achieves the reduction in bitrate.

Pavan Gajjala et al (2013) [6] presents a system where the angular intraprediction makes use of the directly adjacent pixels as reference instead of the traditional reference pixel calculation for HEVC that is using the interpolated reference pixels array. It performs prediction by pixel by pixel approach and achieves the reduction in encoding time.

Maxime Bichon et al (2018) [7] proposed a method where author calculates the dominant edge in the current prediction unit and creates the subset of modes for intraprediction with the help of preliminary experiments about the texture directions of prediction units. If the information about the texture related information is same as the previous Prediction Unit then the same candidate subset can be reused.

Maxim P. Sharabayko et al (2013) [8] presents iterative method for intra mode decision. It makes use of at the most 15 angular modes instead of all 35 modes. They have found out the few fastest modes along with their +/-2 neighboring modes and equidistant modes by experimentation. Thereby reducing the encoding time involved in Intraprediction.

Gaoxing Chen et al (2013) [9] calculates the gradient statistics of the prediction units. And based on the gradient computations the mode of the prediction unit is decided with help of the threshold values except the DC and planar mode. This method is assisting the mode refinement after the rough mode decision.

Mengmeng Zhang et al (2017) [10] The author has used nine equidistant modes along with Planar and DC instead of 35 modes. For first phase of rough mode decision it uses the SATD in the RMD evaluation and SSE in the second phase.

Sookyung Ryu et al (2018) [11] presents a random forest-based technique to speed up the mode decision process. It uses pixel-based approach. It uses four random pixels from four quadrants of the prediction unit as feature to the algorithm.

Andreas Heinde et al (2016) [12] suggests the use of calculating mean of reference pixels of the current prediction units along with the existing fast intra mode decision algorithms.They presented global and local exclusion of angular modes to form the subsets. Global method analyses the variance of the reference samples to decide which modes to include in the candidate mode set.

Haijun Lei et al (2013)[13] calculates sum of absolute differences by down sampling method and search for the best intra mode. In the three steps procedure first they find out the candidate set and its 1-distance and 2-distance neighbors. Whether all the candidate sets undergo RDO process is dependent on the weighted cost.

IV. PRELIMINARY EXPERIMENTS

H.264 specified 9 intra prediction modes with Planar and DC mode. In order to incorporate the finest details of the high-resolution picture along the different directions HEVC came up with 35 intra-picture prediction modes along with Planar and DC mode. These are intended for the homogeneous areas of the picture. In natural imagery major area of the picture is homogeneous in nature. So Planar and DC modes must be used widely as compared to other angular modes.

Retrieval Number: K22360981119/19©BEIESP
DOI: 10.35940/ijitee.K2236.0981119
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

Experiment was carried out to understand the selection of various modes in the different high-resolution images with level of details.

The prediction units (4x4) of frames of *Lena, Park Joy, controlled_burn, flowers* and *pillar* were used to understand the frequency of intra modes.

It can be observed from statistics of Fig 3 that the Planar and DC modes were dominant among all the 35 Intra-picture prediction modes.

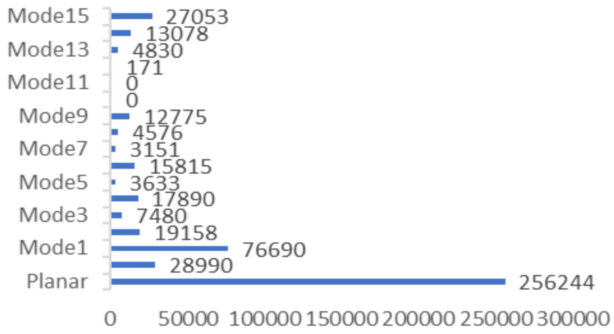

Fig 3: Frequency of occurrence of IntraPrediction modes

But still every Prediction unit of all sizes till 4x4 has to undergo all the 33 modes intra mode calculations and majority of the computation is wasted as the statistics shows that many of the intra modes remain underutilized.

V. DECISION TREES

Machine learning is the hot topic of today's era. It is capable of giving solutions to almost every topic where large amount of data is available for analysis. The huge amount from data available from the picture in terms of prediction units and dominant use of particular prediction mode motivated us to make use of machine learning technique. Classification techniques of machine learning are helpful in intelligent decision making. Decision trees are the optimum choice as We would not like to overburden the encoder further. Decision trees can be implemented in the form of simple if-else statements which are computationally inexpensive. These are used to predict the intra mode of current prediction unit so that it doesn't have to check for all the 35 directional modes. Decision tree induction is constructing decision trees from the class labeled training dataset. Decision tree consist of tree like structure which originates at root node and each internal node represent the attributes used for classification. The leaf nodes at the end of the branch resent the final outcome in the form of predicted class label. Attribute selection measure is the criteria for selection of splitting attribute at the nodes. It gives ranking to the attributes as per the highest value of attribute selection measure. One of the measures is Gini Index. Gini index is used for measuring the impurity of D which is a training dataset.

$$Gini(D) = 1 - \sum_{i=1}^{m} p_i^2,$$

The probability of the tuple to belong to classs Ci is given by piand it is calculated as |Ci|/|D|.The sum is calculated over number of classes . The reduction in impurity due to split on attribute A is

$$\Delta Gini(A) = Gini(D) - Gini_A(D).$$

The attribute with maximum reduction in impurity or minimum Gini index values is taken as splitting attribute.

VI. PROPOSED METHODOLOGY

Use of 33 angular modes along with Planar and DC mode of intraprediction maintains the quality of the picture by taking care of directional structure but also is the reason behind increased computational complexity. This can be overcome if instead of checking all the 35 modes and finding the most suitable mode for intra prediction if we can predict the appropriate mode with minimal computational cost.

The proposed methodology presents a low complexity model for speeding up Intra-picture prediction in High efficiency Video coding. The model consists of decision tree which is trained offline. The trained model predicts the most suitable mode during run time. The method gives substantial encoding time saving during run time with minimal loss of Peak Signal to Noise ratio and Structural Similarity Index.

The model proposed here is a low complexity model because introduction of machine learning technique for prediction of intra picture mode will introduce additional computation. That must outperform the time required for computationally cumbersome exhaustive process for 35 modes. The factors behind the low complexity model are

- Use of simpler machine learning model for classification like decision tree which is easy to build, understand and the model is fast at prediction during run time because these are essentially if-else statements requiring nominal computation time.
- Features used are mean and variance of current prediction unit, Left and Top neighboring Coding Units. These are inexpensive in terms of computational efforts.

Feature engineering was carried out to select the best features in terms of time complexity and accuracy of mode prediction. Various features along with the above-mentioned features correlation, entropy, image gradient magnitude and direction, variance and mean of left and top reference pixels, difference between diagonal pixels of prediction unit were under consideration but features like correlation, entropy, gradient values were accurate but involved significant computational complexity. Others features like mean and variance of reference pixels, difference between diagonal pixels of prediction units were simpler but did not gain much in terms of accuracy of the classification algorithm. Final Intra prediction acceleration algorithm is given as,

Algorithm for Intra-frame Prediction Acceleration

Input: raw I frame on which spatial prediction is to be carried out.

Output: Intra-predicted compressed and decompressed frame

Step1: Prepare the training dataset with values of above-mentioned features of prediction units

Step2: Train the Decision tree learning model offline.

Step3: Export the learned model

Step4: Run the Intra picture prediction for test dataset by using mode predictions from decision tree classifier

Step5: Perform transform, Quantization and entropy encoding

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

The intra prediction speed up process is presented in Figure 4. In the training dataset, the feature information of every prediction unit is gathered. The classifier s trained with satisfactory accuracy. The input I frame is divided into Coding units and further it id decomposed into prediction units. For every prediction unit the leaned classifier will predict the angular mode.

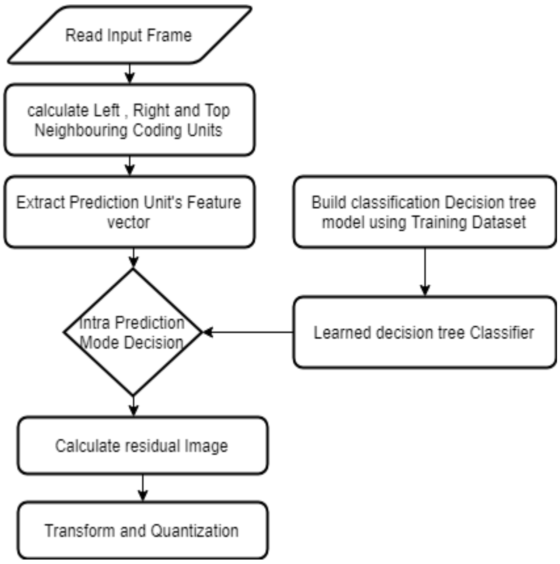

Fig 4: Flowchart of the Proposed Method

The compressed frame will undergo inverse transform and inverse quantization at the decoder end. Hence the decision tree classifier predicted the intra mode based on the values of the feature vector.

VII. RESULTS AND DISCUSSION

Experimental Set up includes the training dataset, Classifier accuracy, test dataset used and size of the prediction units. *Lena, Rainier, Ducks_take_off, life_1080p* of 1920x1080 resolution were the raw frames used for training the classifier. The training Dataset consists of 4,91,520 records. Each for 4x4 Size of the prediction unit. The testing was carried out for different values of quantization parameter. i.e. QP=4,22,27. Entire experiment was carried out in the Matlab environment.

The metrics used for evaluation includes Encoding Time, Peak Signal to Noise ratio and Structural Similarity Index. All the parameters were verified against the standard HEVC intraprediction which checks for all the 33 angular modes along with Planar and DC mode.

% Encoding Time Saving=100-($ET_{proposed/}$ $ET_{standard}$*100)
Where $ET_{proposed\ is}$ the Encoding time required by proposed method and $ET_{standard}$ encoding time required by the HEVC standard method. Difference in PSNR is given by

$$\Delta\ PSNR=PSNR_{standard}-PSNR_{proposed}$$

where $PSNR_{standard}$ and $PSNR_{proposed}$ are the PSNR values for standard method and proposed methods respectively.

$$\Delta\ SSIM=SSIM_{standard}-SSIM_{proposed}$$

where $SSIM_{standard}$ and $SSIM_{proposed\ are}$ the SSIM values for HEVC standard method and proposed methods respectively.

Features like Mean and Variance of Left and Top Coding Unit, Mean and Variance of Current prediction unit were collected for every 4x4 PU from the Training dataset frame and Decision tree Classifier was built with Gini index as attribute selection method.

Fig 5 shows the Encoding time saving for the test dataset Which consists of test data of various resolutions and various levels of details, different amount of homogeneous regions. It is evident from Fig 5 that we save substantial amount of encoding time as we save 35 complex angular mode calculations per Prediction unit of size 4x4 in high resolution images. Such way Prediction Units of size 32x32,16x16,8x8 can save enormous complexity arising out of Intra angular mode calculation during quad tree composition.

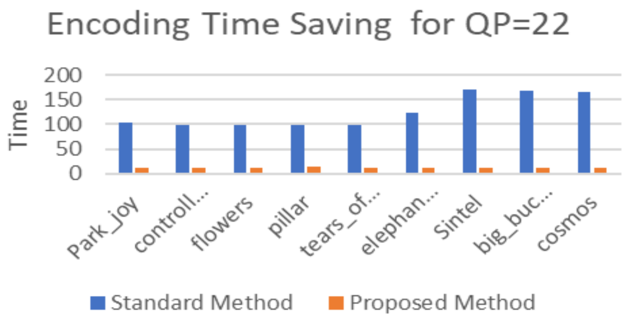

Fig 5: Encoding Time saving for QP=22

Fig 5: 'Sintel' movie frame compressed by Standard method

Fig 6:'Sintel' movie frame predicted by Proposed method

Fig 5 and 6 shows the frame encoded by HEVC standard method and the proposed method. Subjective evaluation does not suggest much difference between the two encoded frames.

Retrieval Number: K22360981119/19©BEIESP
DOI: 10.35940/ijitee.K2236.0981119
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

Table 1: Performance Analysis of Proposed Method for QP=4,22,27

	QP=4			QP=22			QP=27		
	Encoding Time Saving %	Δ PSNR	Δ SSIM	Encoding Time Saving %	Δ PSNR	Δ SSIM	Encoding Time Saving %	Δ PSNR	Δ SSIM
Park_joy	86.923	0.03	0.0034	88.207	0.02	0.0004	88.064	0.02	0.0074
controlled_burn	87.596	0	0.0005	87.485	0.01	0	88.864	0.01	0.0002
flowers	87.347	0.02	0.0103	87.190	0.02	0.0024	83.631	0.02	0.0003
pillar	88.120	0.04	0.0017	86.835	0.03	0.0053	84.060	0.03	0.022
tears_of_steel	90.596	0.07	0.0015	87.487	0.06	0	76.448	0.06	0.0003
elephant_dream	87.450	0.07	0.0033	90.122	0.06	0.0019	85.386	0.05	0.0279
Sintel	87.679	0.08	0.003	92.774	0.07	0.001	84.068	0.06	0.0027
big_buck_bunny	80.177	0.02	0.0003	92.445	0.01	0.0002	84.759	0.01	0.0002
cosmos	87.853	0.09	0.0026	92.614	0.06	0.0026	84.957	0.08	0
Average	**87.959**	**0.04429**	**0.003386**	**87.429**	**0.02**	**0.00203**	**84.409**	**0.03571**	**0.0022**

Table 1 shows the results for Savings in encoding time for different values of QP ie.4,22,27. The proposed method achieves average 86.59% time saving with 0.033 loss of PSNR and 0.0023 loss of structural similarity index. One of the contributing factor for reduced computational complexity is instead of SATD(Sum of transform differences) we have used SAE(Sum of Absolute Error) as the error measure as it is computationally lighter than SATD.

VIII. CONCLUSION

We presented a low complexity machine learning based model for speeding up HEVC intraprediction. The proposed method replaces the exhaustive method by directly predicting the mode with help of decision tree. Decision tree predicts the mode based on the features of the prediction Unit. The method achieves significant amount of encoding time saving against the HEVC standard method which checks for all the 35 intraprediction modes. On average it achieves 87% time saving at the cost of minimal loss in PSNR and SSIM values.

The low complexity is achieved by the use of offline trained decision tree which involves minimal time requirement during run time. Also, SAE used as the error measure further reducing the computational time.

This method can be used as the solution for reducing the computational complexity of high efficiency video coding and along with other solutions for inter picture prediction will enhance its widespread acceptability in the applications. This significant gain can also reduce the power requirements of the encoding devices which includes mainly mobile devices.

REFERENCES

1. Heming Sun, Dajiang Zhou and Satoshi Goto, "A Low-Complexity Hevc Intra Prediction Algorithm Based On Level And Mode Filtering", 2012 IEEE International Conference on Multimedia and Expo, PP-1085-1090
2. HaoLiuandYuexinJie, "Fast HEVC Intramode Decision Based on Hybrid Cost Ranking", Hindawi Publishing Corporation Advances in Multimedia Volume 2016, Article ID 5164025, 9 pages http://dx.doi.org/10.1155/2016/5164025
3. Jakub Siast, Jakub Stankowski, Marek Domański," Hierarchical Fast Selection of Intraframe Prediction Mode in HEVC", Intl Journal of Elctronics And Telecommunications, 2016, Vol. 62, No. 2, Pp. 147-151
4. Anis BenHajyoussef1,2*, Tahar Ezzedine1 and Ammar BouallègueE," EURASIP Journal on Image and Video Processing (2017) DOI 10.1186/s13640-016-0159-9
5. Yu Zhang and Yinyi Lin, "Improving HEVC Intra Prediction with PDE-Based Inpainting" APSIPA 2014.
6. Pavan Gajjala, K. R. Rao," Efficient HEVC Lossless Coding Using Sample Based Angular Intra Prediction", Majlesi Journal of Multimedia Processing, Vol. 2, No. 3, September 2013
7. Maxime Bichon, Julien Le Tanou, Michael Ropert, Wassim Hamidouche, Luce Morin, Lu Zhan," Low Complexity Joint RDO of Prediction Units Couples for HEVC Intra Coding", International Conference on Acoustics, Speech and Signal Processing (ICASSP), Apr 2018, Calgary, Canada.
8. Maxim P. Sharabayko, Nikolay G. Markov," Iterative Intra Prediction Search for H.265/HEVC", 2013 International Siberian Conference on Control and Communications (SIBCON)
9. Gaoxing Chen∗, Lei Sun∗, Zhenyu Liu†, and Takeshi Ikenaga," Fast Mode and Depth Decision HEVC Intra Prediction Based on Edge Detection and Partitioning Reconfiguration" 2013 International Symposium on intelligent signal processing and communication systems.
10. Mengmeng Zhang*, Xiaojun Zhai and Zhi Liu," Fast and adaptive mode decision and CU partition early termination algorithm for intra-prediction in HEVC", EURASIP Journal on Image and Video Processing 2017, DOI 10.1186/s13640-017-0237-7
11. Sookyung Ryu and Je-Won Kang," Machine Learning-Based Fast Angular Prediction Mode Decision Technique in Video Coding", DOI 10.1109/TIP.2018.2857404, IEEE Transactions on Image Processing
12. Andreas Heindel, Christoph Pylinski, and Andr´e Kaup," Two-Stage exclusion of angular intra prediction modes for fast mode Decision in hevc",ICIP 2016
13. Haijun Lei, Zhongwang Yang, "Fast Intra Prediction Mode Decision for High Efficiency Video Coding," 2nd International Symposium on Computer, Communication, Control and Automation (3CA 2013). Atlantis Press,pp. 34-37
14. Gary J.Sullivan,Jens-Rainer Ohm, Woo-Jin Han, "Overview of High Efficiency Video Coding (HEVC) Standard", IEEE Transactions on Circuits and Systems for Video Technology,Vol:22, Issue;12,Dec2012
15. Frank Bossen, Benjamin Bross,Karsten Suhring, David Flynn," HEVC Complexity and Implementation Analysis", IEEE Transactions on Circuits and Systems for VideoTechnolgy,Vol.22,Issue12,2012
16. Vivienne Sze, Madhukar Budagavi, Gary Sullivan," High Efficiency Video Coding (HEVC)Algorithms and Architectures", Springer Publication
17. Iain E.G.Richardson," H.264 and MPEG-4 Video Compression Video Coding for nest generation Multimedia", Wiley Publication

Retrieval Number: K22360981119/19©BEIESP
DOI: 10.35940/ijitee.K2236.0981119
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

AUTHORS PROFILE

Mrs.Chhaya S Pawar received B.E and M.E degree from University of Mumbai in 2006 and 2013 respectively. She has published 28 papers in various national and international journals and conferences and is member of Computer Society of India and International Association of Engineers. She is currently pursuing Ph.D degree from University of Mumbai. Her research interest includes Artificial Intelligence, Video Coding.

Dr.Sudhir Deoraoji Sawarkar received B.E and M.E in Electronics Engineering and Ph.D degree in Computer Engineering. He has more than 29 years of Experience. He has published more than 182 research papers in national and international journals and conferences and is member of Computer Society of India and ISTE. His Research interests include Artificial Intelligence, Neural Networks, Image Processing.

Retrieval Number: K22360981119/19©BEIESP
DOI: 10.35940/ijitee.K2236.0981119
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)