Fast Intra Prediction Mode Decision for HEVC Using Random Forest

Zhuge Yan
Electrical and Electronic Engineering Dpt.
The University of Nottingham Ningbo China
zhuge.yan@nottingham.edu.cn

Siu-Yeung CHO
Electrical and Electronic Engineering Dpt.
The University of Nottingham Ningbo China
+86 574 8818 0000-3631
david.cho@nottingham.edu.cn

Sherif Welsen
Electrical and Electronic Engineering Dpt.
The University of Nottingham Ningbo China
+86 (0) 574-8818 0000 (ext 9383)
Sherif.welsen@nottingham.edu.cn

ABSTRACT
In this paper, we extracted specific image features that represent CU texture, incorporate a machine learning technique, namely random forest, in HEVC intra prediction mode selection, to improve the performance of intra coding of HEVC. Compared with similar algorithms, our method extracts very specific features of image texture changes in terms of angle. Therefore the proposed method can achieve very high prediction accuracy. Having similar reduction in complexity, the proposed algorithm can gain higher video quality compared with similar algorithms.

CCS Concepts
• **Computing methodologies → Image compression**

Keywords
Intra Prediction; HEVC; Random Forest.

1. INTRODUCTION
Machine Learning (ML) is a multi-disciplinary field that involves multiple disciplines such as probability theory, statistics, approximation theory, convex analysis, and algorithm complexity theory [1], [2]. Specializing in how computers simulate or imitate human learning behavior to acquire new knowledge or skills, and reorganize existing knowledge structures to continuously improve their performance. It is the core of artificial intelligence and the fundamental way to make computers intelligent. Its application covers all areas of artificial intelligence. It mainly uses induction, synthesis rather than deduction. Machine learning has a wide range of applications, such as: data mining, computer vision, natural language processing, biometrics, search engines, medical diagnostics, detection of credit card fraud, stock market analysis, sequencing of DNA sequences, speech and handwriting recognition, and strategies Games and robots are used [1], [2].

IVSP 2019, February 25–28, 2019, Shanghai, China

ACM ISBN 978-1-4503-6175-0/19/02…$15.00
DOI: https://doi.org/10.1145/3317640.3317645

Learning skills of machine learning technique are through big data (large amount of training sample). As a classifier, its universality is better than a human designed one. In this paper random forest [3] is used to accurate the process of HEVC intra prediction mode decision.

HEVC intra prediction includes 35 modes, 33 of which are angle dependent (Figure 1). The intra prediction coding process of HEVC firstly performs rough mode selection (RMD), performs cost function calculation on all 35 intra prediction modes, selects N candidate mode entry rate distortion optimization (RDO) processes, and finally selects the optimal frame.

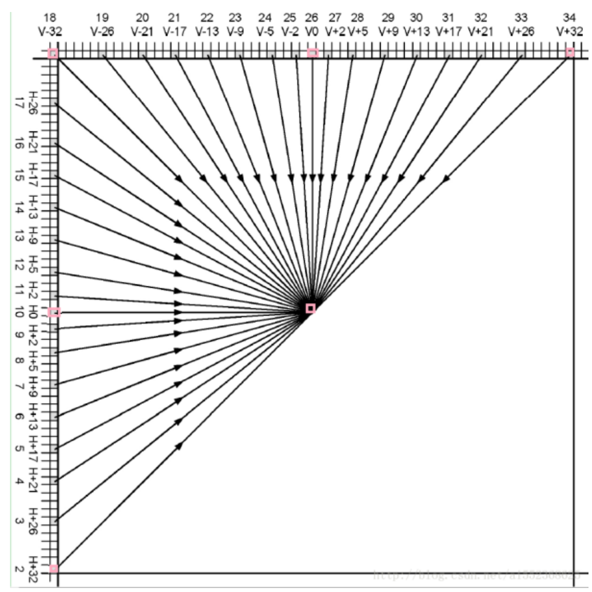

Figure 1. HEVC intra prediction angular modes

In the intra prediction process, the amount of calculation of the RMD cost function is very large. If the number of intra prediction modes that need to be calculated by the cost function can be reduced by pre-judging, the calculation amount of the intra-frame prediction process can be reduced. Many scholars have proposed some pre-judging algorithms. [4] proposed a pre-process method to preclude some of the less probable intra prediction modes without transform and quantization. [5] proposed a method to reduce the intra prediction mode by simplifying the rate distortion model. On the basis of [5], [6] improved and improved the accuracy of the prediction [7]. Reduce computational complexity by reducing the number of prediction modes in 64*64 blocks. These methods are all studied how to reduce the number of intra prediction modes to be traversed for a given size of intra prediction unit type.

[8] assumed the texture of images to have consistent features on consecutive video frams, and proposes intra prediction using the Co-located PU and adjacent PU block intra prediction modes to reduce the RMD algorithm process. The number of modes. Based on this information, the number of intra prediction modes in the RMD selection process was reduced to 17, but this method causes a significant increase in the bit rate.

[9] used the transform domain edge detection to reduce the selection range of the intra prediction mode. For different PU block sizes, different class coefficients are used to calculate the sum of the absolute values of the selected class coefficients; then, according to the calculated sum values, different weights are selected to select the intra prediction mode. Thus, only a portion of the intra prediction mode is calculated. However, this method has an average bit rate increase of 2.8% compared to the original code.

[10] proposed an algorithm that reduces the number of candidate prediction modes using image edge detection and texture features. The algorithm used the edge matching detector to select the intra prediction mode, and used the kernel density estimation method similar to the statistical histogram to improve the accuracy of the edge matching detector, which reduced the encoding time by an average of 37.6%. The code rate was increased by an average of 1.65%.

In [11], a gradient-based HEVC intra prediction fast mode selection algorithm was proposed. The main idea is to use the Sobel operator to represent the gradient assignment of each mode in the form of a histogram, and then select the corresponding number of candidate modes according to the size of different PUs. While reducing the encoding time by 20%, the average bit rate is increased by 1%.

The HEVC intra prediction algorithms proposed above improve the coding efficiency of HEVC to a certain extent, but did not fully utilize the texture characteristics and spatial correlation of the image block itself, and both cause a significant decrease in coding quality. There are also texture based fast intra prediction mode selection algorithms, but uses different features or classifier. Their performance will be compared with the proposed algorithm in Section 3. The algorithm we proposed for the fast selection of HEVC intra prediction modes makes full use of the texture characteristics, and uses machine learning algorithm to accurately predict the intra prediction mode of the image block of any size, which improves the encoding speed and maintains the speed compared with the previous algorithm. Better coding quality. Experiments show that by using the image texture direction and spatial correlation to predict the intra prediction mode, the number of intra prediction modes entering the RMD and RDO processes is reduced, and the coding speed is improved. Section 2 of the paper describes the principle of the algorithm how it is trained. Section 3 shows the test results and comparison with similar algorithms. Section 4 briefly concludes the work.

2. METHODOLOGY

2.1 Image feature extraction

There are 35 different prediction modes in Intra luma prediction, Intra Planar mode, Intra DC mode and 33 Intra Angular modes. Intra Planar is suitable for smooth color transitions, Intra DC is suitable for unchanging colors, and Intra Angular is for predicting images from different angles.

Intra luma prediction has a total of 5 modes, including Planar, vertical mode (equivalent to Intra Angular 1), horizontal mode (equivalent to Intra Angular 10), DC mode and derived mode. The derived mode uses the same prediction mode as Intra luma prediction.

Therefore, to determine the mode of use through machine learning, we need to find features that relate to angles. The principle of intra angular mode coding is assuming that the Figure 2 shows examples of 8x8 luma prediction blocks generated with all the HEVC intra prediction modes. It is obvious to observe that the color transition is relevant to prediction angles.

We used the variance of the pixel values along the prediction direction as the feature of the image. Take mode 2 in Figure 1 as an example, in an 8*8 PU, the considered pixels are along the lines, as is shown in Figure 3 (left). The 3 variance values of the pixel values are the 3 features for mode 2. In order to reduce the computational responsibility, only three sets of features are extracted for each prediction angle. The same process goes on for mode 3, 4… In total 99 features will be extracted for each PU. If a line goes through 2 pixels in the same row or column, the one closer to the line is taken into account. Take mode 31 which is shown in Figure 3 (right) as an example, one of the lines (the one in the middle), representing the prediction angle, went through 12 pixels, the gray ones were selected. No interpolation algorithm is used in this case. As for the Intra Planar mode and the Intra DC mode (the upper two in Figure 2), although there is no relationship between the prediction angle and the angle, the pixel value of the prediction angle can also represent the smoothness of the image, so it still has reference value for the above two modes.

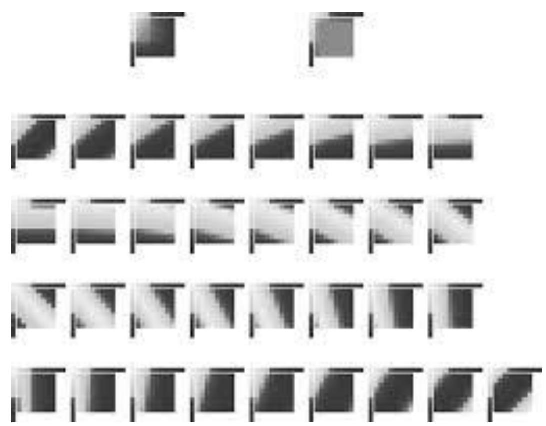

Figure 2. HEVC intra prediction angle texture

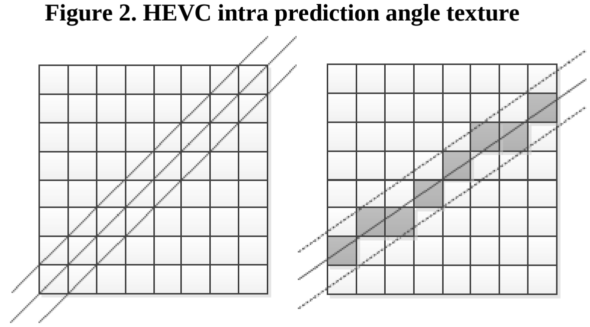

Figure 3. HEVC intra prediction angle feature

2.2 Database setup

The training database is extracted from the YUV test sequence. We say that the test sequence is run with HM. In the process, the generated PU is output as a bitmap, and its size, location and prediction mode are marked as a training sample. Some samples are shown in Figure 4.

We extracted the PUs of several YUV test sequences as training samples. 5 YUV sequences: BasketballDrive, Traffic, PeopleOnStreet, Tennis and PartyScene were used for PU sample extraction. In total there were 4,500,000 PU samples stored in the database. The database can be used not only as a training set for current projects, but also as samples for similar projects on the future.

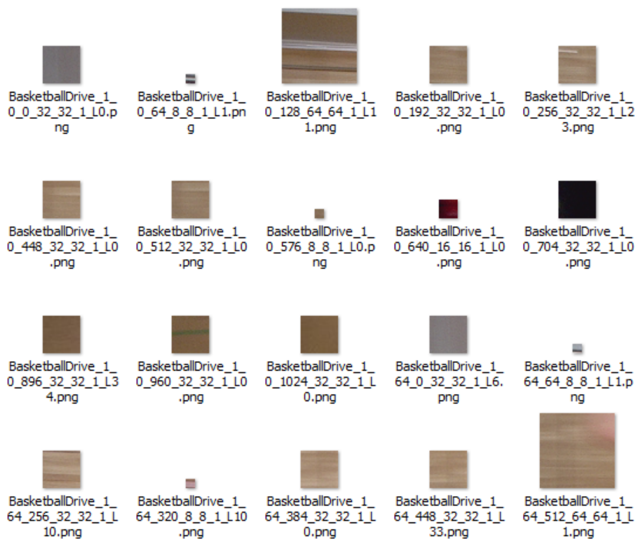

Figure 4. Example of CU samples for training

The training sample extraction codes were based on HM 16.0, as well as the test codes. The codes in HM 16.0 that select the prediction mode (check RDC) was replaced with a trained classifier. In the original program, the algorithm finds the best prediction mode by traversing all prediction modes. The proposed method uses machine learning algorithms to analyze the color and texture of the image and directly determines the prediction mode used, which can greatly save the required running time. For the test 11 different YUV sequences were used and is detailed in the results section.

2.3 Prediction mode decision

Since HEVC defines a large number of modes, it is unrealistic to perform rate-distortion calculations for all modes in most cases. In the HEVC reference code, SATD is used to filter candidate luma intra prediction modes prior to rate-distortion optimization. The number of prediction modes entering the Full RDO is determined according to the size of the corresponding PU, as is shown in Table 1. For these modes and candidate modes in the MPM, prediction and transform operations are performed to obtain the required amount of encoded data and corresponding distortion. The mode with the lowest rate distortion is selected. For chroma intra coding, all possible prediction modes are evaluated based on the rate distortion penalty due to the small number of modes [12].

The method proposed in this paper firstly work out a prediction mode using machine learning base on the feature of the current PU, then the resultant mode is compared with the three most probable mode modes. If the predicted mode is a member of the MPMs, then this mode is used to encode the CU. Otherwise run SATD to calculate the 8 or 3 candidates. If the predicted mode is a member of the candidates, then this mode is used to encode the CU. Otherwise run RDO to find out the prediction, as the original algorithm does. Algorithm 1 shows the pseudo code of prediction mode decision.

Chroma prediction is simpler than luma prediction and will not run either SATD or RDO (Algorithm 2).

Table 1. Number prediction candidates chosen by SATD

PU size	Number
4*4	8
8*8	8
16*16	3
32*32	3
64*64	3

Algorithm 1: Pseudo code for Luma prediction mode prediction using proposed method

Run proposed machine learning prediction;
Run getIntraDirLumaPredictor() (work out MPMs);
if predicted mode is a member of the MPMs **then**
    Encode current PU using predicted mode;
else
    run SATD;
    if predicted mode is a member of the SATD candidates **then**
        Encode current PU using predicted mode;
    else
        Run original RDO to find out the prediction;

Algorithm 2: Pseudo code for chroma prediction mode prediction using proposed method

Run proposed machine learning prediction;
Run getIntraDirPredictor() (work out MPMs);
if predicted mode is a member of the MPMs then
    Encode current PU using predicted mode;
else
    Encode current PU using mode 4
    (same as corresponding luma prediction mode);

3. RESULTS AND DISCUSSION

After extracting the features of the images, we found that the difference in using different kinds of machine learning algorithms is actually not very large.

Because the basis of the judgment mainly comes from the extracted features, various machine learning algorithms can make a more accurate judgment on the newly input image after proper training.

However, there are still some performance differences between different machine learning algorithms, mainly reflected in training time. Table 2 shows the accuracy of the four failed machine learning algorithms and the time required for training, including a simple neural network model. MATLAB was used for the simulation, 1000 samples were used for training and all the settings remained default.

Table 2. Prediction performance of different machine learning techniques

	Accuracy	Time(s)
Random forest	98.23%	577
SVM	95.75%	867
kNN	97.56%	1243
BP neural network	97.88%	2452

We finally chose random forest as it appears to have the fastest training speed, and the accuracy is also high enough compared with other methods in this particular case. The random forest model sets up 400 decision trees. The number of features selected to train on each tree is 11. To reduce the complexity of the operation, 1/10 training samples will be randomly used for training. When the instances on each leaf node exceeds 10, the training will stop. 100,000 random samples in the database were used as training samples. The configuration file "encoder_intra_main.cfg" were used for HM 16.0. This trained model was then used as our intra prediction mode selector, and processed several YUV sequences. The results, including video quality in terms of BDBR [13], and reduced time, i.e. reduced complexity, are shown in Table 3.

Table 3. BD-rate performance of the proposed algorithm

YUV sequence	BDBR	Reduced complexity (time saved)
Map	2.01%	53.34%
Programming	2.32%	55.25%
SlideShow	2.54%	53.45%
WebBrowsing	2.34%	52.67%
Robot	1.32%	54.23%
BasketballScreen	2.16%	64.33%
MissionControl2	2.18%	58.54%
Console	2.78%	56.73%
Desktop	2.64%	57.66%
FlyingGraphics	1.23%	68.48%
MissionControl3	1.47%	59.47%

The search results have been compared against other similar algorithms in Table 4. The fast software-oriented algorithms, namely [6], [14], citemin2015fast, provide a higher coding quality. However, the above method has the following obstacles: First, the CTU-die maximum coding complexity is not pruned, which prevents the fast algorithm from helping to optimize the hardware cost of the encoder. For example, in order to maintain coding quality, [6], [15] defines a conservative threshold. The CU split test is terminated only when one CU meets these thresholds; otherwise, the CU mode is determined to be executed as the original full RDO process. We can see that the fast CU mode method in [6] only reduces the overall coding time by 26%. Second, the CU-level data dependency in [14], which uses current CU depth coding information to prun deeper CU-level RDO processing, leads to the inconvenience of coding parallelism. That is, if this algorithm is employed, the encoding throughput will be degraded.

Table 4. BDBR performance and complexity comparison with similar algorithms

Algorithm	BDBR	Reduced complexity (time saved)
Effective CU size decision for HEVC intracoding [14]	1.08%	47%
Fast intra mode decision for high efficiency video coding (HEVC) [6]	1.06%	26%
Gradient-based fast decision for intra prediction in HEVC [11] /[19]	5.10%	52%
HEVC intra prediction acceleration based on texture direction and prediction unit modes reuse [15]	N/A	44%
HDTV1080p HEVC intra encoder with source texture based CU/PU mode pre-decision [16]	4.53%	62%
VLSI friendly fast CU/PU mode decision for HEVC intra encoding: Leveraging convolution neural network [17]	3.39%	61%
CU Partition Mode Decision for HEVC Hardwired Intra Encoder Using Convolution Neural Network [18]	2.67%	61%
Proposed in this paper	2.09%	58%

[15], [16], [17], [18] considered the texture of PU blocks developed acceleration algorithms based on that. However [15] and [16] did not use machine learning as their classifiers, which means the classifiers were not trained with large date and the accuracy might be increased if a machine learning technique was used. [17] and [18] used CNN to accelerate the mode decision process. CNN does not require specific features but finds them during the training process, makes the feature finding process in black-box. That leads to good results in terms of quality and efficiency, but lacks explanatory of how the texture was processed.

References [16], [17], [19] have VLSI solution because these methods only use source image texture analysis to predict promising CU pattern competitors. For the CU mode, the pre-decision engine can be allocated in the advanced CTU pipeline stage without hindering the rhythm of subsequent RDO-based CTU coding. However, since [16] and [19] do not use the topology information of the feature points, their coding efficiency loss is obvious (BDBR = + 5.10 % and BDBR = + 4.53 %).

As is shown in Table 4, some algorithms have higher image quality but not as much amount of reduced complexity ([6], [14], [15]), in other words, they require longer processing duration. Our algorithm has very similar complexity as [16], [17] and [18] , but has better quality than either of them. Because we used features that has close relationship with the texture of the images, therefore can better represent the color transition in the images. As a result, the prediction accuracy of our algorithm is quite high, and leads to good video quality.

4. CONCLUSION

This paper discussed how to use machine learning to accelerate intra prediction in HEVC. The proposed method uses a machine learning algorithm to determine the value of the angle prediction by extracting the image features of the PUs. Compared with the original RDO method, this method greatly reduces the complexity of coding under the premise of little influence on the picture quality, thus greatly saving the time required for coding.

The key to this prediction method is the extraction and training of image features. 99 features were extracted from each image to ensure that each angle is accurately predicted. Compared with CNN that does not need assigned features and finds features itself, and some other machine learning techniques that choose features have not as relevant to the angle of texture, our features lead to more accurate prediction results, as well as higher video quality.

The proposed algorithm was compared a lot of similar algorithms. The proposed method is the first one uses both angle feature and machine learning technique. It not only Considers the characteristic of HEVC intra prediction - the angle of the texture, but also offers the selector high accuracy and universality through a large number of sample training. Some of similar algorithms used machine learning and improved the coding efficiency of HEVC to a certain extent, but did not fully utilize the texture characteristics of the image block itself. The proposed algorithm makes full use of the texture characteristics of the image block itself, which improves the accuracy of prediction and achieves a good balance between image quality and efficiency.

5. REFERENCES

[1] Buczak, A.L. and Guven, E., 2016. A survey of data mining and machine learning methods for cyber security intrusion detection. *IEEE Communications Surveys & Tutorials*, 18(2), pp.1153-1176.

[2] Baydin, A.G., Pearlmutter, B.A., Radul, A.A. and Siskind, J.M., 2017. Automatic differentiation in machine learning: a survey. *Journal of Machine Learning Research*, 18(153), pp.1-153.

[3] Breiman, L., 2001. Random forests. *Machine learning*, *45*(1), pp.5-32.

[4] Min, Sunil Lee, Il-Koo Kim, Woo-Jin Han, Jani Lainema, and Kemal Ugur. 2010. Unification of the directional intra prediction methods in TMuC. *JCTVC-B100, Geneva, Switzerland* (2010).

[5] Yinji Piao, Junghye Min, Jianle Chen, et al. 2010. Encoder improvement of unified intra prediction. *JCTVC-C207* 7674 (2010), 568–577.

[6] Hao Zhang and Zhan Ma. 2014. Fast intra mode decision for high efficiency video coding (HEVC). *IEEE Transactions on circuits and systems for video technology* 24, 4 (2014), 660–668.

[7] K Sugimoto, A Minezawa, and S Sekiguchi. 2011. Reduced Number of Intra 64* 64 Prediction Mode.

[8] Ajit Singh Motra, Ankuj Gupta, Mahesh Shukla, Pankaj Bansal, et al. 2012. Fast intra mode decision for HEVC video encoder. In *Software, Telecommunications and Computer Networks (SoftCOM), 2012 20th International Conferenc*e on. IEEE, 1–5.

[9] Yi-Ching Ting and Tian-Sheuan Chang. 2012. Fast intra prediction algorithm with transform domain edge detection for HEVC. In *Circuits and Systems (APCCAS), 2012 IEEE Asia Pacific Conference on*. IEEE, 144–147.

[10] Guang Chen, Zhenyu Liu, Takeshi Ikenaga, and Dongsheng Wang. 2013. Fast HEVC intra mode decision using matching edge detector and kernel density estimation alike histogram generation. *In Circuits and Systems (ISCAS), 2013 IEEE International Symposium on*. IEEE, 53–56.

[11] Wei Jiang, Hanjie Ma, and Yaowu Chen. 2012. Gradient based fast mode decision algorithm for intra prediction in HEVC. In *Consumer Electronics, Communications and Networks (CECNet), 2012 2nd International Conference on*. IEEE, 1836–1840.

[12] Dilip Prasanna Kumar. 2013. *Intra frame luma prediction using neural networks in HEVC*. Ph.D. Dissertation. University of Texas at Arlington.

[13] Gisle Bjontegaard. 2001. Calculation of average PSNR differences between RD-curves. *VCEG-M33* (2001).

[14] Liquan Shen, Zhaoyang Zhang, and Zhi Liu. 2014. Effective CU size decision for HEVC intracoding. *IEEE Trans. Image Processing* 23, 10 (2014), 4232–4241.

[15] da Silva, T.L., Agostini, L.V. and da Silva Cruz, L.A., 2014. HEVC intra prediction acceleration based on texture direction and prediction unit modes reuse. *APSIPA Transactions on Signal and Information Processing*, 3.

[16] Jia Zhu, Zhenyu Liu, Dongsheng Wang, Qingrui Han, and Yang Song. 2014. HDTV1080p HEVC Intra encoder with source texture based CU/PU mode pre-decision. In *Design Automation Conference (ASP-DAC), 2014 19th Asia and South Pacific*. IEEE, 367–372.

[17] Xianyu Yu, Zhenyu Liu, Junjie Liu, Yuan Gao, and Dongsheng Wang. 2015. VLSI friendly fast CU/PU mode decision for HEVC intra encoding: Leveraging convolution neural network. In *Image Processing (ICIP), 2015 IEEE International Conference on*. IEEE, 1285–1289.

[18] Zhenyu Liu, Xianyu Yu, Yuan Gao, Shaolin Chen, Xiangyang Ji, and Dongsheng Wang. 2016. CU partition mode decision for HEVC hardwired intra encoder using convolution neural network. *IEEE Transactions on Image Processing* 25, 11 (2016), 5088–5103.

[19] Yongfei Zhang, Zhe Li, and Bo Li. 2012. Gradient-based fast decision for intra prediction in HEVC. In *Visual Communications and Image Processing (VCIP)*, 2012 IEEE. IEEE, 1–6