

# A Fast Rough Mode Decision Algorithm for HEVC

Wei-Xin Yao\*, Dan Yang\*\*, Gui-Fu Lu\*\*, and Jun Wang\*\*

## Abstract

HEVC is the high efficiency video coding standard, which provides better coding efficiency contrasted with the other video coding standard. But at the same time the computational complexity increases drastically. Thirty-five kinds of intra-prediction modes are defined in HEVC, while 9 kinds of intra prediction modes are defined in H.264/AVC. This paper proposes a fast rough mode decision (RMD) algorithm which adopts the smoothness of the up-reference pixels and the left-reference pixels to decrease the computational complexity. The three step search method is implemented in RMD process. The experimental results compared with HM13.0 indicate that the proposed algorithm can save 39.7% of the encoding time, while Bjontegaard delta bitrate (BDBR) is increased slightly by 1.35% and Bjontegaard delta peak signal-to-noise ratio (BDPSNR) loss is negligible.

## Keywords

HEVC, Intra Prediction, Rough Mode Decision, Video Coding

## 1. Introduction

High efficiency video coding (HEVC) standard uses various coding tool to enhance compression performance. Among them, intra prediction introduces 35 angular prediction modes and a flexible coding unit (CU) size based on quadtree. Coding tree unit (CTU) is defined in HEVC, whose size is  $64 \times 64$ . As root of encoding unit, CTU is divided into smaller CUs. A CU is further divided into prediction unit (PU) and transform unit (TU) [1].

To decide the optimal prediction mode and CU size, HEVC makes use of rate distortion optimization (RDO). However, HEVC increases computational burden compared with H.264. At present, a research hotspot is focused on the research of fast intra prediction mode. The research works are lying in two aspects: fast intra prediction mode decision algorithms (FIPMD) and the fast CU or (PU) size decision algorithms [2].

FIPMD algorithms decrease the number of candidate modes before compute RD cost. In [3], the authors proposed improved FIPMD based on gradient. The prediction mode is decided by the texture direction, which is denoted by the intensity gradient. In [4], progressive rough mode search (PRMS) which is based on the calculation of Hadamard cost is proposed. Rate-distortion optimized quantization (RDOQ) is implemented by fewer effective candidates, which are selected by PRMS. In [5], edge of the

※ This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.  
Manuscript received September 19, 2018; first revision December 6, 2017; accepted January 29, 2018

Corresponding Author: Wei-Xin Yao (1285371089@qq.com)

\* School of Electrical Engineering, Anhui Polytechnic University, Wuhu, China (1285371089@qq.com)

\*\*School of Computer Science and Information, Anhui Polytechnic University, Wuhu, China (503308723@qq.com, {luguifu\_jsj, wangjunahpu}@163.com)

PU was also used to speed-up the best selection procedure. In order to obtain the minimal direction differences, Wang and Siu [6] made use of the features of every directional prediction mode to calculate the strength of directional errors of initial spatial domain. Na et al. [7] made use of spatial correlation to detect edges. The candidate modes are adaptively chosen by the edge map. Lee et al. [8] analyzed the probability distribution of local binary patterns to obtain textural information of the most probable mode.

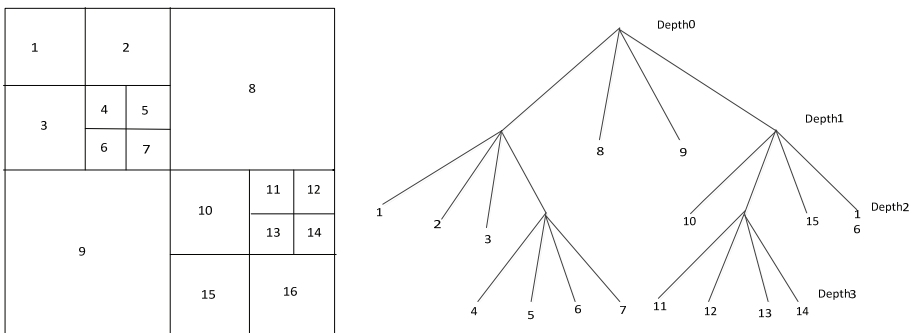
The fast CU or PU size decision algorithms can skip unnecessary CU size partition [9-14]. In [9], the proposed method makes use of motion homogeneity to decide CU depth range. Unnecessary depth levels do not need to be traversed. In [10], according to regular updates of the statistical parameter, CU depth decision was early terminated. In [11], a novel CU size decision strategy is presented by the authors, which is based on the local and global edge complexities. In order to reduce complexity of the screen content coding, Duanmu et al. [12] proposed machine learning based on CU statistics and sub-CU homogeneity. Yu et al. [13] adopted convolution neural network to reduce the amount of CU/PU candidate modes, which decrease the complexity of the real-time hardwired encoder. According to the depth of adjacent CUs, CU size decision was early terminated. In terms of correlation between the current layer mode and the higher layer mode in PU, the authors propose a new fast PU mode decision algorithm [14].

In this paper, we analyze the reference pixels of the PU to exclude some angular intra prediction modes before intra prediction. The amount of candidate modes is decreased from 35 to 19. The three step search method is implemented by 19 modes. The proposed algorithm has many advantages such as easy implement, good performance and high robustness.

The rest of the paper is arranged as followed. A brief review of intra coding process is described in Section 2. In Section 3, a fast rough mode decision algorithm is presented. Experiments and the corresponding result analysis of the proposed algorithm are presented in Section 4. Finally, some inspiring conclusions are given in Section 5.

## 2. INTRA-PREDICTION IN HEVC

HEVC makes use of quadtree structure to split large code unit (LCU) into CU sizes ranging from  $8 \times 8$  to  $64 \times 64$ . The numbers denote depth levels in Fig. 1. Depth 3, 2, 1, and 0 correspond to the CU size of  $8 \times 8$ ,  $16 \times 16$ ,  $32 \times 32$ , and  $64 \times 64$ , respectively. A CU is divided into PUs, which are used as basic units for intra-prediction. According to the adjacent decoded pixels, the predicted pixels are obtained by linear interpolation. Two PU sizes are  $M \times M$  and  $2M \times 2M$ . Therefore, the PU sizes range from  $4 \times 4$  to  $64 \times 64$ .



**Fig. 1.** LCU partition and its corresponding quadtree.

The CUs are also divided into TUs according to quadtree structure. TU is used for transformation and quantization. The TU sizes range from 32×32 to 4×4. Fig. 2 shows the process of dividing a CU of 32×32 size into PU and TU.

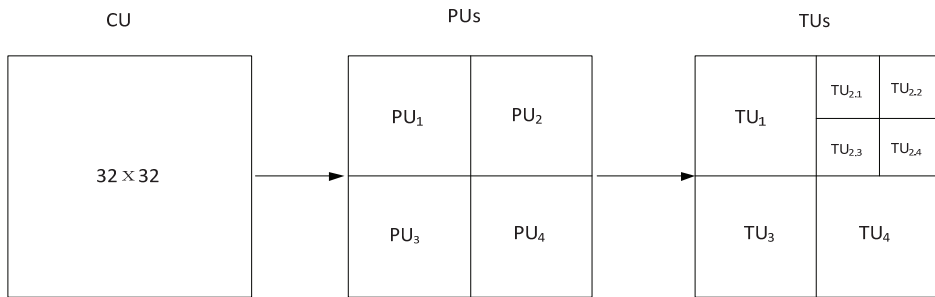


Fig. 2. The process of dividing a CU of 32×32 size into PU and TU.

### 2.1 Rough Mode Decision

HEVC supports 35 intra-prediction modes which are shown in Fig. 3. The order number 0 and 1 represent planar mode and DC mode. The order numbers from 2 to 34 represent 33 angular prediction modes. The N most possible candidate modes are obtained by the rough mode decision (RMD). In the RMD processes, the Hadamard cost and the bit cost of 35 prediction modes are computed. N is related to PU size. N is 8, 8, 3, 3, and 3, respectively corresponding to the PU size of 4×4, 8×8, 16×16, 32×32, and 64×64. The N lowest Hadamard cost is selected to a subset of candidate, as defined in Eq. (1) [15].

$$H_{cost} = SATD + \lambda R_{mode} \tag{1}$$

where Hadamard transformation is a generalized Fourier transformation and SATD denotes the sum of absolute Hadamard transformed residual signal,  $R_{mode}$  represents bit costs for the prediction mode, and  $\lambda$  is the Lagrange multiplier.

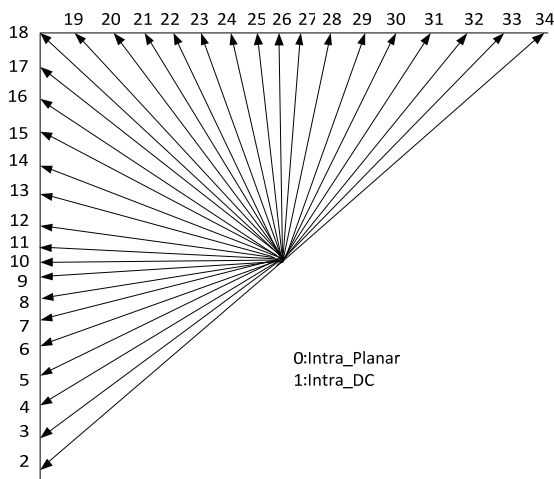


Fig. 3. Thirty-five intra-prediction modes in HEVC.

After the RMD processes, we apply RDO to select the optimal prediction mode for PU. RDO is implemented in the subset of candidate. The least RDO cost is selected by function (2).

$$RDO_{cost} = SSE + \lambda R \tag{2}$$

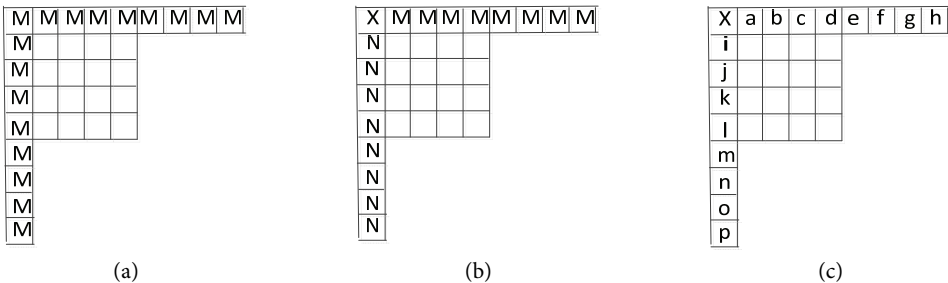
SSE represents the sum of squared difference of the predicted against original PU. R and  $\lambda$  denote bit-rate for encoding process and Lagrange parameter, respectively.

### 2.2 Most Probable Modes

Intra-prediction mode is fast decided by Zhao et al. [16]. According to spatial correlation, most probable modes (MPMs) of the current PU have a great correlation with the prediction modes of neighboring PUs. The optimal prediction mode of PU exists in the MPMs, which is very probable. Even in different types of test sequences, this probability fluctuation is very small. MPMs are inserted into the subset of candidate if modes of above and left neighboring PUs are not included in the subset of candidate. RDO is applied to N+MPM modes.

### 3. Fast Rough Mode Decision

However, 35 kinds of prediction mode are calculated in RMD process, which requires too much time. It is required that a fast RMD algorithm is proposed.



**Fig. 4.** Three cases of reference pixels for the prediction block: (a) all the same reference pixels, (b) the same up-reference pixels and the same left-reference pixels, and (c) different reference pixels.

Fig. 4 shows the reference pixels for the 4x4 block. When the reference pixels are the same, the current PU block is a smoothing block. Planar mode is selected, and the other modes are skipped. The Planar mode adopts the mean of linear interpolation in horizontal and vertical directions to obtain the prediction values of PUs. DC mode adopts the mean of reference pixels to obtain the prediction values of PUs. DC mode and planar mode are suitable for blocks containing smooth texture. However, planar mode is appropriate for blocks with gradient texture. The planar mode maintains the maximum continuity of block boundaries. We skip the unnecessary modes based on smoothness of the reference pixels. We can define the smoothness of the reference pixels as following [17].

$$Var_{up} = \frac{1}{2N+1} \sum_{k=0}^{2N} (f_k^{up} - U_{up})^2 \tag{3}$$

$$Var_{left} = \frac{1}{2N+1} \sum_{K=0}^{2N} (f_k^{left} - U_{left})^2 \tag{4}$$

where  $U_{up}$  denotes average of the up-reference pixels,  $U_{left}$  denotes average of the left-reference pixels,  $N$  represents size of the current block,  $f_k^{up}$  denotes the up-reference pixel associated with index  $K$ , and  $f_k^{left}$  denotes left-reference pixel associated with index  $K$ . If  $Var_{up}$  is equal to  $Var_{left}$ , Planar mode is selected for the current block. If  $Var_{up}$  is larger than  $Var_{left}$ , the up-reference pixels are smoother than the left-reference pixels and RMD process is performed by 18–35 prediction modes. If  $Var_{up}$  is smaller than  $Var_{left}$ , the left-reference pixels are smoother than the up-reference pixels and RMD process is performed by 0–18 prediction modes.

The three step search method is implemented by using  $0, 1, 2+4\delta$ , where  $\delta \in [0,4]$  or  $\delta \in [4,8]$ . The subset of candidate mode is selected roughly by using step size 4. Then the subset of candidate mode is expanded by using step size 2 and 1. Taken  $16 \times 16$  as an example, 0–18 prediction modes are selected according to smoothness of the reference pixels. Firstly, step size 4 for search is performed in range of mode index 0–18, which constituted the subset of candidate mode. Hadamard cost is calculated by using the subset and the ascending order of Hadamard cost is  $\{10, 6, 14, 18, 2, 1, 0\}$ . We select the three best mode as  $\xi_1\{10, 6, 14\}$ . Secondly, the neighboring modes of two distances are added in  $\{10,6,14\}$ . Then, the subset is updated as  $\{8, 10, 6, 4, 12, 14, 16, 18, 2, 1, 0\}$  and selects the three best mode as  $\xi_2\{8,10,6\}$ . At last, the neighboring modes of one distance are added in  $\{8, 10, 6\}$  and the subset is updated as  $\{8, 9, 10, 7, 6, 11, 5, 4, 12, 14, 16, 18, 2, 1, 0\}$ . RDO process is implemented by  $\xi_3\{8, 9, 10\}$ . The three step search process is shown in the Fig. 5.

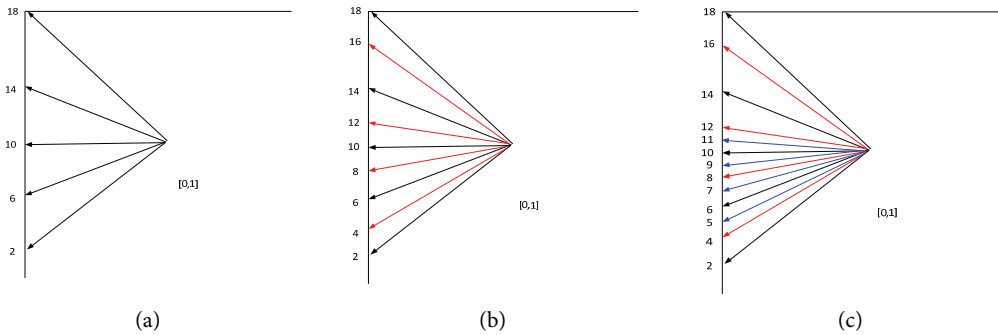


Fig. 5. The three step search method: (a) the first step, (b) the second step, and (c) the third step.

### 4. Experimental Results

It is performed on HM13.0 reference software that fast mode decision proposed. The experiment environment is Intel Core i7-6500 2.5 GHz and RAM of 4 GB. This paper makes use of 100 frames of the video sequences to experiment. It is presented by JCJ-VC the video sequences, containing five classes. Different QPs (37, 32, 27, 22) are tested in the experiment. To evaluate performance of the proposed algorithm, the parameters are measured according to Bjontegaard delta bitrate (BDBR), Bjontegaard delta peak signal-to-noise ratio (BDPSNR) and encoding time saving. BDBR and BDPSNR are used to represent the average differences. Encoding time saving represents the total time reduction in percentage compared with HM13.0. Encoding time saving is defined by formula (5).

$$Time(\%) = \frac{T_{prop} - T_{HM}}{T_{HM}} \times 100\% \quad (5)$$

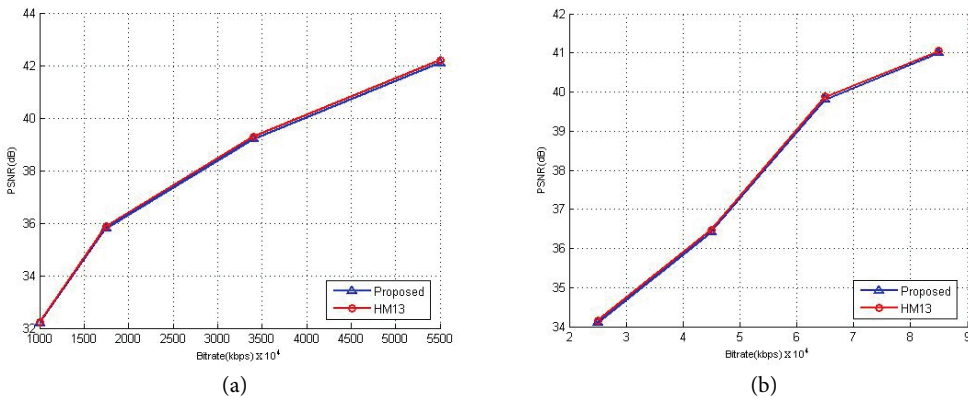
where  $T_{HM}$  and  $T_{prop}$  present the encoding time which is required by HM13.0 and the proposed algorithm.

Table 1 shows that experiment results compared between Gao algorithm and the proposed algorithm. The fast rough mode decision algorithm can save time from 30.0 to 39.7, while it increases BDBR by between 0.56 and 2.64, and decreases BDPSNR by between 0.01 and 1.01. Gao algorithm [18] increases less BDBR than our algorithm, but our algorithm can save more time. It is negligible that BDBR increase and BDPSNR loss.

**Table 1.** Comparison of coding performance (%) of two different algorithms

Video classes	Sequence	Proposed algorithm			Gao algorithm		
		BDBR	BDPSNR	$\Delta T$	BDBR	BDPSNR	$\Delta T$
Class A (2560×1600)	Traffic	1.10	-0.07	35.1	0.92	-0.05	27.3
	People on street	0.7	-0.05	38.2	0.91	-0.06	24.7
Class B (1920×1080)	KimonoI	0.56	-0.06	30.0	1.31	-0.04	24.1
	Cactus	0.81	-0.07	34.2	1.07	-0.03	24.2
Class C (832×1080)	BasketballDrill	2.01	-0.07	38.4	1.21	-0.05	28.6
	BQMall	0.85	-0.01	31.5	0.93	-0.08	34.2
Class D (416×240)	BasketballPass	1.98	-0.12	39.7	1.2	-0.08	32.5
	BQSquare	2.64	-0.22	36.6	2.78	-0.18	30.1
Class E (1280×720)	FourPeople	1.91	-0.86	30.4	1.10	-0.06	22.3
	Johnny	0.94	-1.01	35.7	1.34	-0.06	29.7
Average		1.35	-0.25	34.98	1.28	-0.07	27.8

We can see the RD performances of the proposed algorithm are close to that of HM13.0 from Fig. 6. The encoding efficiency of the proposed algorithm is approximately equal to that of HM13.0, while the proposed algorithm can save more time.



**Fig. 6.** Comparison of the RD curve of the proposed algorithm and HM13.0: (a) basketball and (b) people on the street.

## 5. Conclusion

The paper proposes the fast RMD decision algorithm. The proposed algorithm is implemented by using two methods. Firstly, we compare the upper and left reference pixels to decrease the amount of candidate

modes from 35 to 19. Secondly, the three step search method is implemented in 19 candidate modes to reduce computational complexity. It is the final result that 12–15 modes are performed in RMD process instead of 35 modes. Our algorithm can save a lot of time, while RD performance is approximately unchanged.

## Acknowledgement

This work is supported by the National Natural Science Foundation of China (No. 61572033) and Key Projects of Natural Science Research Fund of Anhui Province (No. KJ2015A311), and Provincial Natural Science Research General Project of Anhui higher Education Promotion Plan in 2014 (No. TSKJ2014B07).

## References

- [1] M. Ramezanpour and F. Zargari, "Fast CU size and prediction mode decision method for HEVC encoder based on spatial features," *Signal, Image and Video Processing*, vol. 10, no. 7, pp. 1233-1240, 2016.
- [2] G. J. Sullivan, J. R. Ohm, J. W. Han, and T. Wiegand, "Overview of the high efficiency video coding (HEVC) standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1649-1668, 2012.
- [3] Y. Zhang, Z. Li, and B. Li, "Gradient-based fast decision for intra prediction in HEVC," in *Proceedings of 2012 Visual Communications and Image Processing*, San Diego, CA, 2012, pp. 1-6.
- [4] H. Zhang and Z. Ma, "Fast intra mode decision for high efficiency video coding (HEVC)," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 24, no. 4, pp. 660-668, 2013.
- [5] T. L. da Silva, L. V. Agostini, and L. A. da Silva Cruz, "Fast HEVC intra prediction mode decision based on EDGE direction information," in *Proceedings of 2012 Proceedings of the 20th European Signal Processing Conference (EUSIPCO)*, Bucharest, Romania, 2012, pp. 1214-1218.
- [6] L. L. Wang and W. C. Siu, "H. 264 fast intra mode selection algorithm based on direction difference measure in the pixel domain," in *Proceedings of 2009 IEEE International Conference on Acoustics, Speech and Signal Processing*, Taipei, Taiwan, 2009, pp. 1037-1040.
- [7] S. Na, W. Lee, and K. Yoo, "Edge-based fast mode decision algorithm for intra prediction in HEVC," in *Proceedings of 2014 IEEE International Conference on Consumer Electronics (ICCE)*, Las Vegas, NV, 2014, pp. 11-14.
- [8] J. H. Lee, K. S. Jang, B. G. Kim, S. Jeong, and J. S. Choi, "Fast intra mode decision algorithm based on local binary patterns in High Efficiency Video Coding (HEVC)," in *Proceedings of 2015 IEEE International Conference on Consumer Electronics (ICCE)*, Las Vegas, NV, 2015, pp. 270-272.
- [9] L. Shen, Z. Liu, X. Zhang, W. Zhao, and Z. Zhang, "An effective CU size decision method for HEVC encoders," *IEEE Transactions on Multimedia*, vol. 15, no. 2, pp. 465-470, 2013.
- [10] S. Cho and M. Kim, "Fast CU splitting and pruning for suboptimal CU partitioning in HEVC intra coding," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 23, no. 9, pp. 1555-1564, 2013.
- [11] B. Min and R. C. Cheung, "A fast CU size decision algorithm for the HEVC intra encoder," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 25, no. 5, pp. 892-896, 2015.
- [12] F. Duanmu, Z. Ma, and Y. Wang, "Fast CU partition decision using machine learning for screen content compression," in *Proceedings of 2015 IEEE International Conference on Image Processing (ICIP)*, Quebec City, Canada, 2015, pp. 4972-4976.

- [13] X. Yu, Z. Liu, J. Liu, Y. Gao, and D. Wang, "VLSI friendly fast CU/PU mode decision for HEVC intra encoding: leveraging convolution neural network," in *Proceedings of 2015 IEEE International Conference on Image Processing (ICIP)*, Quebec City, Canada, 2015, pp. 1285-1289.
- [14] X. Shang, G. Wang, T. Fan, and Y. Li, "Fast CU size decision and PU mode decision algorithm in HEVC intra coding," in *Proceedings of 2015 IEEE International Conference on Image Processing (ICIP)*, 2015, pp. 1593-1597.
- [15] Q. Zhang, Y. Yang, H. Chang, W. Zhang, and Y. Gan, "Fast intra mode decision for depth coding in 3D-HEVC," *Multidimensional Systems and Signal Processing*, vol. 28, no. 4, pp. 1203-1226, 2017.
- [16] L. Zhao, L. Zhang, S. Ma, and D. Zhao, "Fast mode decision algorithm for intra prediction in HEVC," in *Proceedings of 2011 Visual Communications and Image Processing (VCIP)*, Tainan, Taiwan, 2011, pp. 1-4.
- [17] L. L. Wang and W. C. Siu, "Novel adaptive algorithm for intra prediction with compromised modes skipping and signaling processes in HEVC," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 23, no. 10, pp. 1686-1694, 2013.
- [18] L. Gao, S. Dong, W. Wang, R. Wang, and W. Gao, W. (2015, May). Fast intra mode decision algorithm based on refinement in HEVC. In *2015 IEEE International Symposium on Circuits and Systems (ISCAS)*, Lisbon, Portugal, 2015, pp. 517-520.



**Wei-Xin Yao** <https://orcid.org/0000-0001-5602-2143>

She received M.S. degree in communication and information system from Gui Zhou University, Guiyang, China, in 2007. She is currently a Lecturer with School of Electrical Engineering, Anhui Polytechnic University. Her research interests include video coding and image processing.



**Dan Yang** <https://orcid.org/0000-0003-3898-8627>

He received M.S. degree in computer application technology from Gui Zhou University, Guiyang, China, in 2007. He is currently an Associate Professor with School of Computer Science and Technology, Anhui Polytechnic University. His research interests include video coding and image processing.



**Gui-Fu Lu** <https://orcid.org/0000-0001-9047-6579>

He received the B.S. degree in 1997 from Hefei University of Technology in China, the M.S. degree in 2004 from Hangzhou Institute of Electronics Engineering, and the Ph.D. degree in 2012 from Nanjing University of Science and Information, Since 2004, he has been teaching in the School of Computer Science and Information, Anhui Polytechnic University, Wuhu, China. His research interests include computer vision, digital image processing and pattern recognition.



**Jun Wang** <https://orcid.org/0000-0002-8159-583X>

He was born in 1975, He is currently an Associate Professor with School of Computer Science and Technology, Anhui Polytechnic University. His research interests include image processing, machine vision and intelligent detection.