Fast CU Size Decision Using Machine Learning for Depth Map Coding in 3D-HEVC

Ruyi Zhang*,+, Kebin Jia*,+ , Pengyu Liu*,+

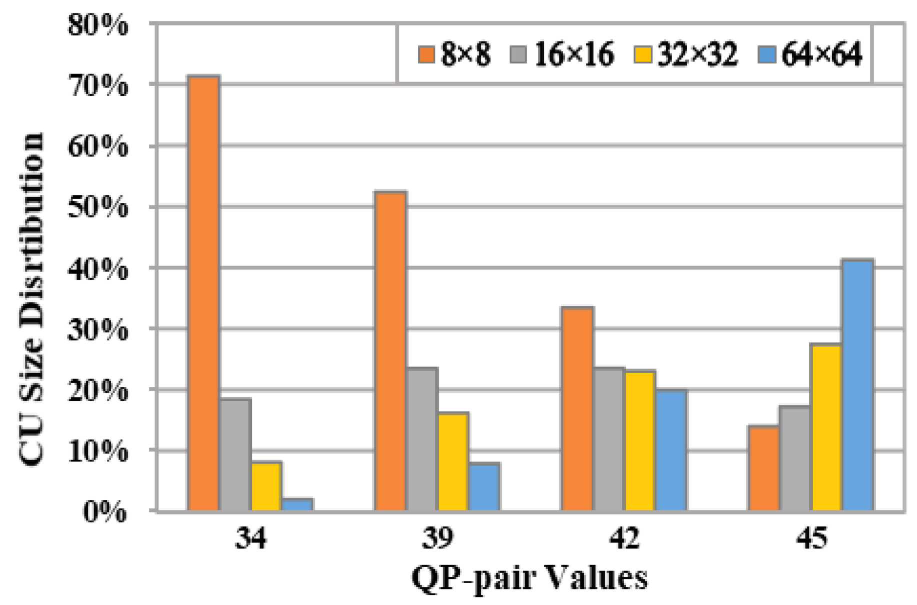
*Faculty of Information Technology, Beijing University of Technology, Beijing, 100124, China

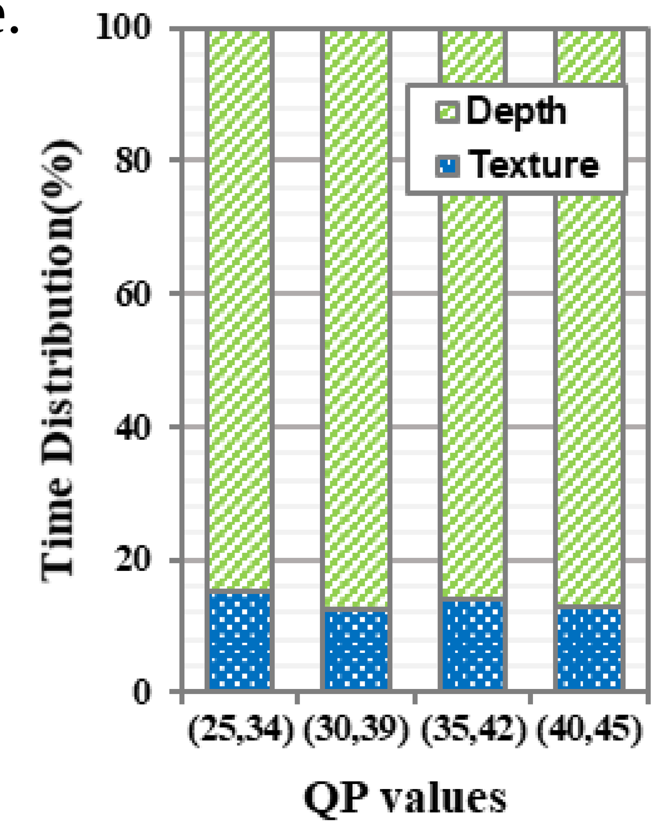
+Beijing Laboratory of Advanced Information Networks, Beijing, 100124, China

zry66882032@sina.com, {kebinj, liupengyu}@bjut.edu.cn

1. INTRODUCTION

- 3D-HEVC inherits the flexible quadtree coding structure of HEVC in both texture and depth maps [1] .This advanced partition structure inevitably leads to a sharp increase in coding complexity while improving coding performance.
- In the All-Intra (AI) configuration of HTM16.0, the depth maps coding occupies about 86% of total 3D-HEVC encoding time and the change in QP values affect the choice of CU size distribution.
- **Motivation:** limiting the continued division of smaller CUs under certain conditions, the coding complexity can be effectively reduced. Although the QP value has a strong correlation with the CU size, it is not feasible to simply remove some depth only based on the QP values. Therefore, it is necessary to design an effective solution to effectively reduce the coding complexity while ensuring the coding performance.

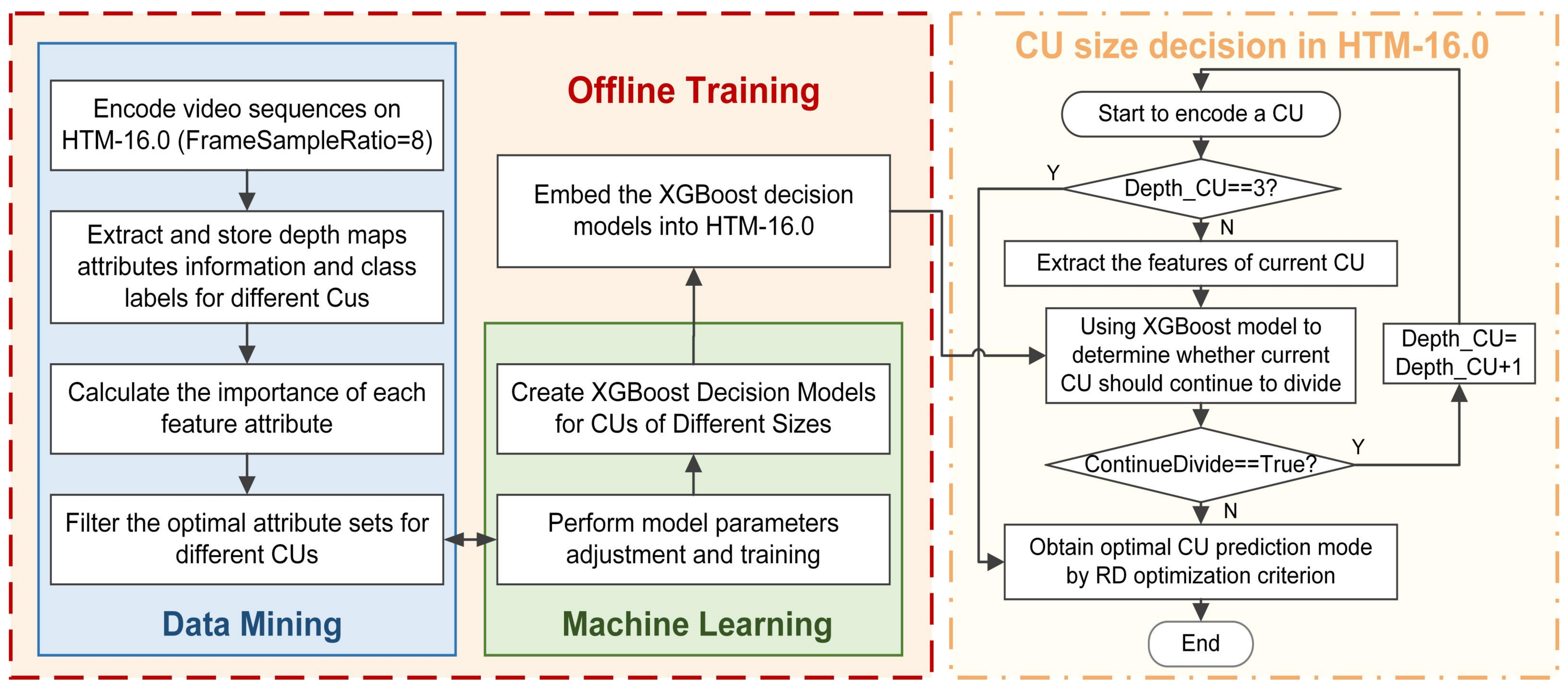




2. PROPOSED METHODS

2.1 Flow Chart

- Extreme Gradient Boosting (XGBoost) [2] models are constructed for different sizes of CUs to adapt to different texture characteristics and make the model more accurate under the AI configuration.



- **Data Mining:** Acquire texture feature attribute information and store the CU divided flags in a one-to-one correspondence.
- **Machine Learning:** Using the optimal feature set selected in the preceding step for different CUs, the XGBoost models are trained by adjusting the model parameters.
- **Model Embedding:** The offline training decision models are embedded in the CU quadtree partitioning process to replace the traditional RDO calculation process.

2.2 Feature attributes selection

Attribute	Description
VAR	$\text{VAR}=\sum_{i=1}^{N}\sum_{j=1}^{N}p(i,j)^2-(\sum_{i=1}^{N}\sum_{j=1}^{N}p(i,j))^2 / N*N$
var_max	Maximum VAR of smaller blocks inside the current CU
MAX_mean	$\text{MAX\_mean}=\max\{p(i,j)-mean\}$
NMSE	$\text{NMSE}=\frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}(p(i,j)-mean)^2$
Complexity	$\text{Complexity}=\sum_{i=1}^{N}\sum_{j=1}^{N}\lvert p(i,j)-mean\rvert$
Ave_Grad	Calculate the mean of the gradient values in each direction according to the Isotropic Sobel operator
Ang_Sobel	Gradient direction calculated from Isotropic Sobel operator
QP	The current QP-depth value

- Considering the impact of model accuracy and the number of feature attributes on the performance of the encoder, different CUs of different sizes have selected different feature attribute subsets.

CU size	Attribute subsets
CU=64	NMSE, VAR, Ang_Sobel, Complexity
CU=32	NMSE, QP, Ang_Sobel, Complexity, var_max, VAR
CU=16	QP, NMSE, C, Ang_Sobel, VAR

3. RESULTS

3.1 Performance comparison with HTM-16.0 (%)

Sequences	V/T	S/T	ETR
Balloons	-0.07	0.24	42.64
Kendo	-0.07	1.01	43.58
Newspaper	-0.18	1.06	45.26
GT_Fly	-0.08	-0.17	49.35
Poznan_Street	-0.12	0.14	39.91
Undo_Dancer	-0.05	0.65	45.87
Shark	-0.28	-0.36	38.06
1024x768	-0.10	0.77	43.82
1920x1088	-0.13	0.07	43.30
Average	-0.12	0.37	43.52

3.2 Performance comparison with similar algorithms (%)

Related Works	S/T	ETR
[3]	-0.01	32.1
[5]	2.67	61.1
[6]	0.51	25.5
This work	0.37	43.52

4. CONCLUSIONS

- This paper proposed a fast 3D-HEVC depth map coding method, which uses the XGBoost decision model to define the partitioning of CUs, replacing the complex RDO calculation in the original encoding process.
- Experimental results demonstrated that this approach achieves on average 43.52% time saving with negligible degradation of coding performance under the all intra configuration, compared with HTM-16.0.
- The proposed algorithm promotes 3D-HEVC to be used in real-time video communication systems.

5. REFERENCES

[1] G. Tech et al., "Overview of the Multiview and 3D Extensions of High Efficiency Video Coding," IEEE Transactions on Circuits and Systems for Video Technology, vol. 26, no. 1, pp. 35-49, Jan. 2016.

[2] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in Proceedings of the 22Nd ACM SIGKDD International Conference on Knowledge Discovery and Data Miningpp. 785–794, 2016 .

[3] K. Peng, J. Chiang, W. Lie, "Low complexity depth intra coding combining fast intra mode and fast CU size decision in 3D-HEVC," IEEE International Conference on Image Processing (ICIP), pp. 1126-1130, 2016.

[4] Z. Liu, X. Yu, Y. Gao, S. Chen, X. Ji and D. Wang, "CU Partition Mode Decision for HEVC Hardwired Intra Encoder Using Convolution Neural Network," in IEEE Transactions on Image Processing (TIP), vol. 25, n. 11, pp. 5088-5103, Nov. 2016

[5] R. Jing, Q. Zhang, B. Wang, P. Cui, T. Yan and J. Huang, "CART-based fast CU size decision and mode decision algorithm for 3D-HEVC," Signal, Image and Video Processing, vol. 13, no. 2, pp. 209-216, 2019.

6. ACKNOWLEDGEMENTS

This work was supported in part by the Project for the National Natural Science Foundation of China under Grants No. 61672064 and Beijing Laboratory of Advanced Information Networks under Grants No.040000546618017.