ICASSP 2017

FAST HEVC INTRA CODING ALGORITHM BASED ON MACHINE LEARNING AND LAPLACIAN TRANSPARENT COMPOSITE MODEL

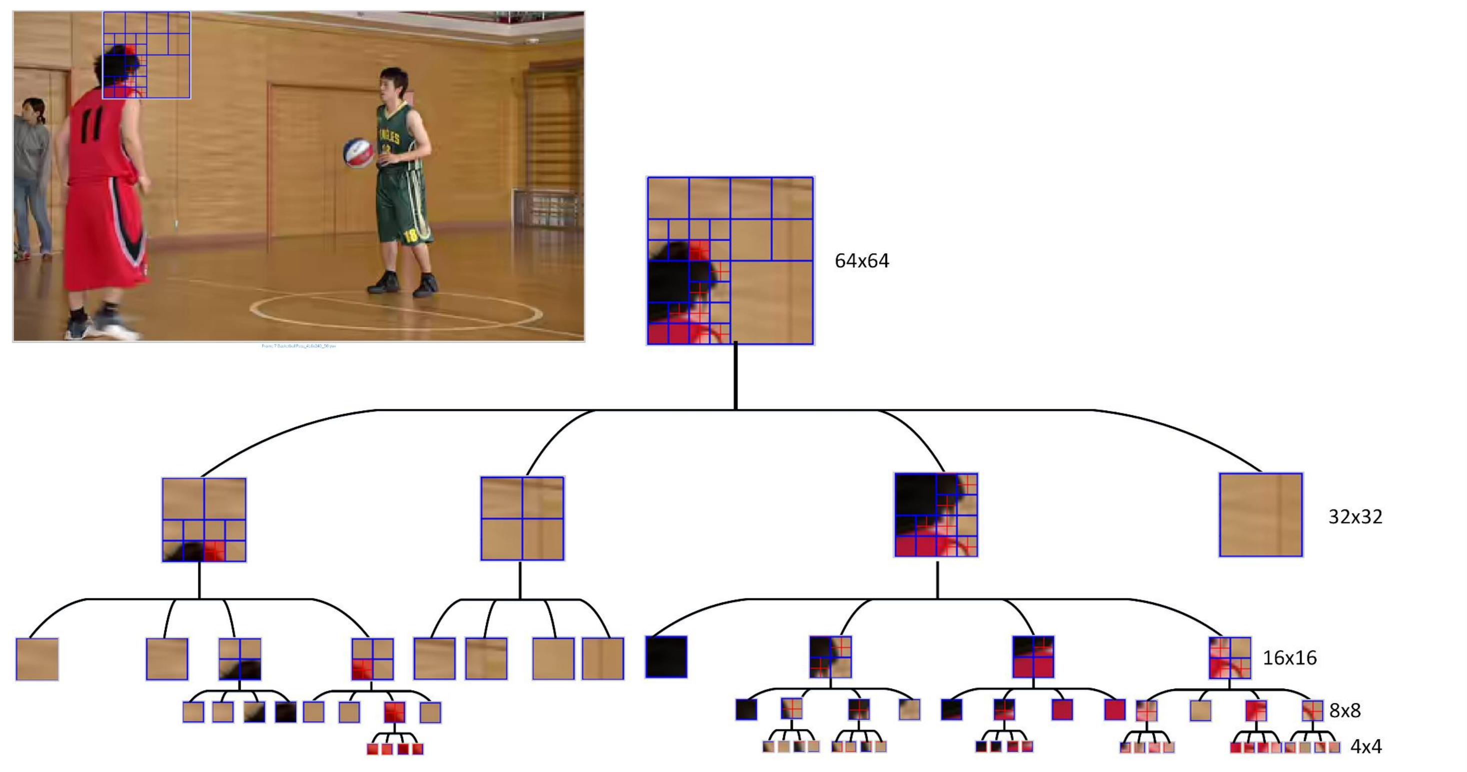
Yi Shan and En-hui Yang, Fellow, IEEE

University of Waterloo, ON, Canada (Email: {y9shan;ehyang}@uwaterloo.ca)

Introduction

High Efficient Video Coding (HEVC) is the latest video coding standard developed by the Joint Collaborative Team on Video Coding (JCT-VC). Compared with H.264, it improves the coding efficiency by 50% at the price of significant increase in encoding time, due to Rate Distortion Optimization (RDO) on large variations of block sizes and prediction modes. In this work, a new fast intra coding algorithm is proposed to alleviate the high computational complexity of HEVC intra-frame coding. The proposed algorithm is based on machine learning and Laplacian Transparent Composite Model (LPTCM). Features called Summation of Binarized Outlier Coefficient (SBOC) vectors are firstly extracted from original frames by using LPTCM and then fed into online trained Support Vector Machine (SVM). Two SVMs are combined to predict Coding Unit (CU) decisions so that the encoding process can be significantly sped up. Additionally, a performance controller is introduced to ensure the robustness of machine learning models. It is shown by experiments that, compared with HM 16.3, the proposed algorithm reduces the encoding time, on average, by 48% with negligible increase in BD-rate.

Block Partition in HEVC

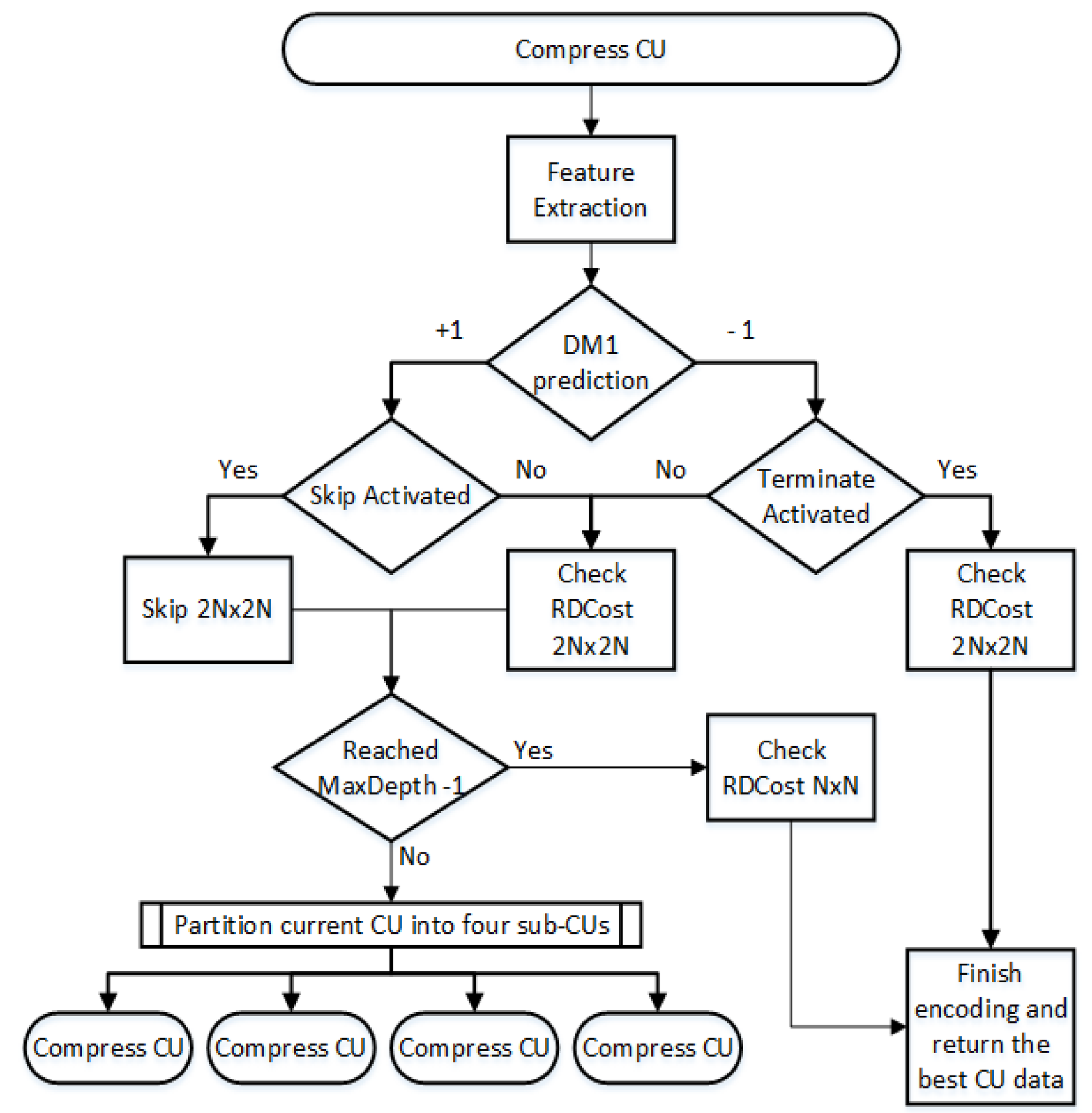


- HEVC captures the variation of image content by partitioning an image into different block sizes via a quad-tree structure. The maximum block size (Coding Tree Unit) is 64×64; the minimum block size (Smallest Transform Unit) is 4×4.
- In the standard method, the best Coding Unit (CU) partition mode is selected by exhaustively searching for the minimum Rate-Distortion (RD) cost:

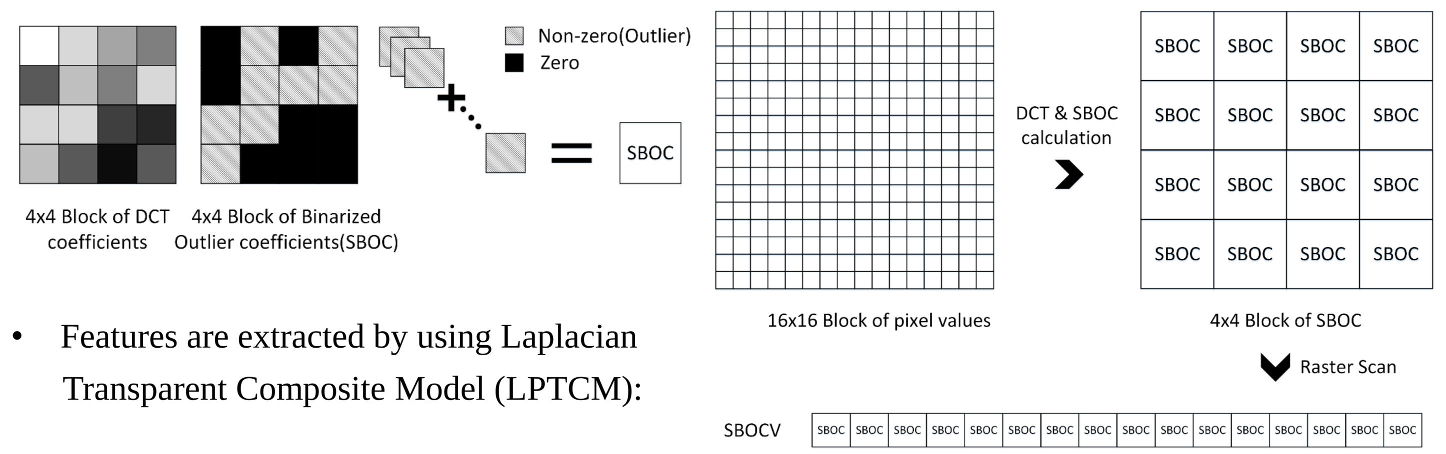
$$J_{RD} = SSE + \lambda R_{total}$$

- Computational complexity on CU size decision can be alleviated by making a prediction for one of two actions: CU Skip and CU Termination.
- CU Skip (+1): skip all the encoding processes on the current depth by jumping into the next depth of the quad-tree and directly splitting the CU into four sub-CUs.
- CU Termination (-1): terminate the CU splitting and keep the current CU size as the final choice.

Skeleton of the Proposed Algorithm



LPTCM & Feature Extraction

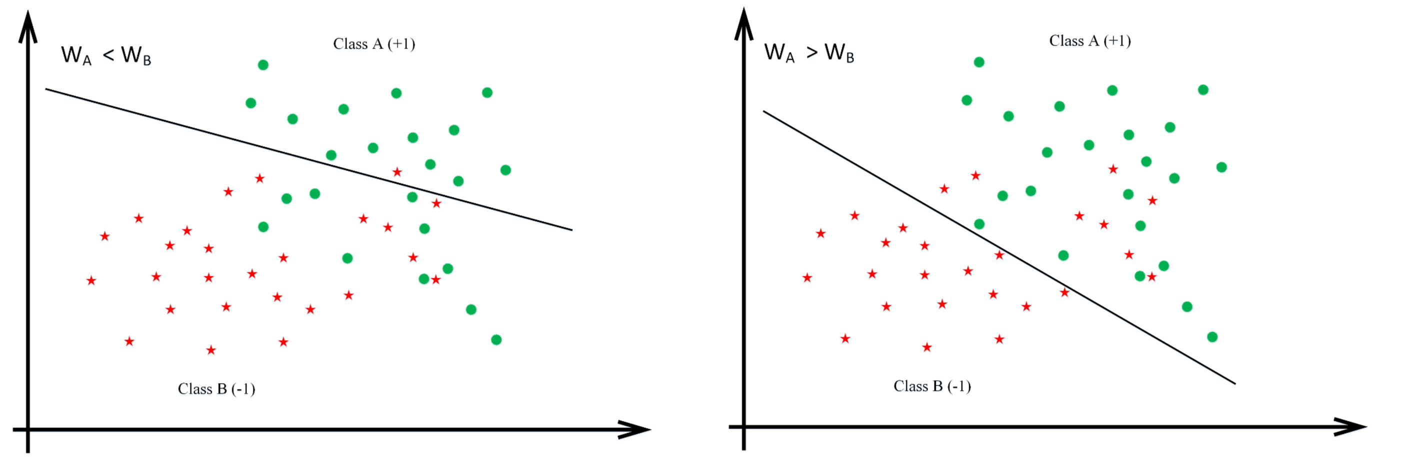


- Features are extracted by using Laplacian Transparent Composite Model (LPTCM):

$$p\,(y|y_c,b,\lambda)=\begin{cases}\frac{b}{2\lambda(1-e^{-y_c/\lambda})}e^{-|y|/\lambda} & \text{if } |y|<y_c\\ \frac{1-b}{2(a-y_c)} & \text{if } y_c<|y|\le a\\ \max\left\{\frac{b}{2\lambda(1-e^{-y_c/\lambda})}e^{-\frac{|y|}{\lambda}},\frac{1-b}{2(a-y_c)}\right\} & \text{if } |y|=y_c\\ 0 & \text{otherwise}\end{cases}$$

- Discrete Cosine Transform (DCT) coefficients that are smaller than y_c will be treated as inlier and suppressed to 0. DCT coefficients larger than y_c will be quantized to 1.
- A feature vector is formed by scanning Summation of Binarized Outlier Coefficient (SBOC) map.
- CU with size $N \times N$ will result in a feature vector with $N^2/16$ dimentions.

Weighted Support Vector Machine



- SVM classifier separates data points into different classes by using hyperplanes:

$$\mathbf{w}^T\mathbf{x} + b = 0.$$

- Given training set $S_t = \{(\boldsymbol{x}_1, y_1), (\boldsymbol{x}_2, y_2) \ldots, (\boldsymbol{x}_M, y_M)\}$, the hyperplane is trained by solving optimization problem:

$$\min_{\boldsymbol{w},b}\left\{\frac{1}{2}\boldsymbol{w}^T\boldsymbol{w} + C(W^+\sum_{i=1}^{|S^+|}\xi_i + W^-\sum_{j=1}^{|S^-|}\xi_j)\right\},$$

subject to

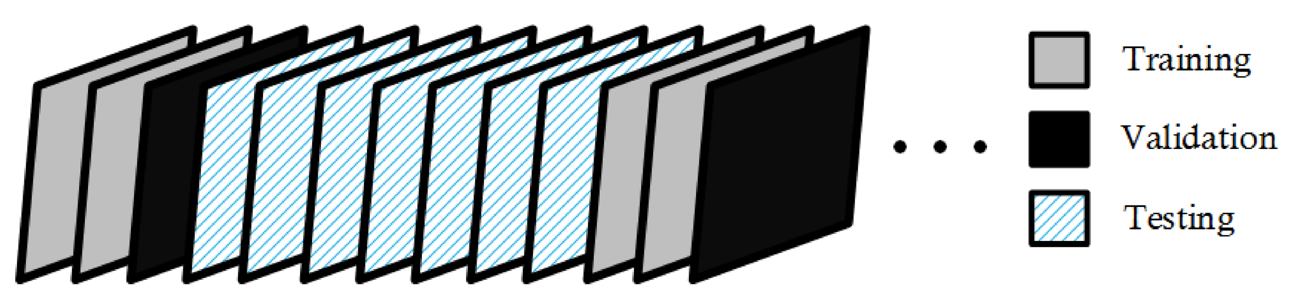
$$y_i(w^T x_i + b) \ge 1 - \xi_i \text{, and } \xi_i \ge 0 \quad \forall (x_i, y_i).$$

- To classify new data samples

$$y_{new} = sign(\boldsymbol{w}^T\boldsymbol{x}_{new} + b).$$

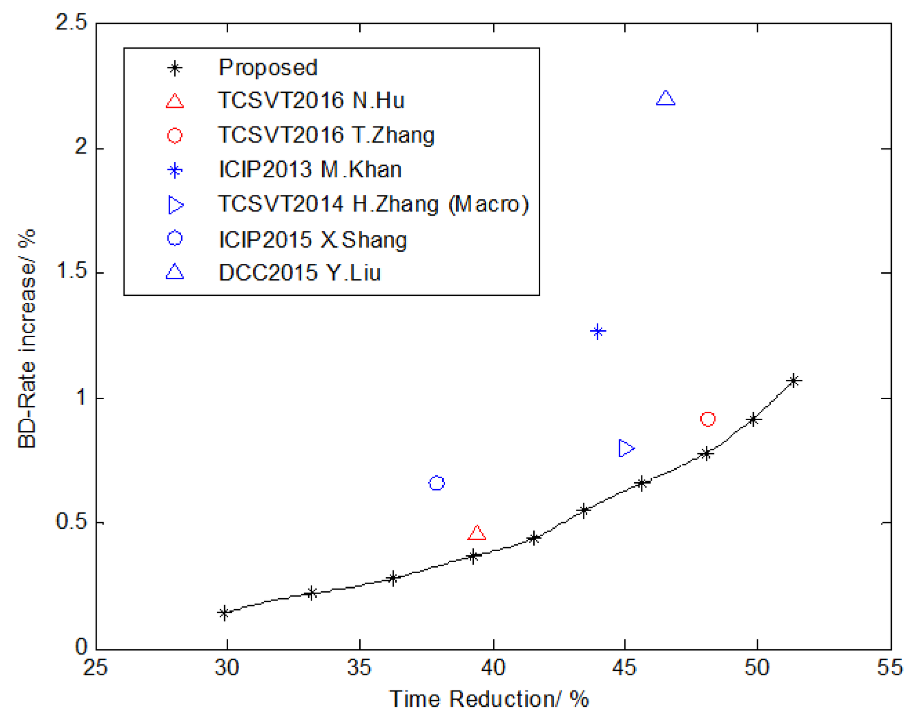
- $W^- > W^+$: larger penalty for false positive error, resulting in high precision on CU skip.
- $W^+ > W^-$: larger penalty for false negative error, resulting in high precision on CU termination.

Performance Control & Dynamic Training Period



- Every P frames as one period, which can be set to different value for each depth. Models will be retrained periodically to ensure the effectiveness of predictions.
- T frames for training: collect feature vectors and labels for all CU sizes; train SVM models at the end of the training stage.
- V frames for validation: validate the performance of prediction models; switch off the decision of some models to avoid significant loss on coding efficiency.
- $P-T-V$ frames for testing: apply prediction models to make decisions on CU partition and boost the encoding speed. Turn SVMs on or off based on their performances in validation stage.

Experiments & Results



- The full test included 24 test sequences of 6 classes and all sequences were encoded under All-Intra-Main configuration.
- Benchmark methods marked with red points are tested on the same machine. Blue points are data reported from the respective papers.
- The time reduction is calculated by comparing with the standard HM 16.3 encoder

$$Time\ Reduction = \frac{T_{HM} - T_{Fast}}{T_{HM}} \times 100\%\,.$$

- The main version of the proposed algorithm can achieve significant time reduction (overall 48.03%) with negligible increase in BD-rate (0.78%). It compares favorably with other fast algorithms for HEVC intra coding proposed recently in the literature.
- Shown the advantage of the proposed new feature vector, online training, performance control and model switching.

References

1. Yang, E. H., Yu, X., Meng, J., & Sun, C. (2014). Transparent composite model for dct coefficients: Design and analysis. *IEEE Transactions on Image Processing*, *23*(3), 1303-1316.
2. Hu, N., & Yang, E. H. (2015). Fast mode selection for HEVC intra-frame coding with entropy coding refinement based on a transparent composite model. *IEEE Transactions on Circuits and Systems for Video Technology*, *25*(9), 1521-1532.
3. Hu, N., & Yang, E. H. (2016). Erratum to “Fast Mode Selection for HEVC Intra-Frame Coding With Entropy Coding Refinement Based on a Transparent Composite Model”. *IEEE Transactions on Circuits and Systems for Video Technology*.
4. Zhang, T., Sun, M. T., Zhao, D., & Gao, W. (2016). Fast Intra Mode and CU Size Decision for HEVC. *IEEE Transactions on Circuits and Systems for Video Technology*.
5. Zhang, H., & Ma, Z. (2014). Fast intra mode decision for high efficiency video coding (HEVC). *IEEE Transactions on circuits and systems for video technology*, *24*(4), 660-668.
6. Shang, X., Wang, G., Fan, T., & Li, Y. (2015, September). Fast CU size decision and PU mode decision algorithm in HEVC intra coding. In *Image Processing (ICIP), 2015 IEEE International Conference on* (pp. 1593-1597). IEEE.
7. Duanmu, F., Ma, Z., & Wang, Y. (2015, September). Fast CU partition decision using machine learning for screen content compression. In *Image Processing (ICIP), 2015 IEEE International Conference on* (pp. 4972-4976). IEEE.
8. Khan, M. U. K., Shafique, M., & Henkel, J. (2013, September). An adaptive complexity reduction scheme with fast prediction unit decision for HEVC intra encoding. In Image Processing (ICIP), 2013 20th IEEE International Conference on (pp. 1578-1582). IEEE.
9. Chang, C. C., & Lin, C. J. (2011). LIBSVM: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology (TIST)*, *2*(3), 27.

Acknowledgement

This work is supported in part by the Natural Sciences and Engineering Research Council of Canada under Grant RGPIN203035-16, and by the Canada Research Chairs Program.