Multiple Linear Regression for High Efficiency Video Intra Coding

Zhaobin Zhang[*], Yue Li[#], Li Li[*], Zhu Li[*] and Shan Liu[+]

[*]*University of Missouri Kansas City*
[#]*University of Science and Technology of China*
[+]*Tencent America*

Introduction

- Background
  1. The prevalence of video/image capture devices generates more internet traffic which poses new challenge to compress and transmit these content.
  2. ISO/IEC MPEG and ITU-T worked together as the Joint Video Exploration Team (JVET) to explore state-of-the-art algorithms and prepare for the next-generation video standards.

- Limitations
  1. Existing intra prediction schemes of HEVC are not able to characterize the spatial variations due to limited prediction modes.
  2. Deep learning methods are better than existing HEVC solution, but often suffer from overwhelming computational complexity which is not tolerable in video codecs.

Introduction Cont'd

- Related work
  1. Y. Li [3] proposed to jointly predict current block by combining intra block copy and existing directional prediction scheme.
  2. C. Noel [5] developed a regression-based intra prediction scheme by iteratively refining predictors through regularized regression.
  3. Y. Li [7] proposed a CNN for HEVC intra coding following a down-sample up-sample pipeline.

- Proposed
  1. We propose a concise design based on Multiple Linear Regression (MLR).
  2. The proposed scheme, dubbed MIP, takes both the **reference pixels** and the best **intra prediction** as inputs, and derives the prediction block through a MLR regression model.
  3. To refine the model, the **intra prediction direction** is leveraged to fit separate regressors.
  4. An average of 0.4% BD-Rate reduction is achieved on HEVC common test sequences.

HEVC Intra Prediction

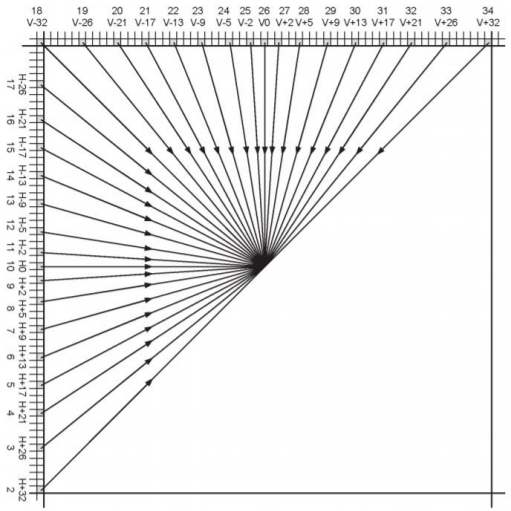


Fig. 1 HEVC intra prediction modes.

- Block-based intra prediction
  1. DC mode
  2. Planar mode
  3. 33 angular modes

- Motivated by [8] which incorporates piecewise linear projections, we attempt the potentials by combining interpolation and multiple linear regression.

Multiple Linear Regression

- Formulation

  The objective of MLR is to predict the outcome $\hat{Y}$ given the observations X and targets Y. A typical MLR model with k predictor variables:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \epsilon$$

  where ϵ is the error term, $\Theta = \{\beta_0, \beta_1, \ldots, \beta_k\}$ is the coefficient set to be solved for.

- Solution

  Least square is used to solve the MLR problem.

$$y = X\beta + \epsilon$$
$$\hat{\beta} = (X^T X)^{-1} X^T y$$

  Then the estimated value of y can be calculated as follows:

$$\hat{y} = X\hat{\beta}$$
$$\epsilon = y - \hat{y}$$

MLR for Intra Prediction

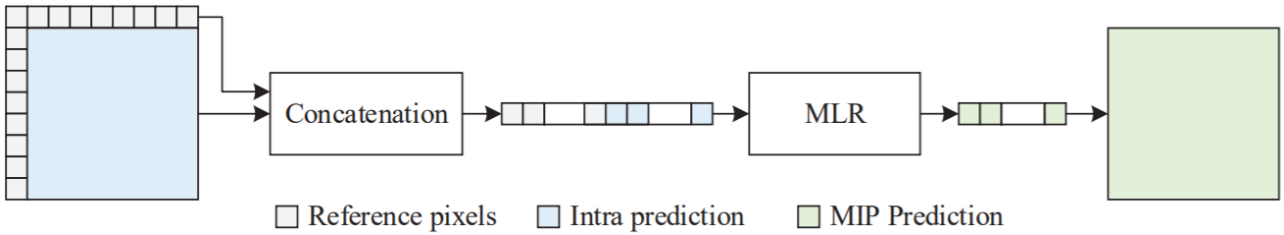


Fig. 2. Framework of MIP scheme. The reference pixels and the best intra prediction are utilized to derive the current block.

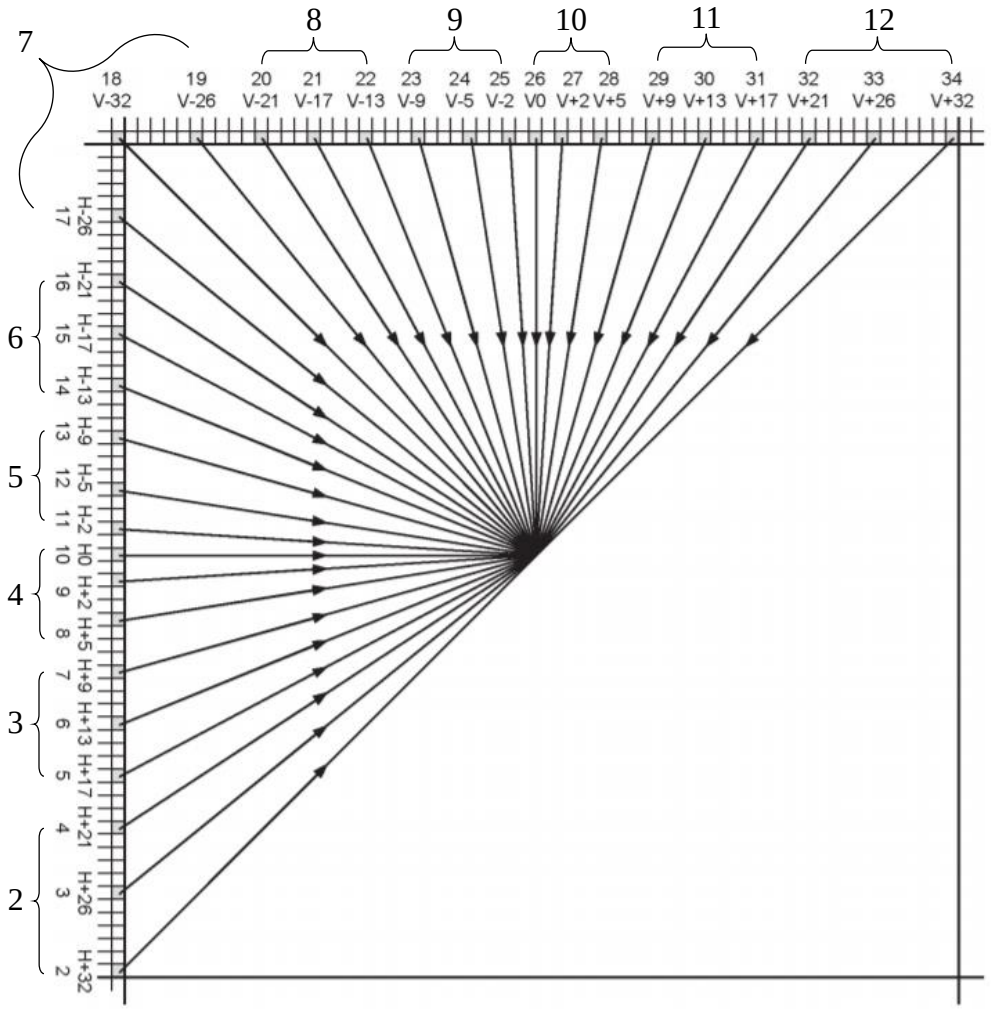
- Framework

  The best intra prediction is obtained through RDO.

  $2N + 1$ reference pixels

  Loss function $\mathcal{L} = ||Y - \hat{Y}||_2^2,\ \hat{Y} = XA + b$

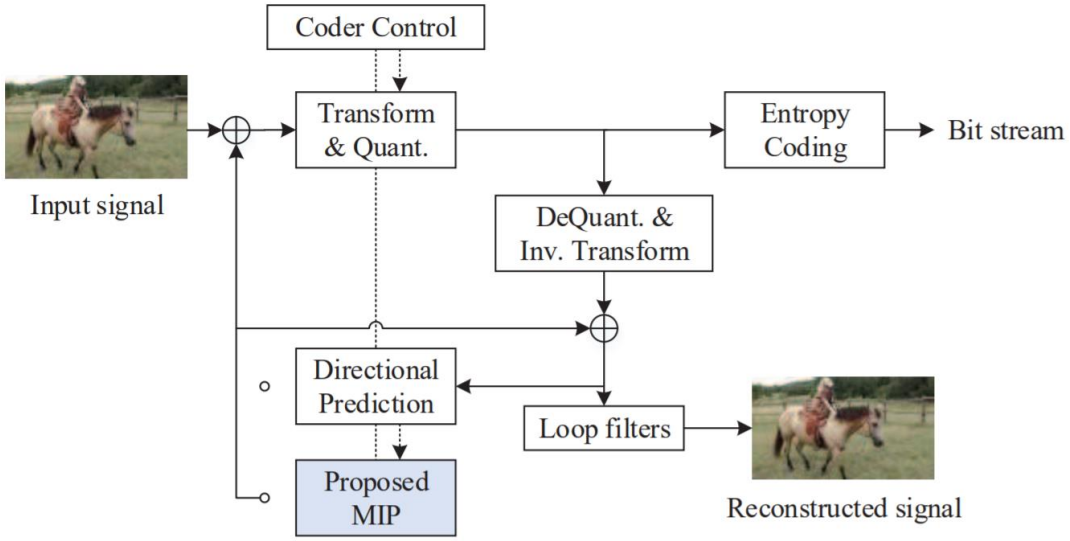
Refined Model



- Separate models w.r.t. prediction direction

$$m = \begin{cases} 0, & if\ n = 0 \\ 1, & if\ n = 1 \\ floor\left(\frac{n-2}{3}\right) + 2, & if\ n > 1 \end{cases}$$

Integration into HEVC



The proposed MIP is integrated into HEVC as an additional intra prediction mode. A flag is transmitted indicating whether it's adopted through RDO.

Experiment Settings

- Environments
  - Implemented in HEVC reference software 16.0.
  - $QP = \{22, 27, 32, 37\}$
  - All Intra (AI) configuration under HEVC CTC.
  - Only the 1st frame is used from each sequence.

- Training
  - Training dataset: cropped blocks from DIV2K.
  - Cropped blocks from 800 2k images for training, 100 2k images for validation and test, respectively.
  - A separate MIP is trained for each $\{block_{size}, QP, m\}$ combination, therefore $4 \times 4 \times 13 = 208$ models.

Experiment Results

Table 2. The BD-Rate results of the proposed method.

Sequence		BD-Rate		
		Y	U	V
Class A	Traffic	-0.9%	-0.6%	-1.3%
	PeopleOnStreet	-0.5%	0.0%	0.1%
	Nebuta	-0.9%	-0.8%	-0.7%
	SteamLocomotive	-0.6%	-0.3%	-0.3%
Class B	Kimono	-0.6%	-1.5%	-1.1%
	ParkScene	-0.6%	-1.1%	-1.0%
	Cactus	-0.7%	-0.3%	-1.1%
	BQTerrace	-0.4%	-1.3%	-0.6%
	BasketballDrive	-1.0%	0.2%	-1.0%
Class C	BasketballDrill	-0.1%	-2.0%	-0.8%
	BQMall	-0.2%	0.1%	-0.4%
	PartyScene	-0.1%	-0.2%	-0.7%
	RaceHorsesC	-0.4%	-0.2%	-0.7%
Class D	BasketballPass	0.1%	1.1%	-0.8%
	BQSquare	0.1%	0.1%	-1.1%
	BlowingBubbles	0.4%	-1.3%	-1.4%
	RaceHorses	-0.3%	-1.6%	0.6%
Class E	FourPeople	-0.4%	0.3%	-1.6%
	Johnny	-0.3%	-1.1%	-1.3%
	KristenAndSara	-0.5%	-0.4%	-0.6%
Class A		-0.7%	-0.5%	-0.5%
Class B		-0.6%	-0.8%	-1.0%
Class C		-0.2%	-0.6%	-0.8%
Class D		0.1%	-0.4%	-0.7%
Class E		-0.4%	-0.4%	-1.2%
Average		-0.4%	-0.6%	-0.8%
Enc Time		487%		
Dec Time		154%		

- An average of -0.4%, -0.6% and -0.8% BD-Rate saving is achieved for Y, U, V component , respectively.
- It performs better on high-resolution content, e.g., -0.9% BD-Rate reduction on *Traffic*.

Conclusions

1. This paper proposes a new method based on multiple linear regression for high-efficiency video intra coding.
2. The proposed scheme MIP, accepts both reference pixels and the best intra prediction and learns an end-to-end projection using a linear regressor.
3. The model is fitted in pixel domain which insures the simplicity.
4. To refine the model, separate model is trained w.r.t. the intra prediction direction.
5. The neat and concise architecture achieves promising gains against HEVC reference software 16.0.

Q & A