Dynamically Reconfigurable Architecture System for Time-varying Image Constraints (DRASTIC) for HEVC Intra Encoding

Yuebing Jiang
Department of Electrical and
Computer Engineering
University of New Mexico
Albuquerque, New Mexico 87111
Email: yuebing@unm.edu

Gangadharan Esakki
Department of Electrical and
Computer Engineering
University of New Mexico
Albuquerque, New Mexico 87111
Email: gesakki@ece.unm.edu

Marios Pattichis
Department of Electrical and
Computer Engineering
University of New Mexico
Albuquerque, New Mexico 87111
Email: pattichis@ece.unm.edu

***Abstract*—We introduce the use of a dynamically reconfigurable architecture system for time-varying image constraints (DRASTIC) and consider its application in HEVC intra encoding. DRASTIC provides a framework for jointly optimizing energy-rate-distortion for different operating modes. DRASTIC optimization involves dynamically reconfiguring parameters (e.g., the quantization parameter, encoding configuration modes) in either software or hardware implementation. However, in this paper, we will not consider the use of hardware cores. DRASTIC implementations for HEVC intra-encoding allows optimized performance in a set of control modes we have provided. We provide a list of Pareto-optimal configurations to provide performance scalability and demonstrate performance on time-varying constraints with DRASTIC control.**

***Index Terms*—HEVC, intra, DRASTIC, complexity, rate, distortion, optimization**

I. INTRODUCTION

HEVC is improved video coding standard with tools such as recursive coding/transforms units, complex intra prediction modes, and asymetric inter prediction unit division, etc. It aims at 50% bit rate reduction for equal perceptual video quality [1]. However, the performance comes with unbearable computing complexity as more computing and comparision are executed. To reduce the inter encoding complexity, several configuration modes were shown as in [2]. For intra encoding complexity, rough mode set(RMS,[3]), gradient based intra prediction [4] and coding unit(CU) depth control[5] were proposed. Problem comes when complexity is treated as a separate performance factor, you can't achieve scalablility in complexity with same RD performance.

Dynamically Reconfigurable Architecture System for Time-varying Image Constraints (DRASTIC) is a multi-objective optimization framework for video compression considering opposing compression performance. Note E as computing energy, Q as image quality, R as compression bitrate and $\mathcal{C}$ is the control space. There are 4 modes in DRASTIC:

- Minimum Complexity: Used when energy left is limited and still have rate and quality considerations.

$$\min_{\mathcal{C}} E,\ \ subject\ to\ Q \geq Q_{min}\ \&\ R \leq R_{max}$$

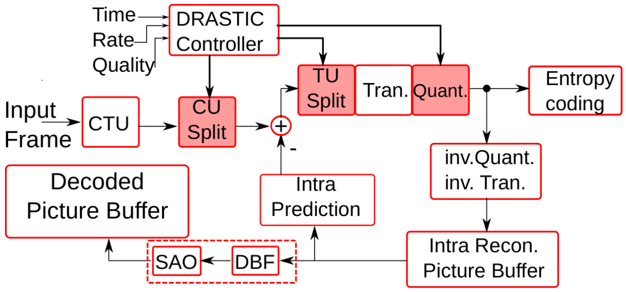


Fig. 1. Diagram for DRASTIC HEVC Intra Encoding System

- Minimum Rate: Used when connection bandwidth is limited.

$$\min_{\mathcal{C}} B,\ \ subject\ to\ Q \geq Q_{min}\ \&\ E \leq E_{max}$$

- Maximum Quality: Used when connection and device energy can be granted at certain level.

$$\max_{\mathcal{C}} Q,\ \ subject\ to\ P \leq P_{max}\ \&\ R \leq R_{max}$$

- Typical: Used when device needs more precise resource allocation strategy.

$$\min_{\mathcal{C}} -\alpha \cdot Q + \beta \cdot E + \gamma \cdot R,$$

$$subject\ to\ Q \geq Q_{max}\ \&\ E \leq E_{min}\ \&\ R \leq R_{min}$$

Our constributations are as follows: 1) build a control space where scalable encoding complexity, bitrate and quality can be provided using HEVC intra encoding. 2) DRASTIC modes are implemented on our performance model on frame level. The paper is organized as follows: In section II, we talked about the control space we have created and shows how the performance will respond to control. DRASTIC control model and parameter is discussed in section III. Conclusion is made in section IV.

II. RECONFIGURABLE PARAMETERS

As derivation of DRASTIC DCT implemented for JPEG system mentioned in [6], our DRASTIC HEVC Intra encoding system is shown in fig.1. For the reference software HM11.0 implementation[7] on intra encoding, the encoding process is as follows: For Luma prediction modes, a rough mode set (RMS[8], [3], 8 modes for 4x4 and 8x8 CU, 3 mode for other CUs) is selected based on their simplified RD performance. At this step, sum of absolute Hadmard transform coefficients is used as **distortion** and mode bit is used as **Rate**. Based on RMS, the best Luma prediction mode is selected in RMS according to a more complex RD performance. **Distortion** is calculated as sum of square reconstruction difference. **Rate** is the bits used when largest TU is deployed. For chroma prediction modes, RMS is not needed because of less pixels and modes to compute. The best chroma prediction mode is selected from 5 available modes according to their RD performance same as second step for luma mode. Given the determined luma and chroma prediction mode. The transform tree and trasform coefficients are determined using an exhaustive subdivision process, where coding tree unit is constructed based on spliting with the best RD performance. The reconstructed pixel values were saved in reconstructed picture buffer, further filter operation such as deblocking filter(DBF) and sample adaptive offset (SAO) is applied before the intra picture is put into decoded picture buffer. A DRASTIC controller is added to realize DRASTIC modes. It takes in the compression performances as feed back from each frame's compression results and control the comprssion of next frame. Note in this paper, we disabled deblocking filter and sample adaptive offset, other configuration were the same as standard intra main profile configuration. This DRASTIC implementation onto HEVC Intra encoding application has everything in software.

To provide Time, Rate and Quality scalability, we use two control parameters, (1) QP (0-51), we know QP plays an import role on RD control, larger QP will lead to smaller bitrate and higher distortion and vice versa, also we know larger QP will lead to less residual coefficients to encode, thus less encoding time. (2) We have defined a customized parameter $config = i,\ i \in [0, 19]$, this parameter is used to control CU and TU sizes. Larger $config$ will lead to more control depth, thus better RD performance and more encoding time. We have $config = 5 * DepthConfig + FinerConfig$. The meaning of $DepthConfig$ and $FinerConfig$ is shown in tbl.I and II.

TABLE I
DEPTH CONTROL FOR THE CU, TU CONTROL USING DEPTHCONFIG.

DepthConfig	Allowed Luma CU size	Allowed Luma TU size
0	64,32	32
1	64,32,16	32,16
2	64,32,16,8	32,16,8
3	64,32,16,8,4	32,16,8,4

TABLE II
FINER DEPTH CONTROL FOR THE CU, TU SIZES USING FINERCONFIG. (NOTE FOR DEPTHCONFIG=0, WE ONLY CONTROL CU SIZES USING FINERCONFIG)

finer_depth_config	Allowed Luma CU size	Allowed Luma TU size
0	20% minimum size	20% minium size
1	40% minimum size	40% minimum size
2	60% minimum size	60% minimum size
3	80% minimum size	80% minimum size
4	100% minimum size	100% minimum size

We have shown the performance space on Time(seconds per sample), Rate(bits per sample) and Quality (PSNR) with RaceHorses 432x240 video in fig.2, 3 and 4. Note we use seconds per sample (SPS) for time measurement and bits per sample (BPS) for rate measurement. The control space have QP in range [0,51) with step of 3 and config in range [0,19] with step of 1. For each control combination, we encode 6 frame from the video and save the averaged results. From the figure, we can conclude that higher configuration will lead better RD performance but higher complexity, higer QP will lead smaller Rate, higher Distortion and lower complexity. We also have tested that all 340 generated points are pareto points, which means we have created a pareto-optimal configurations.

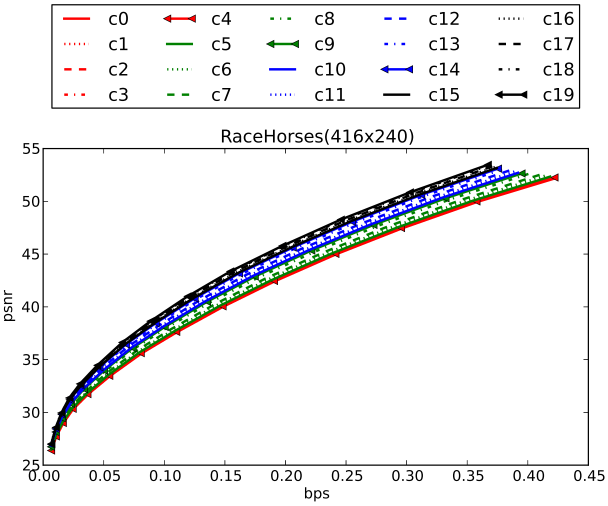


Fig. 2. Rate,Distortion performance with different QP and config control on RaceHorses 432x240. BPS stands for bits-per-sample.

III. DRASTIC CONTROL

A. Initialization Process

To realize DRASTIC control, we define 3 profile (low, median and high) to simulate real-time constraints. Each profile comes with a control pair ($Config_{init}$,QP_{init}). We set (42,0) for low profile, (28,7) for medium profile and (14,14) for high profile. The constraints are selected on the average performance of 3 neighbored control pair around initialization control, they are ($Config_{init}$+2,QP_{init}),($Config_{init}$,$QP_{init} - 2$)

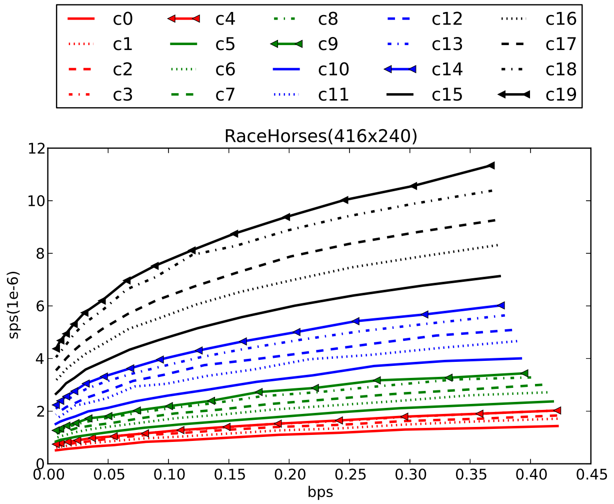


Fig. 3. Rate,Complexity performance with different QP and config control on RaceHorses 432x240. SPS stands for seconds-per-sample.

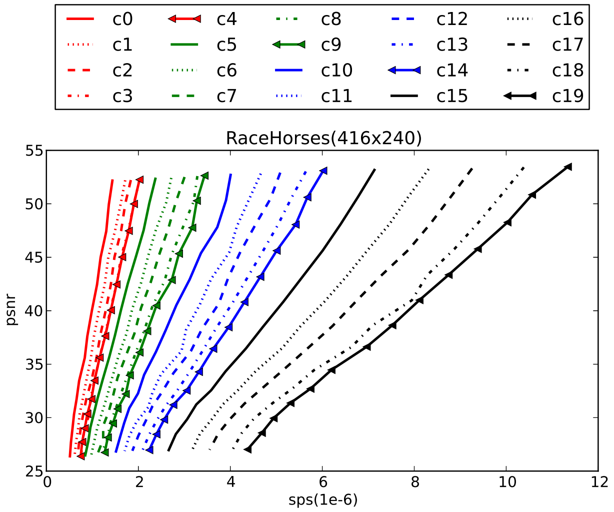


Fig. 4. Complexity,Distortion performance with different QP and config control on RaceHorses 432x240

and $(Config_{init}+2, QP_{init}-4)$. The initialization process is shown in fig.5(a). Note the averaged performance of 3 neighbored control as $PSNR_{init}$,$Rate_{init}$ and $Time_{init}$, they can be used as constraints according to DRASTIC modes.

B. Initialization and Hold Control

A simple solution for controling the encoding system to meet initialized constraints is just to initialize and hold. We change the profile constraints every 40 frames. The initialization and hold policy is shown in fig.5, this method will create a smooth performance sequence, we know that the performance will not change too much as long as there is no scene change. Problem with initialization and hold control policy is that the control space is not used to achieive optimized performance, also it can't adapt to image contents, scene change will lead to performance ripple.

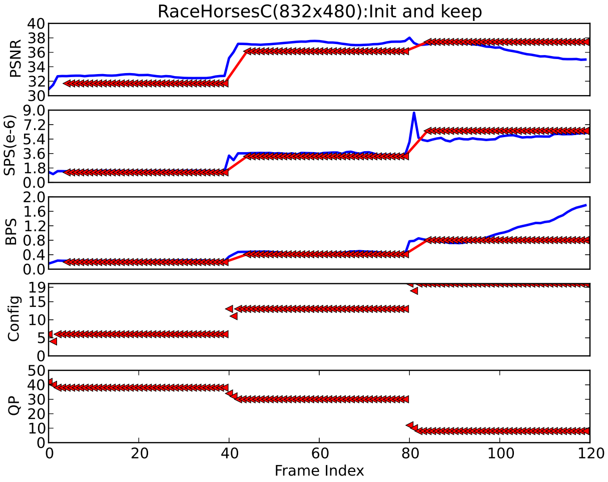


Fig. 5. Initialization and hold control results

C. Prediction Model and Model Update

We propose a set of simple performance model in eq.1 to predict frame level performance on PSNR, Time and Rate. We need 3 control and performance samples to calculate the encoding system coefficients, and the system coefficients will be used to predict performance with control variation. The model analysis process is shown in eq.2,3and4.

$$\begin{aligned} PSNR &= a1\cdot QP + b1\cdot Config + c1 \\ Time &= a2\cdot QP + b2\cdot Config + c2 \\ Rate &= a3/(2^{QP-4/6}) + b3\cdot Config + c3 \end{aligned}, A = \begin{bmatrix} a1 & b1 & c1 \\ a2 & b2 & c2 \\ a3 & b3 & c3 \end{bmatrix} \tag{1}$$

$$\begin{bmatrix} a1 \\ b1 \\ c1 \end{bmatrix} = \begin{bmatrix} QP_1 & Config_1 & 1 \\ QP_2 & Config_2 & 1 \\ QP_3 & Config_3 & 1 \end{bmatrix}^{-1} \cdot \begin{bmatrix} PSNR_1 \\ PSNR_2 \\ PSNR_3 \end{bmatrix} = C^{-1}\cdot\overrightarrow{PSNR} \tag{2}$$

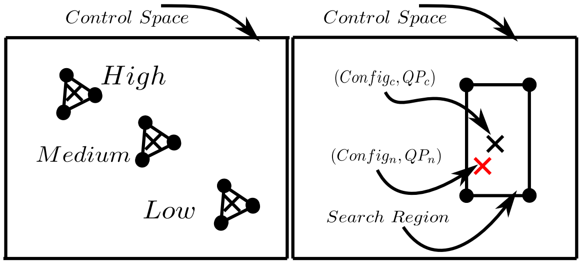
$$\begin{bmatrix} a2 \\ b2 \\ c2 \end{bmatrix} = \begin{bmatrix} QP_1 & Config_1 & 1 \\ QP_2 & Config_2 & 1 \\ QP_3 & Config_3 & 1 \end{bmatrix}^{-1} \cdot \begin{bmatrix} Time_1 \\ Time_2 \\ Time_3 \end{bmatrix} = C^{-1}\cdot\overrightarrow{Time} \tag{3}$$

$$\begin{bmatrix} a3 \\ b3 \\ c3 \end{bmatrix} = \begin{bmatrix} 2^{(QP_1-4)/6} & Config_1 & 1 \\ 2^{(QP_2-4)/6} & Config_2 & 1 \\ 2^{(QP_3-4)/6} & Config_3 & 1 \end{bmatrix}^{-1} \cdot \begin{bmatrix} Rate_1 \\ Rate_2 \\ Rate_3 \end{bmatrix} = C^{-1}\cdot\overrightarrow{Rate} \tag{4}$$

Note also the inverse of control matrix C don't always exist, we use a simple fix that will detect dependent rows in C and update the latest control pair only. If C^{-1} still cannot be obtained and the last model is not precise since it misses estimated Time-Rate-PSNR being by more than 5% , we will perturb the control by adding random offsets from $[-1,+1]$ to the control parameters.

D. Search for Optimal Control

Once the system performance model is updated or labeled as precise, we will calcuate next performance vector $[PSNR_n, Time_n, Rate_n]$ using next control pair $(Config_n, QP_n)$. With current control pair as $(Config_c, QP_c)$, seach region for $(Config_n, QP_n)$ is defined as $[Config_c - 2, Config_c + 2] \times [QP_c - 5, QP_c + 5]$. According to different DRASTIC control mode, different objective will be compared. Optimal control pair will be used as encoding control for next frame. If the system model is not precise and can't be constructed, we will make encoding control for next frame a perturbed version of current encoding control as stated above.



(a) (b)

Fig. 6. (a) Profile initialization process (b) Control pair search and comparison process

IV. DRASTIC Implementation And Results

A. Minimum Complexity Mode

Minimum Complexity mode will use $PSNR_{init}$ and $Rate_{init}$ as constraints Q_{min} and R_{max}. In the process of searching for the optimal control, (1) If there exist $PSNR_n >= Q_{min}$ and $Rate_n <= Q_{max}$. We will select the optimal control as the pair with minimum $Time_n$.(2) If No estimation performance satisfies above constraints. We will build an objective function as eq.5 and select optimal control with the minimum objective. We suggest the coefficients as $\alpha = 1$, $\beta = 20$ and $\gamma = 20$. An example controlled compression sequence is shown in fig.7. Compare to initialization and hold policy, we can see that the encoding time is properly depressed, especially at high profile.

$$\begin{aligned}&\alpha(Time_n/Time_{init}) + \beta(\text{abs}(PSNR_n - Q_{min})/Q_{min})\\&+\gamma(\text{abs}(Rate_n - R_{max})/R_{max})\end{aligned} \tag{5}$$

B. Minimum Rate Mode

Minimum Rate mode will use $PSNR_{init}$ and $Time_{init}$ as constraints Q_{min} and E_{max}. In the process of searching for the optimal control, (1) If there exist $PSNR_n >= Q_{min}$ and $Time_n <= E_{max}$. We will select the optimal control as the pair with minimum $Rate_n$.(2) If No estimation performance satisfies above constraints. We will build an objective function as eq.6 and select optimal control with the minimum objective. We suggest the coefficients as $\alpha = 20$, $\beta = 20$ and $\gamma = 1$. An example controlled compression sequence is shown in fig.8. Compare to initialization and hold policy, minimum rate tends to have smaller bitrate while satisfying quality and SPS constraints.

$$\begin{aligned}&\alpha(\text{abs}(Time_n - E_{max})/E_{max})\\&+\beta(\text{abs}(PSNR_n - Q_{min})/Q_{min}) + \gamma(Rate_n/R_{max})\end{aligned} \tag{6}$$

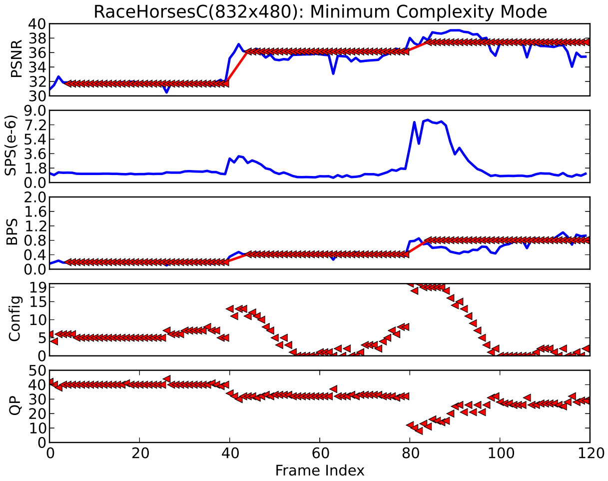


Fig. 7. minimum complexity mode with 3 profiles

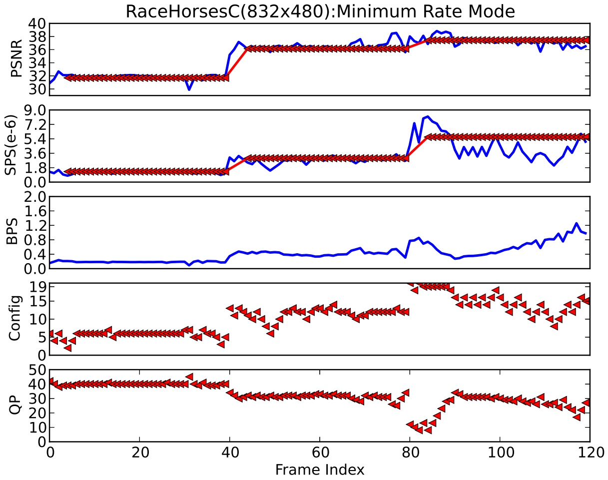


Fig. 8. minimum rate mode with 3 profiles

C. Maximum Quality Mode

Maximum Quality mode will use $Rate_{init}$ and $Time_{init}$ as constraints R_{max} and E_{max}. In the process of searching for the optimal control, (1) If there exist $Rate_n <= R_{max}$ and $Time_n <= E_{max}$. We will select the optimal control as the pair with maximum $Quality_n$.(2) If No estimation performance satisfies above constraints. We will build an objective function as eq.7 and select optimal control with the minimum objective. We suggest the coefficients as $\alpha = 20$, $\beta = -1$ and $\gamma = 20$. An example controlled compression seqeuence is shown in fig.9. Compare to initialization and hold policy, we can see that the compression quality strive to maintain a high level.

$$\begin{aligned}&\alpha(\mathrm{abs}(Time_n - E_{max})/E_{max})\\&+\beta(PSNR_n/Q_{init}) + \gamma(\mathrm{abs}(Rate_n - R_{max})/R_{max})\end{aligned} \tag{7}$$

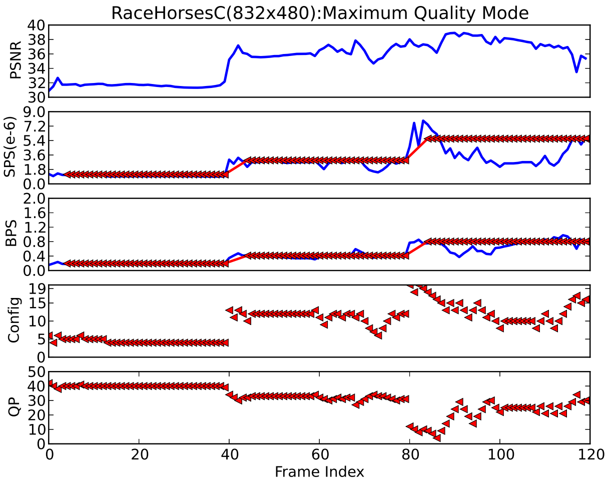


Fig. 9. maximum quality mode with 3 profiles

D. Typical Mode

Typical mode will use $PSNR_{init}$, $Rate_{init}$ and $Time_{init}$ as constraints Q_{min}, R_{max} and E_{max}. In the process of searching for the optimal control, (1) If there exist $PSNR_n >= Q_{min}$,$Rate_n <= Q_{max}$ and $Time_n <= E_{max}$. We will select the optimal control as the pair with minimum objective in eq.8. The coefficients are suggested as $\alpha = 1$,$\beta = -1$,$\gamma = 1$. (2) If No estimation performance satisfies above constraints. We will build an objective function as eq.9 and select optimal control with the minimum objective. We suggest the coefficients as $\alpha = 20$, $\beta = 20$, and $\gamma = 20$. Compare to initialization and hold policy, typical mode provide a more consistant performance for each profile.

$$\alpha(Time_n/E_{max}) + \beta(PSNR_n/Q_{min}) + \gamma(Rate_n/R_{max}) \tag{8}$$

$$\begin{aligned}&\alpha(abs(Time_n - E_{max})/E_{max})\\&+\beta(abs(PSNR_n - Q_{min})/Q_{min})\\&+\gamma(abs(Rate_n - R_{max})/R_{max})\end{aligned} \tag{9}$$

V. CONCLUSION

In this paper, we presented a DRASTIC implementation for HEVC intra encoding system. The system can be used to reach different multi-objective optimization in video communication applications. A set of performance model are proposed to estimate encoding performance, and the implementation results has shown reasonable optimization is reached.

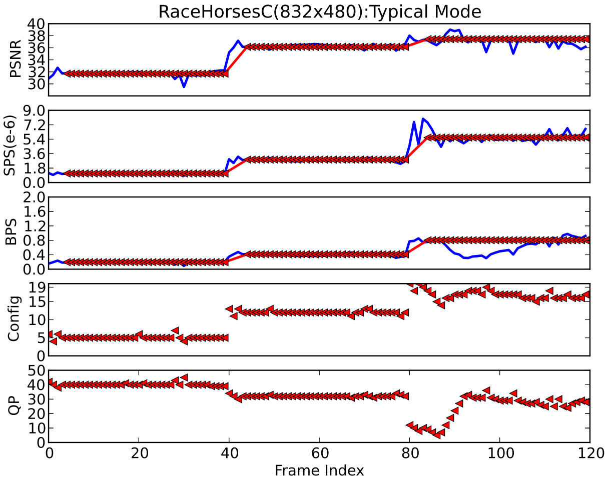


Fig. 10. typical mode with 3 profiles

REFERENCES

[1] G. Sullivan, J. Ohm, W.-J. 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.

[2] F. Bossen, B. Bross, K. Suhring, and D. Flynn, "Hevc complexity and implementation analysis," *IEEE Transactions on Circuit and Systems for Video Technology*, vol. 22, no. 12, pp. 1685–1696, 2012.

[3] L. Zhao, L. Zhang, S. Ma, and D. Zhao, "Fast mode decision algorithm for intra prediction in hevc," in *2011 IEEE Visual Communications and Image Processing (VCIP),*, nov. 2011, pp. 1 –4.

[4] W. Jiang, H. Ma, and Y. Chen, "Gradient based fast mode decision algorithm for intra prediction in hevc," in *2012 2nd International Conference on Consumer Electronics, Communications and Networks (CECNet),*, april 2012, pp. 1836 –1840.

[5] G. Correa, P. Assuncao, L. Agostini, and L. da Silva Cruz, "Complexity control of high efficiency video encoders for power-constrained devices," *IEEE Transactions on Consumer Electronics,*, vol. 57, no. 4, pp. 1866–1874, November.

[6] Y. Jiang and M. Pattichis, "Dynamically reconfigurable dct architectures based on bitrate, power, and image quality considerations," in *2012 19th IEEE International Conference on Image Processing (ICIP),*, 2012, pp. 2465–2468.

[7] I.-K. Kim, K. McCann, K. Sugimoto, B. Bross, and W.-J. Han, "High efficiency video coding (hevc) test model 11 (hm11) encoder description," 2013.

[8] J. C. Yinji Piao, Junghye Min, "Encoder improvement of unified intra prediction, jctvc-c207," 2010.