OPTIMAL DEPENDENT BIT ALLOCATION FOR AVS INTRA-FRAME CODING VIA SUCCESSIVE CONVEX APPROXIMATION

Chao Pang, Oscar C. Au, Feng Zou, Xingyu Zhang, Wei Hu and Pengfei Wan

Department of Electronic and Computer Engineering
The Hong Kong University of Science and Technology
Clear Water Bay, Kowloon, Hong Kong
Email:{pcece, eeau, fengzou, xyzhang, huwei, leoman}@ust.hk

ABSTRACT

We consider the optimal dependent bit allocation strategy for AVS intra-frame coding. Due to the block-based predictive coding, the rate-distortion (R-D) characteristics of neighboring blocks are dependent with each other. However, the interblock coding dependency is neglected in most of the existing bit allocation methods. Different from the conventional methods, the proposed method fully exploit the interblock coding dependency and carefully leverage it in the problem formulation. Then successive convex optimization techniques are employed to convert the original nonconvex optimization problem into a series of convex optimization problems which can be solved efficiently and optimally. Experimental results have proved the superiority of the proposed method in terms of significant R-D performance improvement.

Index Terms— AVS intra-frame coding, dependent bit allocation, successive convex approximation

1. INTRODUCTION

AVS video coding standard [1] is working group of audio and video coding standard in China, which established in 2002. Targeting at various video applications, AVS-video coding standards define different profiles, combining advanced video coding tools with trade-off between coding efficiency and encoder/decoder implementation complexity as well as functional properties.

In AVS intra-frames, the frames are divided into non-overlapped blocks and coded block by block. Predictive coding techniques are employed to encode each block. The prediction signal is derived from the previously reconstructed neighboring pixels. Because of the transform characteristics and equal quantization parameters for all the frequency positions, the quantization distortion is not evenly distributed inside the frame. Typically the distortion of the block boundary pixels is much larger than that of the interior pixels. A typical distortion distribution result is shown in Fig. 1. From the figure, we can see that the average PSNR value of the boundary pixels is about 1dB lower than that of the interior pixels. This biased distortion distribution has two disadvantages: one is the blocking artifact which leads to unpleasant visual experience; the other one is potential degradation of coding efficiency, since the larger distortion of the boundary pixels will propagate to the following blocks through intra-prediction. In [2], deblocking filter is proposed to recover the boundary pixels and further ease the blocking artifact. However, the deblocking filter is performed after all the blocks are encoded. Thus, the error propagation is still inevitable and the intra-prediction results are not changed. To improve the intra-prediction performance, more sophisticated intra-prediction techniques are proposed in [3], where two intra-prediction modes are combined to get the final prediction value. Although the prediction is improved, the distortion of the neighboring pixels used for intra-prediction is still relatively large. In [4], an equal distortion distribution strategy is proposed such that the distortion inside the frame is evenly distributed. However, to maximize the coding efficiency, the boundary pixels are expected to have relatively smaller distortion, since they will be used as the reference pixels for the prediction of the following blocks. In our previous work [5], an optimal

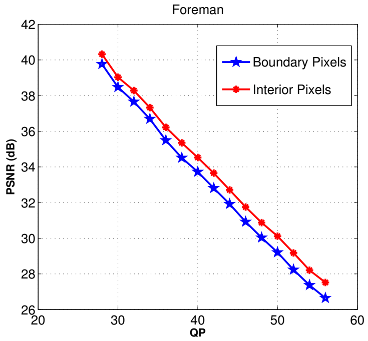


Fig. 1. Biased distortion distribution in AVS intra-frames

This work has been supported in part by the Research Grants Council (GRF Project no. 610112) and HKUST (HKUST Project no. FS-GRF12EG01) of the Hong Kong Special Administrative Region, China.

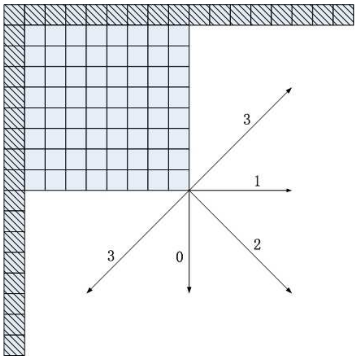


Fig. 2. Intra-prediction in AVS

distortion redistribution method is proposed. However, that work is used for H.264 intra-frame coding. In this paper, an optimal dependent bit allocation method is proposed, which is dedicated to AVS intra-frame coding. In our proposed method, the coding dependency and its impact to the final coding efficiency are fully investigated.

This paper is organized as follows: in Section II, the coding dependency of neighboring blocks in AVS intra-frame coding is investigated; the proposed optimal dependent bit allocation strategy is introduced in Section III; the experimental results are shown Section IV and Section V concludes the paper.

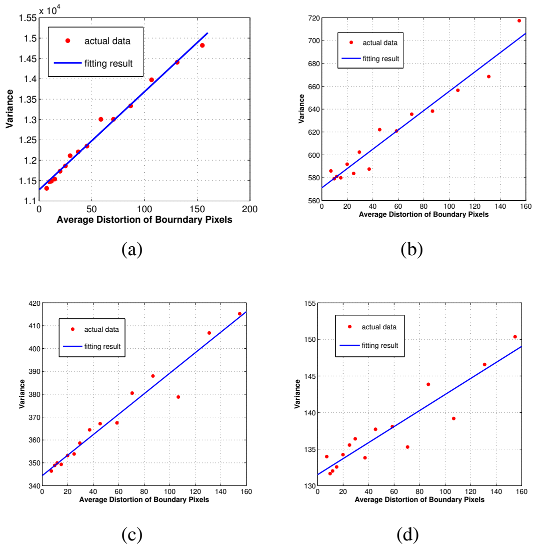
2. INTERBLOCK CODING DEPENDENCY IN AVS INTRA-FRAME CODING

Block-based predictive coding is employed in AVS intra-frame coding. For each coding block, the neighboring reconstructed pixels are used for intra-prediction. To accommodate various video contents and further enhance the prediction performance, 5 different prediction modes are defined for each 8*8 block in AVS Jizhun profile, including a DC prediction mode and 4 directional prediction modes. As shown in Fig. 2. For each prediction mode, the neighboring reconstructed pixels will be copied along the corresponding direction to form the prediction result.

After the intra-prediction, the residue block can be calculated by subtracting the prediction block from the original block in pixel domain. Then, the DCT transform is performed to obtain the DCT coefficient block. Let us assume the DCT coefficients at different frequency positions are independent and follow Gaussian distribution. Based on the classic R-D theory, the total bitrate can be estimated as

$$R = \sum_{i=0}^{7}\sum_{j=0}^{7} \log \frac{\sigma_{i,j}^2}{D_{i,j}} \qquad \sigma_{i,j}^2 > D_{i,j} \tag{1}$$

where $\sigma_{i,j}^2$ and $D_{i,j}$ are the variance and the distortion of the DCT coefficient at frequency position (i, j), respectively.



(a) (b) (c) (d)

Fig. 3. Fitting results of (2) at different frequency positions: (a) (0,0); (b) (0,2); (c) (3,0); (d) (2,2).

From (1), we can learn that the R-D function can be fully depicted if $\sigma_{i,j}^2$ are estimated.

In this paper, an empirical approach is employed to estimate $\sigma_{i,j}^2$. From extensive experiments, we found that $\sigma_{i,j}^2$ can be calculated as

$$\sigma_{i,j}^2 = \alpha_{i,j}\bar{d} + \beta_{i,j} \tag{2}$$

where $\alpha_{i,j}$ and $\beta_{i,j}$ are frequency position dependent parameters. $\bar{d}$ is the average distortion of the neighboring reconstructed pixels. Some typical experimental results are shown in Fig. 3. For the DCT coefficients at each frequency position, $\sigma_{i,j}^2$ increases monotonically with $\bar{d}$, which is consistent with our expectation. If $\bar{d}$ is large, the neighboring pixels used for intra-prediction tend to have large quantization distortion. This will lead to worse prediction and more residue. For $\beta_{i,j}$, it can be viewed as a measure of the intrinsic difference, the so-called innovation signal, between the original neighboring pixels and current block. In other words, even the neighboring pixels are free of quantization distortion, the residue signals are often not zero. Typically the value of the innovation signal depends on the characteristics of the image content itself. It should be noted that the DCT coefficients at different frequency positions correspond to different values of $\alpha_{i,j}$ and $\beta_{i,j}$. (2) is an interesting finding, as it provides us a quantitative, instead of conventionally conceptual, measure of the coding dependency inhabits in the intra-prediction based image coding.

3. OPTIMAL DEPENDENT BIT ALLOCATION FOR AVS INTRA-FRAME CODING

In this section, the proposed optimal dependent bit allocation strategy for AVS intra-frame coding will be introduced. A typical rate constrained distortion minimization problem can be mathematically presented as

$$\begin{aligned} \min_{R_i} \quad & \sum_{i=1}^{n} D_i \\ \text{s.t.} \quad & \sum_{i=1}^{n} R_i \leq R \end{aligned} \tag{3}$$

where (R_i,D_i) is the R-D pair of each coding unit, and R is the total rate constraint. In the proposed method, the coding unit is selected to be the DCT coefficients at each frequency position. In practical implementations, (R_i,D_i) can be adjusted by tuning the quantization parameter, which corresponds to the frequency position, in the quantization matrix.

Another critical issue needs to address is the relationship between the distortion in pixel domain and frequency domain. Because the intra-prediction is performed in pixel-domain, $\bar{d}$ also refers to the average pixel-domain distortion. However, $D_{i,j}$ is frequency-domain distortion. Let d and D be the distortion matrix, measured in mean squared error (MSE), in pixel domain and quantization domain respectively. Assuming that the DCT coefficients at different frequency positions are independent, we have

$$vec(d) = (T' \otimes T') \circ (T' \otimes T') vec(D) \tag{4}$$

where T is the DCT transform matrix. '$\otimes$' denotes the Kronecker product. '$\circ$' is the componentwise multiplication. '$vec(.)$' is the operation of converting a matrix into a vector by stacking all its columns one by one.

Then, the dependent bit allocation optimization problem can be formulated as

$$\begin{aligned} \min_{R_{i,j}} \quad & \sum_{i=0}^{7}\sum_{j=0}^{7} D_{i,j} \\ \text{s.t.} \quad & \sum_{i=0}^{7}\sum_{j=0}^{7} R_{i,j} \leq R \\ & R_{i,j} = \underbrace{\log \frac{\sigma_{i,j}^2}{D_{i,j}}}_{g_1} \\ & \sigma_{i,j} = \alpha_{i,j}\bar{d} + \beta_{i,j} \\ & \bar{d} = \frac{1}{15}\Big(\sum_{i=0}^{7} d_{7,i} + \sum_{i=0}^{6} d_{i,7}\Big) \\ & vec(d) = (T' \otimes T') \circ (T' \otimes T') vec(D) \\ & \sigma_{i,j} \geq D_{i,j} \end{aligned} \tag{5}$$

Unfortunately, the above problem is not a convex optimization problem, since g_1 is not a convex function of $\sigma_{i,j}^2$. Thus, it is difficult to find the optimal bit allocation strategy directly. In this paper, successive convex optimization techniques are employed to solve this optimization problem. Instead of solving (5), a series of convex optimization problems are solved. During each iteration, the nonconvex function g_1 is approximated with a convex function. As pointed out in [6], this iterative approximation will converge to a point satisfying the Karush-Kuhn-Tucker (KKT) conditions of the original problem if the approximation of $f_t(\mathbf{x})$ meets the following 3 requirements:

- $f_t(\mathbf{x}) \leq \tilde{f}_t(\mathbf{x})$ for all $\mathbf{x}$
- $f_t(\mathbf{x}) \leq \tilde{f}_t(\mathbf{x})$ where $\mathbf{x}_0$ is the optimal solution of the approximated problem in the previous iteration.
- $\nabla f_t(\mathbf{x}) = \nabla \tilde{f}_t(\mathbf{x})$

Here we approximate g_1 as

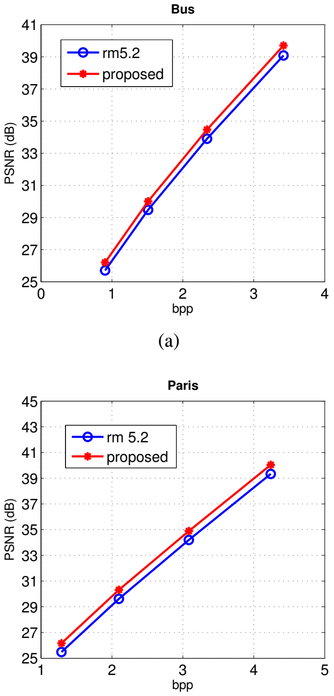
$$g_1 = log(\hat{\sigma}_{i,j}^2) - 1 + \frac{\sigma_{i,j}^2}{\hat{\sigma}_{i,j}^2} \tag{6}$$

where $\hat{\sigma}_{i,j}^2$ is the optimal solution of the approximated problem in previous iteration. To keep the approximation accuracy, $\sigma_{i,j}^2$ is restricted as $(1-\epsilon)\hat{\sigma}_{i,j}^2 \leq \sigma_{i,j}^2 \leq (1+\epsilon)\hat{\sigma}_{i,j}^2$.

With the above convex approximation, we will solve a series of following optimization problem to obtain the solution of (5).

$$\begin{aligned} \min_{R_{i,j}} \quad & \sum_{i=0}^{7}\sum_{j=0}^{7} D_{i,j} \\ \text{s.t.} \quad & \sum_{i=0}^{7}\sum_{j=0}^{7} R_{i,j} \leq R \\ & R_{i,j} = log(\hat{\sigma}_{i,j}^2) - 1 + \frac{\sigma_{i,j}^2}{\hat{\sigma}_{i,j}^2} - \log(D_{i,j}) \\ & \sigma_{i,j} = \alpha_{i,j}\bar{d} + \beta_{i,j} \\ & \bar{d} = \frac{1}{15}\Big(\sum_{i=0}^{7} d_{7,i} + \sum_{i=0}^{6} d_{i,7}\Big) \\ & vec(d) = (T' \otimes T') \circ (T' \otimes T') vec(D) \\ & \sigma_{i,j} \geq D_{i,j}; \\ & (1-\epsilon)\hat{\sigma}_{i,j}^2 \leq \sigma_{i,j}^2 \leq (1+\epsilon)\hat{\sigma}_{i,j}^2 \end{aligned} \tag{7}$$

It can be proved that the problem of (7) is a convex optimization. The optimal solution of (7) can be obtained efficiently [7]. Here the scientific software CVX [8] is employed to get the optimal solution of (7). After the solution of (5) is obtained, the optimal dependent bit allocation strategy is achieved as well.



(a)

(b)

Fig. 4. R-D performance comparison of the proposed method with the conventional method

4. EXPERIMENTAL RESULTS

We evaluate the performance of the proposed dependent bit allocation algorithm using the AVS reference software rm5.2. Various video sequences with different content characteristics are employed in the experiments. For each video sequence, the first frame is encoded as intra-frame. Two different bit allocation strategies are compared: one is the default bit allocation strategy in the reference software, where all the DCT coefficients are quantized with the same quantization parameter; the other one is the proposed dependent bit allocation method.

The experimental results of two typical video sequences, bus and paris, are shown in Fig. 4. It can be seen that compared with the default bit allocation method, the proposed method significantly improves the R-D performance. One thing needs to mention is that the bitrate is calculated as in (1). The coding performance improvement is mainly due to the careful employment of the interblock coding dependency. Under the total bitrate constraint, the proposed method tends to favor the pixels at the block boundaries, since they will be used for the intra-prediction for their following neighbor blocks.

5. CONCLUSIONS

In this paper, we proposed an analytic approach for the dependent bit allocation algorithm for AVS intra-frame coding. Different from the conventional bit allocation method, the interblock coding dependency is fully exploited and leveraged in the bit allocation. After careful formulation, successive convex optimization techniques are employed to solve the original nonconvex problem. A series of convex optimization problems are solved to achieve the optimal solution which satisfies the KKT conditions of the original optimization problem. Experimental results have demonstrated the effectiveness of the proposed method with significant coding performance improvement.

6. REFERENCES

[1] L. Yu, S. Chen, and J. Wang, “Overview of avs-video coding standards,” *Signal Processing: Image Communication*, vol. 24, no. 4, pp. 247–262, Apr. 2009.

[2] P. List, A. Joch, J. Lainema, G. Bjntegaard, and M. Karczewicz, “Adaptive deblocking filter,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 614–619, Jul. 2003.

[3] Y. Ye and M. Karczewicz, “Improved H.264 inta coding based on bi-directional intra prediction, directional transform, and adaptive coefficient scanning,” in *Proc. of IEEE International Conference on Image Processing*, Oct. 2008, pp. 2116–2119.

[4] X. Yu, D.-K. He, and E.-H. Yang, “Adaptive quantization with balanced distortion distribution and its application to H.264 intra coding,” in *Proc. of IEEE International Conference on Image Processing*, Nov. 2009, pp. 1049–1052.

[5] C. Pang, O. C. Au, F. Zou, J. Dai, and R. Cha, “Optimal distortion redistribution in block-based image coding using successive convex optimization,” in *Proc. of IEEE International Conference on Multimedia and Expo*, Jul. 2011, pp. 11–15.

[6] M. Chiang, C. W. Tan, D. P. Palomar, D. O’Neil, and D. Julian, “Power control by geometric programming,” *IEEE Trans. Wireless Commun.*, vol. 6, no. 7, pp. 2640–2651, Jul. 2007.

[7] S. Boyd and L. Vandenberghe, Eds., *Convex Optimization*, Cambridge University Press, 2004.

[8] M. Grant and S. Boyd, “CVX: Matlab software for disciplined convex programming, version 1.21,” `http://cvxr.com/cvx`, Apr. 2011.