DETENTED QUANTIZATION TO SUPPRESS FLICKER ARTIFACTS IN PERIODICALLY INSERTED INTRA-CODED PICTURES IN H.264 VIDEO CODING

Keiichi Chono Yuzo Senda Yoshihiro Miyamoto

Media and Information Research Laboratories, NEC Corporation

ABSTRACT

This paper describes a new flicker suppression method for periodically inserted Intra-coded Pictures (I-pictures) in H.264 video coding. An H.264 encoder periodically inserts I-pictures for channel hopping and random access. The coding noise pattern for I-pictures differs from that of previously appearing Predictive-coded Pictures (P-pictures) because, unlike with P-pictures, inter-frame prediction is not used with I-pictures. This discontinuity in coding noise patterns generates intra-flicker and heavily degrades subjective video quality at low bit rates. We propose Detented Quantization (DQ) to reduce the discontinuity in coding noise patterns between P-pictures and I-pictures. DQ stabilizes the representation levels of input coefficients in I-pictures on the basis of a derivation of those of the inter-coded images produced from previous P-pictures. Simulation results show that DQ reduces intra-flicker by more than 50% in H.264 video coder JM8.6, and it significantly improves subjective video quality.

***Index Terms*—**Video coding, Transforms, Quantization, Motion compensation.

1. INTRODUCTION

Video coding standards are based on motion-compensated prediction, i.e., inter-frame prediction. Residual differences in inter-frame prediction are encoded at each block in a picture. The newest video coding standard, H.264, also employs directional spatial intra-prediction for intra-coding [1]. This spatial intra-prediction is conducted on the basis of the neighboring pixels of adjoining coded blocks. Though spatial intra-prediction improves the intra-coding performance of H.264, it results in more intra-flicker than previous video coding standards [2]. Methods for reducing intra-flicker in H.264 video encoding have been proposed in [2] and [3]. In [2], quantization is applied in the frequency domain to the intra-predicted image prior to subtracting that image from the input image. The method in [3] revises the rate-distortion-optimized cost function in [4], which is used for intra-estimation. This revision is done by adding to that cost function a new parameter: the difference between a current Macro Block (MB) and the MB at the same location in the previous I-picture. Further, this method also refines the Quantization Parameter (QP) and re-encodes a current MB until the difference between it and the previous MB becomes smaller than a pre-determined threshold value. While the methods introduced in both [2] and [3] successfully reduce the amount of intra-flicker that intra-prediction produces between two I-pictures, they also create their own problems. They do not, for example, take into account the fact that intra-flicker results from the difference in coding noise patterns between inter-coding and intra-coding (see Figure 1.), which makes it impossible for them to effectively suppress the flicker in periodically inserted I-pictures. Further, since the method in [2] is not compatible with the H.264 standard, it would be necessary to modify not only the encoder side and but also the decoder side. By way of contrast, while the method in [3] requires no such modification for compatibility, it does require a large calculation-overhead in order to use the rate-distortion-optimized cost function needed to obtain reconstructed images.

This paper proposes Detented Quantization (DQ) to suppress flicker in periodically inserted I-pictures in H.264 video coding. DQ reduces the discontinuity in noise patterns between inter-coding and intra-coding by stabilizing the representation levels of input coefficients in I-pictures, and it effectively suppresses the flicker. Neither modification of the H.264 standard nor a rate-distortion-optimized cost function are required with this proposed method.

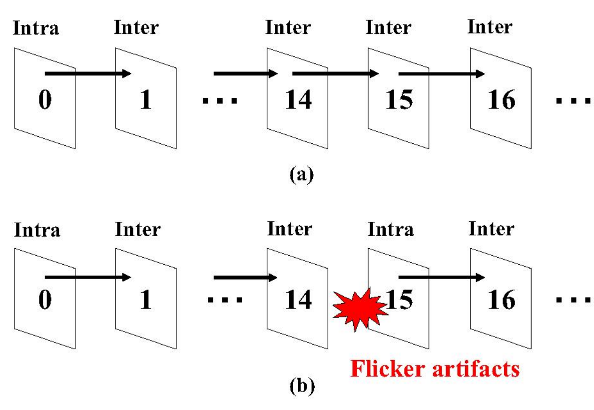


Figure 1: Example of video coding structure: (a) without periodic I-picture and (b) with periodic I-picture.

2. FLICKER ARTIFACTS IN PERIODICALLY INSERTED I-PICTURES

In order to consider why flicker artifacts occur in periodically inserted I-pictures, let us first define reconstructed image R on H.264 as

$$R(O,P) = W(O-P) + P \ , \quad (1)$$

where O denotes the original image and P denotes a predicted image. W(X) is $T^{-1}Q^{-1}QT(X)$. T represents a transform and Q refers to quantization.

Inter-frame predicted images in static and panning regions in successive frames tend to be almost the same ($O15 \cong O14$, $P15 \cong P14$). Reconstructed images in such regions are, therefore, also almost the same ($R15 \cong R14$), as may be deduced from Eq. (1). Hence, coding noise patterns will be continuous in the video coding structure shown in Figure 1 (a). However, coding noise patterns will be discontinuous with periodically inserted I-pictures in the video coding structure shown in Figure 1 (b). Since intra-predicted images in periodically inserted I-pictures differ from inter-frame predicted images in previous P-pictures, reconstructed I-picture images will differ from previously reconstructed P-picture images even if their original images were identical ($O15 \cong O14$, $P15 \neq P14$, $R15 \neq R14$). When, as a result of quantization, this discontinuity has been increased, it comes to be perceived as flicker.

Figure 2 shows flicker power transitions for three video coding structures. The horizontal axis is frame number (time), and the vertical axis is flicker power in [2]. Though there are no peaks in (a), there are strong peaks for I-picture periods in (b). These peaks are an objective representation of flicker. Further, (c) shows that such strong peaks exist even if intra-prediction is restricted to the DC mode, for which increases in intra-flicker are the least among all H.264 intra-modes. This indicates that the flicker in periodically inserted I-pictures is mainly due to the difference in coding noise patterns between inter-coding and intra-coding.

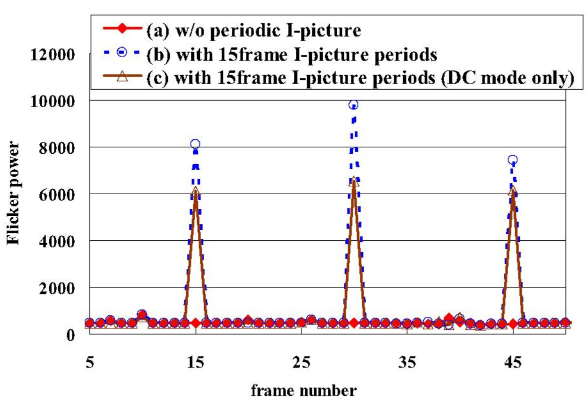


Figure 2: Flicker power transitions in "Woman in Flowers" (JM8.6 QP=32)

3. PROPOSED FLICKER ARTIFACTS SUPPRESSION METHOD

The proposed method includes a modification of the quantization process for MBs in periodically inserted I-pictures. It consists of three steps. In the first step, the inter-coded image for a current MB is derived from the previous P-picture. Next, a representation level (hereafter referred to as the *detent position*) is derived from that inter-coded image. Finally, detented quantization is conducted for input coefficients on the basis of their respective corresponding detent positions. These steps are performed on the encoder side only, and there is no need to modify for the H.264 standard.

3.1. Derivation of inter-coded image

The Inter-coded image C for a current MB in an I-picture is derived as:

$$C = W(O_t - MCP(R_{t-1})) + MCP(R_{t-1}) \ , \quad (2)$$

where t denotes the frame number of the current I-picture. MCP(X) represents the motion-compensated predicted image produced from coded frame X. Flicker tends to be particularly noticeable in panning scenes and in static regions of other types of scenes for which inter-frame prediction is highly accurate. Since the first term of Eq. (2) will then be zero, this equation may be simplified to:

$$C = MCP(R_{t-1}) \ . \quad (2')$$

Coded image C will have a coding noise pattern similar to that of the previous P-picture. This means that discontinuity will be significantly reduced if each of the MBs in an I-picture is encoded so as to resemble as closely as possible its corresponding MB in inter-coded image C. This will not drastically reduce the coding efficiency of the subsequent P-pictures so long as the previous I-picture has been processed in the manner described in 3.3 below.

It will not be necessary to add any circuits in order to perform the MCP in this step on H.264 hardware encoders; the circuit for MCP can be used since it is not otherwise employed in intra-coding. In order to reduce the computation required for determining motion vectors used in the MCP, zero motion vectors can be used as the motion vectors; flicker in static regions will still be suppressed.

3.2. Derivation of the detent position

The detent position D for an input coefficient of a current MB in I-pictures may be expressed as:

$$D = Q^{-1}Q^{roff}(c) = q(k) \ , \quad (3)$$

where c is the inter-coded coefficient derived from inter-coded image C, k denotes the quantization index of inter-coded coefficient c, and q(k) is the value of the corresponding representation level. Coefficient c is defined as $c = T(C - SIP)$, where SIP represents the intra-predicted image produced from adjoining coded blocks. Q^{roff} refers to quantization with round off. Eq. (3) therefore shows that the detent position is the representation level closest in value to the coefficient c which maximizes the continuity of coding noise patterns.

3.3. Detented quantization

DQ is performed to replace the output quantization index L of an input coefficient o, which is defined as $o = T(O_t - SIP)$, with the quantization index k of the detent position q(k) when o is in a detent range. Specifically,

$$L = \begin{cases} k & \cdots \quad \textit{if o is in detent range} \\ Q(o) & \cdots \quad \textit{otherwise}. \end{cases} \tag{4}$$

The detent range is the overlapped range of the inter dead-zone centered at c and the inter dead-zone centered at q(k), as it illustrated in Figure 3. The first dead-zone is introduced to reduce the discontinuity of coding noise patterns between current and the previous pictures, while the second one narrows the range to avoid producing undesirable non-zero coefficients in the coding of the subsequent P-picture, particularly in static regions. When input coefficients lie within the detent range, their representation levels are locked in the detent position q(k). The representation level of an input coefficient o, for example, will be q(k) even though it would actually be closer to q(k+1) than to q(k).

DQ works to reduce the discontinuity in coding noise patterns between inter-coding and intra-coding without a serious degrading of coding efficiency.

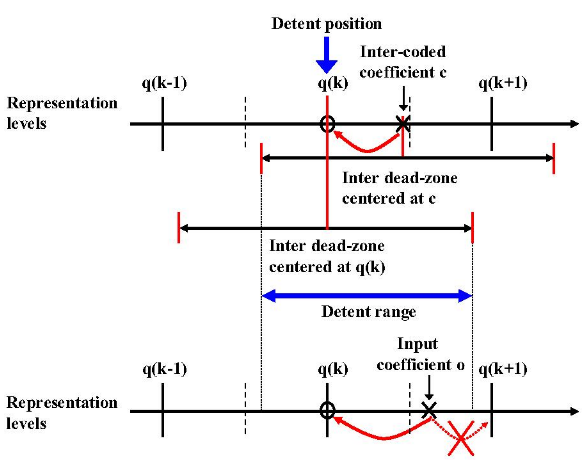


Figure 3: Detented quantization

4. EXPERIMENTS

We have evaluated the effectiveness of DQ in computer simulations. Table 1 shows the simulation conditions. We used H.264 video coder JM86 as an anchor coder. DQ was implemented into JM86 as a proposed coder. The revised intra cost function in [3] was also implemented into JM86 as a conventional coder. It should be noted here that, since the conventional coder uses a rate-distortion-optimized cost function to achieve best performance, it is necessary to use a cost function when the target MB is an MB in an I-picture. To measure flicker for periodically inserted I-pictures, we used the following flicker power S:

$$S = \underset{i,j(SSD(O(i-1,j),O(i,j))<500)}{AVG} SSD(R(i-1,j)-R(i,j), O(i-1,j)-O(i,j)) \tag{5}$$

where AVG denotes the calculation of an average, SSD is Sum of Squared Difference, O(i,j) is the j-th MB of the i-th original frame to be coded as an I-picture, and R(i,j) is the corresponding reconstructed MB. Thus, S represents the increase in the difference between successive frames that can be observed in reconstructed images (it is noted in [2] that subjective testing has shown S to be an effective measure).

Table 1: Test conditions

Software	**JM86[5][6]**
Frame rates	**30 Hz**
I-picture period	**15 frame**
Quantization	**Fixed(QP=20, 24, 28, 32, 36)**
Hadamard	**Used**
RD optimization	**Not used**
Entropy coding	**CABAC**

(a) **(b)**

Figure 4: Test sequences: (a) "Crowded Crosswalk" (SIF, 150frames) and (b) "Woman in Flowers" (SIF, 150frames).

The first frames in two test sequences from [7] are shown in Figure 4. There are static regions in the both of the sequences since the camera does not move. The static region seen at the top of "Crowded Crosswalk" is relatively flat-textured, while that in the upper region of "Woman in Flowers" is highly textured. The quality of these regions tends to be degraded by the flicker.

Figure 5 and 6 show the experimental results for S and PSNR versus bit rate. We can see that S for Proposed is less than half those for Anchor and Conventional. This improvement occurs because DQ reduces the discontinuity in coding noise patterns at periodically inserted I-pictures while Anchor and Conventional cannot do that. Though DQ may not produce representation levels which are optimized in terms of PSNR, as may be seen in the graphs, differences among PSNR values are negligible. Figure 7 illustrates examples of coded I-pictures and the temporal coding noise differences in each coder. We can see that DQ successfully reduce flicker while maintaining picture quality. In this way, DQ significantly improves total subjective video quality.

5. CONCLUSION

In this paper, we have proposed Detented Quantization (DQ) as a new flicker suppression method for periodically inserted I-pictures in H.264 video coding. DQ stabilizes the representation levels of input coefficients in I-pictures on the basis of a derivation of those of the inter-coded images produced from previous P-pictures. Simulation results showed that DQ can reduce intra-flicker by more than 50% in H.264 video coder JM8.6, and it significantly improves total subjective video quality.

6. REFERENCES

[1] Thomas Wiegand et al.: "Overview of the H.264/AVC Video Coding Standard, " IEEE Trans. on Circuits and Systems for Video Technology, Vol. 13, No. 7, pp. 560-576, 2003.

[2] X. Fan et al.: "Flicking Reduction in All Intra Frame Coding," JVT-E070, Joint Video Team of ISO/IEC MPEG & ITU-T VCEG Meeting, 2002.

[3] Shinichi Sakaida et al.: "Adaptive quantization control for reducing flicker of AVC/H.264 intra frames," PCS2004, 2004.

[4] Gary J. Sullivan et al.: "Rate-Distortion Optimization for VIDEO COMPRESSION," IEEE Signal Proc. Magazine, pp.74-90, 1998.

[5] Gary J. Sullivan et al.: "Joint Model Reference Encoding Methods and Decoding Concealment Methods," JVT-J049, Video Team of ISO/IEC MPEG & ITU-T VCEG Meeting, 2003.

[6] JVT Reference Software: http://iphome.hhi.de/suehring/tml/

[7] "Standard test sequence for subjective assessment," ITE, the Institute of Image Information and Television Engineers, Japan.

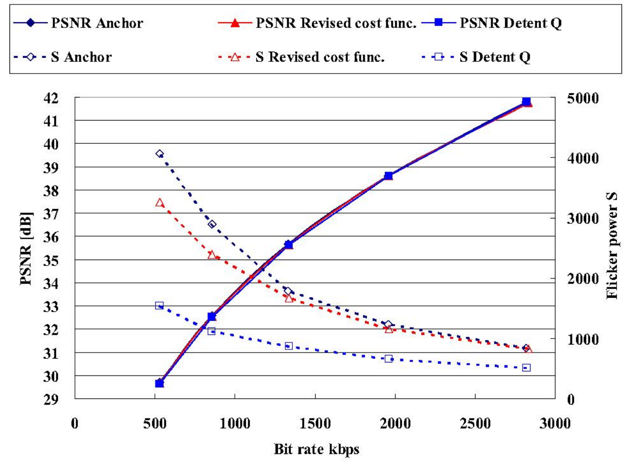


Figure 5: Results for "Crowded Crosswalk"

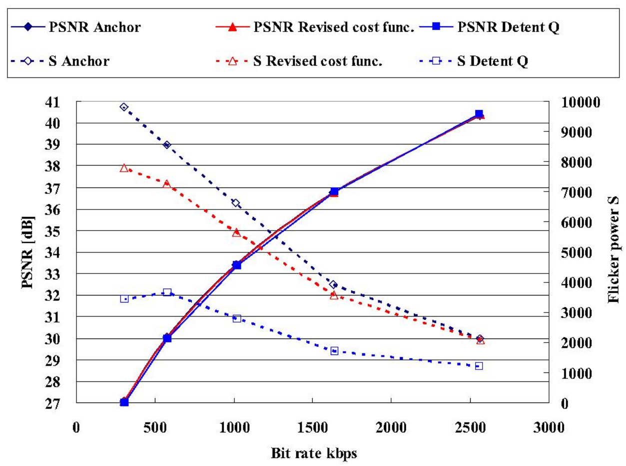
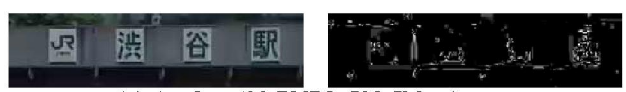


Figure 6: Results for "Woman in Flowers"



(a) Anchor (29.7[dB], 528.5kbps)

(b) Conventional (29.7[dB], 529.2kbps)

(c) Proposed (29.6[dB], 529.2kbps)

Figure 7: Example of coded I-pictures and the temporal coding noise differences for "Crowded Crosswalk".