Title

Flicker suppression in JPEG2000 using segmentation-based adjustment of block truncation lengths

Journal

Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on, 1

Authors

Leontaris, A
Tonomura, Y
Nakachi, T
et al.

Publication Date

2007-04-01

Peer reviewed

FLICKER SUPPRESSION IN JPEG2000 USING SEGMENTATION-BASED ADJUSTMENT OF BLOCK TRUNCATION LENGTHS

Athanasios Leontaris[†], *Yoshihide Tonomura*[‡], *Takayuki Nakachi*[‡], *and Pamela C. Cosman*[*]

[†]Dolby Laboratories, Inc., Burbank, CA 91505, U.S.A.
[‡]NTT Network Innovation Laboratories, Yokosuka, Kanagawa 239-0847, JAPAN
[*]University of California, San Diego, La Jolla, CA 92093, U.S.A.

ABSTRACT

Flickering is a temporal visual artifact that affects compressed video. It is prominent in intra-frame video coders and is largely the result of content variations and quantization. We concentrate on flickering due to quantization. JPEG2000 uses Post-Compression Quantization which is applied through the EBCOT algorithm. EBCOT has been found, however, to cause significant flickering in the reconstructed video. In this work, we evaluate existing flicker metrics, investigate the causes of flicker, and propose a new rate-distortion optimal algorithm that suppresses flicker. The proposed algorithm suppresses temporal flicker at a negligible cost in spatial image quality.

***Index Terms*—** JPEG2000, flicker artifact, rate control, quality metrics, rate-distortion optimization.

1. INTRODUCTION

Flickering is a temporal artifact that arises due to variations in reconstructed low-motion areas across consecutive frames of a compressed image sequence. It is very perceivable at low and medium bit rates in video encoded using JPEG2000. While JPEG2000 has been primarily standardized for use in still imaging applications, it has recently been mandated as the codec of choice for digital cinema applications [1]. JPEG2000, as a still image coder, is especially prone to flickering. However, subjective observations show that, in static areas, hybrid video codecs such as MPEG-2 and H.264/AVC produce flickering comparable to that of JPEG2000.

In [2] flicker artifact was categorized into: (a) flickering due to small temporal variations in the input signal luminance, that give rise to large differences in the de-quantized coefficients [3], and (b) flicker artifact due to uneven rate allocation among collocated blocks across subsequent frames. We are interested in this second type. Type (a) artifact is content-dependent and is perceived as ripples mainly in edges. Type (b) artifact is perceivable in static areas (often the background) across consecutive frames; it is also augmented when intra-frame coders (JPEG2000) are used. It was first suggested in [4] and [5] that flickering is a result of dynamic quantization. In particular, for the case of JPEG2000, it is a result of the post-compression (PC) quantization (EBCOT). The EBCOT rate allocation scheme of JPEG2000 performs PC quantization through code-block bit-plane truncation on an *intra-frame* basis. In [6] it was noted that even if most of the frame is static, a slight content change (e.g. object motion) can significantly alter the rate allocation among static blocks of two subsequent frames. As a result, collocated JPEG2000 code-blocks of bitplanes of wavelet coefficients, corresponding to identical content in subsequent frames, are often truncated at different lengths.

Kuge [3] and Kato et al. [5], studied the type (a) and (b) artifacts, respectively, but did not propose a solution. More recently, Becker et al. [6] proposed an approach for type (a) artifact suppression. A family of methods was proposed in [2] to suppress type (b) flicker. One method called TRUNC sought to keep the quantization error distribution equal within a single frame, and also keep the coding pass truncation lengths almost constant across consecutive frames. The other proposed method called HYBRID used a segmentation algorithm to identify those blocks in a video sequence that are most prone to flicker (relatively static content) and applied EBCOT on blocks that are deemed active and TRUNC to blocks that contribute to flicker. However, aliasing artifact was observed in TRUNC and also HYBRID in [2].

In this work, we address the type (b) flicker artifact. We propose a JPEG2000 standard-compliant solution, that modifies the post-compression rate control scheme and suppresses flickering. The proposed algorithm is rate-distortion optimal and does not suffer from aliasing. The paper is outlined as follows: We discuss existing flicker artifact metrics and causes of flickering in Section 2. In Section 3 we discuss a new rate-distortion optimal rate allocation scheme that takes into account flickering using a novel flicker artifact heuristic metric. Experimental results and conclusions are presented in Section 4.

2. FLICKERING: METRICS AND CAUSES

To address flickering, it is critical to measure it reliably. The metric used in [2] is based on a classification map that is derived using change detection. Using the original image blocks, the map categorizes code-blocks into those prone to flicker and those that are not. A simple No-Reference metric [5, 2], examines the variance of MSE/PSNR for the reconstructed blocks that were classified as prone to flickering during the initial segmentation phase. Variance of the quality of these blocks is an indication of flickering.

Fan et al. [7] proposed a Full-Reference metric that is nonzero for blocks deemed static according to their similarity to their collocated counterpart in the previous frame. For those blocks the metric is calculated as the Sum of Squared Differences (SSD) between the original block difference and the reconstructed block difference. The block difference is obtained by subtracting the collocated block in the previous frame from the current block.

Drelie et al. in [8] use a segmentation map to differentiate blocks prone to flicker. For those blocks, the metric is calculated as a frac-

This work was performed while the first author was at UCSD. It was supported in part by the Center for Wireless Communications at UCSD, the Office of Naval Research, and the UC Discovery Grant program of the State of California.

tion: The numerator is the difference of the SSD (a) of the original and reconstructed blocks for the current frame, and the SSD (b) of the original and reconstructed blocks for the previous frame. The denominator is the sum of the above two SSDs.

The last metric we evaluated was proposed by Winkler et al. in [9]. It uses the 2D Fourier Transform to perform frequency analysis of the signal and then calculates two sub-metrics for low and high frequencies. The final metric is the sum of those sub-metrics divided by terms that account for temporal and spatial activity.

bit rate (bpp)	PSNR	[7]	[8]	[9]
PSNR	1.0000	0.1421	0.1705	-0.0091
Fan et al. [7]		1.0000	0.8696	0.2959
Drelie et al. [8]			1.0000	0.2043
Winkler et al. [9]				1.0000

Table 1. Cross-correlations of metric outputs for JPEG2000 EBCOT-coded sequence at 0.50bpp.

The cross-correlations between the metrics per-frame output for the synthetic Circular Zone Pattern (CZP) sequence encoded at 0.50 bits per pixel (bpp) with JPEG2000 EBCOT are presented in Table 1. Although not shown, we note that these correlation values are almost the same for bit rates ranging from 0.2 to 1.4bpp. We observe that the three flicker metrics are not correlated with PSNR. The Fan et al. and Drelie et al. metrics are well correlated, however. This is also visually confirmed by the plot of the metric responses in Fig. 1. The Winkler et al. metric however performs badly. It seems to be more correlated with quantization errors rather than flicker artifacts. We conducted small scale subjective tests and confirmed that all of the spikes registered in Fig. 1 by both the Fan et al. metric and the Drelie et al. metric correspond to actual visual spikes of flickering. We thus believe that these two metrics perform well in registering flicker. We use them in Section 4 to evaluate our proposed algorithm.

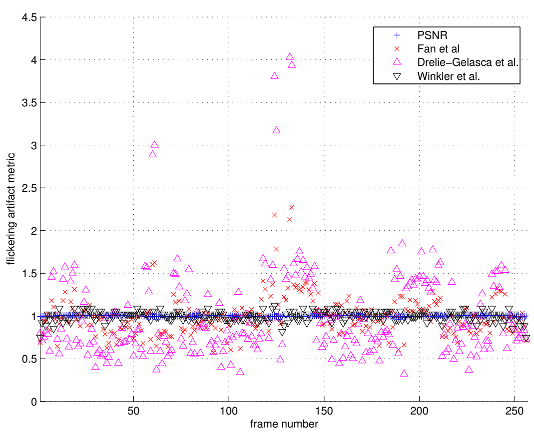


Fig. 1. PSNR and three flicker metrics versus frame number for the CZP image encoded with JPEG2000 (EBCOT) at 0.125 bpp.

Causes of Flickering: The influence of dynamic quantization noise in video quality was first discussed in [4]. To gain insight into the causes of flicker, we modified the EBCOT codec to output information on its choice of coding pass truncation lengths on a block and subband basis. The frames were decomposed into six levels. Level 0 contains the LL subband, while levels 1 through 5 each consist of three subbands, LH, HL, and HH. Level 0 is the coarsest resolution level and level 5 is the highest resolution level. Two such "frames" of coding pass truncation lengths are shown in Fig. 2. They come from frames 60 and 61 of the CZP sequence (a synthetic gray-scale sequence with 257 frames at a resolution of 512×256 pixels) encoded at 0.125bpp. Higher color intensity indicates a higher level of bitplane transmitted. In addition, the map is color-coded: we use three colors for the three types of coding passes in JPEG2000: significance (red), magnitude-refinement (blue), and cleanup (green) pass. The spatial output of the Fan et al. metric (we note that the Drelie et al. is quite similar qualitatively) for these two frames is shown in Fig. 3. The intensity of the color (here only green) denotes the strength of the artifact.

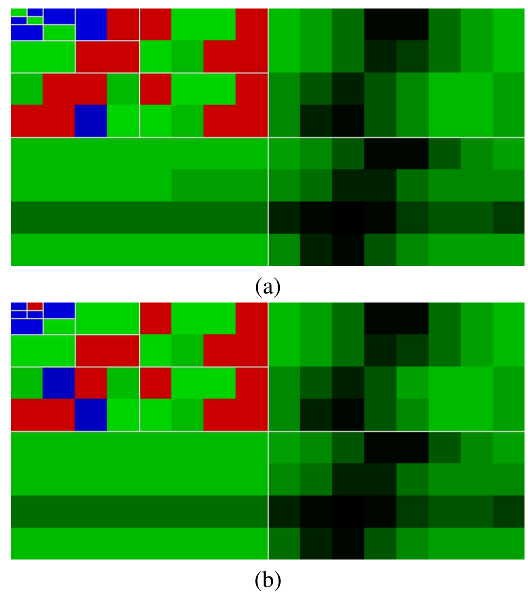


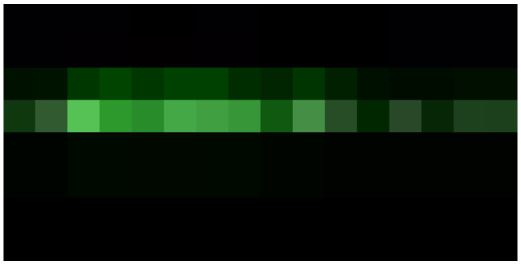
Fig. 2. EBCOT coding-pass truncation length map for frames (a) 60, (b) 61.

We observe in Fig. 3 a large spike in flicker artifact. In Fig. 2, we find that going from frame 60 to frame 61 there is a change in the truncation length in the second from the top 32×32 block of the bottom left subband of level 4. The block's color turns from red to blue. We repeated this heuristic search for other cases of flicker artifact spikes and were able to tie all such spikes to specific changes of coding pass truncation lengths. All these spikes were also highly visible.

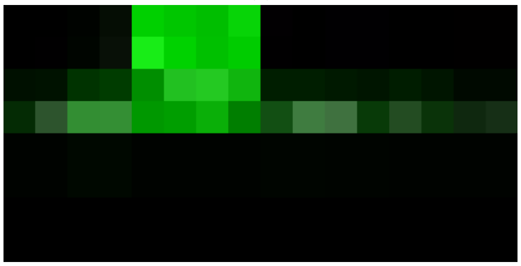
Specifically, we noticed that the severe flicker artifacts, that become obvious during subjective evaluation, are associated with specific changes in pass truncation lengths mainly at levels 4 and 5. This finding agrees with subjective observations conducted at NTT Labs that pointed to levels 4 and 5 as the biggest sources of flicker. Therefore, we will also experiment with a simple codec called "4-5" where levels 0 through 3 were encoded with EBCOT and the rest with TRUNC from [2].

3. A NEW APPROACH

The use of TRUNC and HYBRID in [2] almost eliminated flicker but this gain came at the cost of a significant drop in PSNR recon-



(a)



(b)

Fig. 3. Fan et al. metric results visualization for frames (a) 60, (b) 61.

struction quality. The truncation of code-block bitplanes at equal lengths with those algorithms suppresses flicker artifact at the cost of increased aliasing artifacts. High-frequencies are not encoded well because bits are starved at these locations. This artifact was decreased considerably with the classification maps of HYBRID, but was not eliminated completely. Hence we were motivated to devise a scheme that can counter both flicker and aliasing, addressing the shortcomings of both HYBRID and TRUNC. Note that subjective experiments show that aliasing is often preferable to flicker.

A rate-distortion optimized solution for flicker artifact suppression was proposed for INTRA frames of H.264/AVC in [10]. Their method effectively minimized the following Lagrangian cost:

$$J = D + \mu F + \lambda R \tag{1}$$

where D is the distortion, F is the flicker artifact measure, and R is the bit rate. μ and λ are constants chosen empirically to fulfill the compression efficiency vs. flicker suppression trade-off. The measure F used was essentially the Fan et al. metric [7]. The method performed well for H.264/AVC but requires the reconstruction of the blocks, thereby increasing complexity.

We adopt an approach better suited for use within the EBCOT post-compression rate allocation framework. First we label the block as prone to flicker (static) or not (active). If it is prone to flicker, then the rate allocation algorithm takes into account the resulting flicker. The spatial activity map is derived as in HYBRID with a tuned change detector. This is fast to compute as it operates on the pixels of the original sequence and does not require additional encoding or decoding.

As we go through each block during EBCOT rate allocation, if the block is classified as active, then the coding pass truncation length is determined with the traditional EBCOT rate allocation. If the block is classified as static then there are two passes: first the usual EBCOT pass truncation length is determined. If the bitplane t_n^* at which the block is truncated using EBCOT is equal to the bitplane t_{n-1} selected for the collocated block in the previous frame, then nothing further is done. If, however, they differ, then we go over the range of potential truncation lengths ranging from t_{n-1} to t_n^*, and select the one that minimizes the following cost:

$$\min_{t \in \left[t_{n-1}, t_n^*\right]} F(t_{n-1}, t) + \mu \left| R_{t_n^*} - R_t \right| \tag{2}$$

The decision is thus a trade-off of flicker artifact and rate difference from the R-D optimal rate. The trade-off is controlled by the constant parameter μ. The flicker artifact $F(\alpha, \beta)$ can be written as:

$$F(\alpha, \beta) = \gamma(c, l, s)\, |\alpha - \beta| \tag{3}$$

The weighting parameter γ may depend on the color component c, the level l, and the subband s. In our implementation it only changes for varying level l, and incorporates our conclusions from the EBCOT flicker study in Section 2. We set $\mu = 10$ and set $\gamma = [1, 2, 5, 10, 6]$ for $l = [1, 2, 3, 4, 5]$. Hence γ has higher values for levels 4 and 5. Our algorithm (denoted "NEW" below) incorporates both the segmentation map that functions as a switch, as well as the empirical conclusions on the significance of the subband levels and our visual inspection (through γ). Still, it is a function of the bitplane truncation lengths difference $\alpha - \beta$. We note that the encoded bit stream is JPEG2000 compliant.

4. RESULTS

For our experiments we employed and modified the Kakadu software implementation of JPEG2000 [11]. The wavelet decomposition levels were set to five and the code-block size to 32×32 pixels, as mandated by the DCI specification [1]. We evaluated two different image sequences: (a) the CZP sequence, and (b) the RGB "Mother-Daughter" sequence with 300 frames at a resolution of 352×288 pixels. We evaluate five JPEG2000 rate allocation schemes: (a) EBCOT as implemented in the Kakadu software, (b) TRUNC as proposed in [2], (c) HYBRID as proposed in [2], (d) "4-5" mentioned in Section 2, and (e) NEW which is proposed in Section 3.

The response of the Drelie et al. flicker metric on all schemes is presented in Fig. 4 for both sequences. EBCOT has the highest flicker and the lowest is achieved by TRUNC. HYBRID performs the worst for "mother-daughter" which is not expected as in other sequences it performs close to TRUNC. The NEW codec however manages to have much better flicker performance compared to the completely heuristic 4-5 codec and EBCOT. It achieves flicker values close to TRUNC.

Figure 5 shows the PSNR for the five different rate allocation schemes. We observe that the NEW codec achieves PSNR performance almost identical to that of EBCOT. All the other codecs have inferior performance by about 1dB compared to NEW.

In conclusion, flickering is a major visual artifact of video coders and is most prominent in coders that make use of intra-frame coding such as JPEG2000 or H.264/AVC Intra coding. In JPEG2000 it is mainly a result of dynamic quantization through the EBCOT post-compression quantization scheme. Previous work addressed this problem by keeping the code-block truncation length constant across frames, but this introduced aliasing. In this work, we studied the relationship of flickering with JPEG2000 truncation lengths. We evaluated existing flicker metrics and proposed a novel standard-compliant JPEG2000 rate allocation scheme that suppressed flickering without sacrificing objective PSNR performance. A major advantage of our scheme is its low complexity as it avoids decoding the blocks at the encoder side.

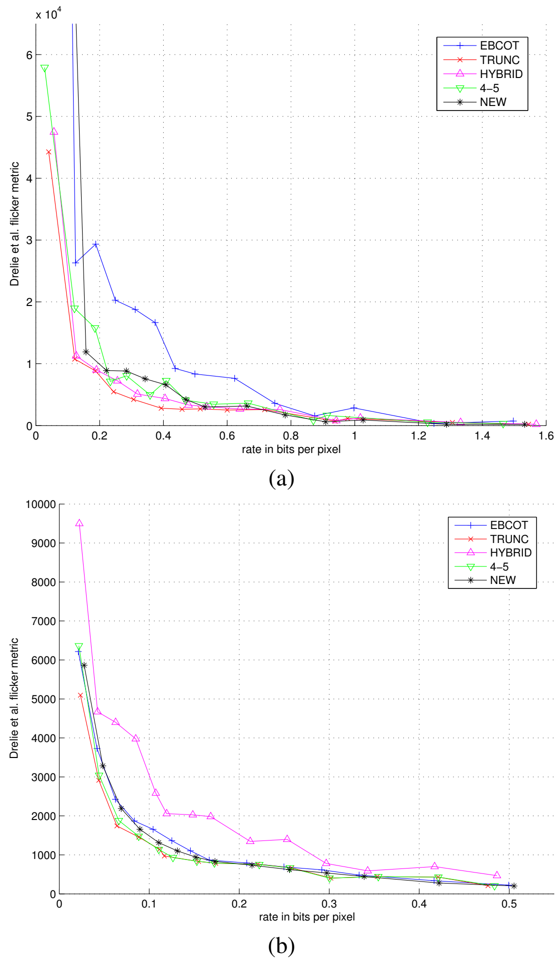


Fig. 4. Flicker Metric Results for (a) CZP, (b) "Mother-daughter" sequence.

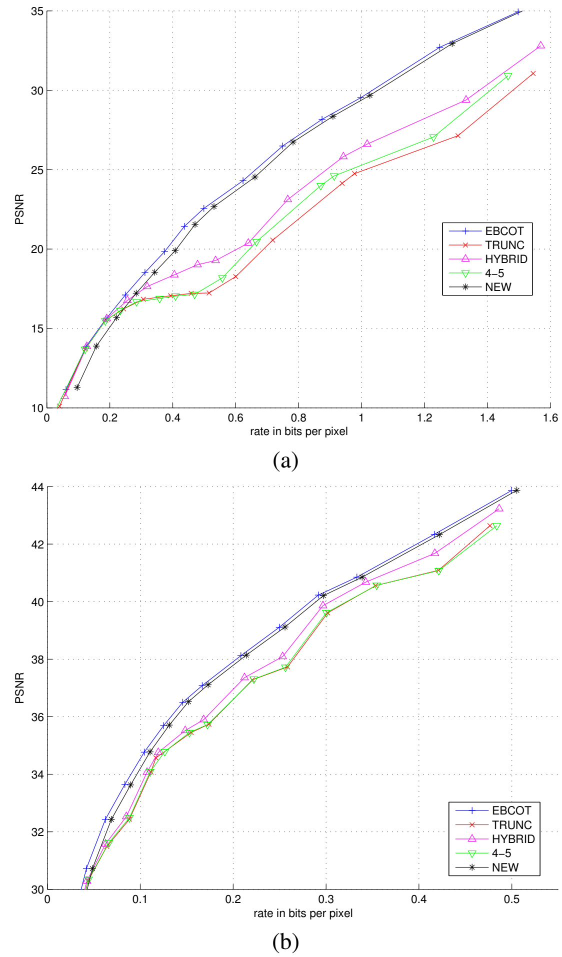


Fig. 5. PSNR Metric Results for (a) CZP, (b) "Mother-daughter" sequence.

5. REFERENCES

[1] "Digital cinema system specification V1.0," Digital Cinema Initiatives, LLC, Member Representatives Committee, July 2005.

[2] A. Leontaris, Y. Tonomura, and T. Nakachi, "Rate control for flicker artifact suppression in motion JPEG2000," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, May 2006, vol. 2, pp. 41–44.

[3] T. Kuge, "Wavelet picture coding and its several problems of the application to the interlace HDTV and the ultra-high definition images," in *Proc. IEEE International Conference on Image Processing*, June 2002, vol. 3, pp. 217–220.

[4] T. Carney, Y.-C. Chang, S. A. Klein, and D. G. Messerschmitt, "Effects of dynamic quantization noise on video quality," in *Proc. SPIE Human Vision and Electronic Imaging IV*, May 1999, pp. 141–151.

[5] T. Kato, S. Tanabe, H. Watanabe, and H. Tominaga, "On a relationship between flicker artifacts and quantization error for Motion JPEG2000," in *Proc. FIT2003, J-039*, Sept. 2003.

[6] A. Becker, W. Chan, and D. Poulouin, "Flicker reduction in intraframe codecs," in *Proc. IEEE Data Compression Conference*, Mar. 2004, pp. 252–261.

[7] X. Fan, W. Gao, Y. Lu, and D. Zhao, "Flicking reduction in all intra frame coding," Joint Video Team of ISO/IEC MPEG and ITU-T VCEG, JVT-E070, Oct. 2002.

[8] E. Drelie-Gelasca and T. Ebrahimi, "On evaluating metrics for video segmentation algorithms," in *Proc. Second Int. Workshop on Video Processing and Quality Metrics for Consumer Applications*, Jan. 2006.

[9] S. Winkler, E. Drelie-Gelasca, and T. Ebrahimi, "Towards perceptual metrics for video watermark evaluation," in *Proc. SPIE Digital Image Processing XXVI*, Nov. 2003, pp. 371–378.

[10] S. Sakaida, K. Iguchi, S. Gohshi, and Y. Fujita, "Adaptive quantization control for reducing flicker of AVC/H.264 intra frames," in *Proc. Picture Coding Symposium*, Dec. 2004.

[11] D. Taubman, "High performance scalable image compression with EBCOT," *IEEE Transactions on Image Processing*, vol. 9, no. 7, pp. 1158–1170, July 2000.