MODIFIED H.264 INTRA PREDICTION FOR COMPRESSION OF VIDEO AND IMAGES CAPTURED WITH A COLOR FILTER ARRAY

Colin Doutre and Panos Nasiopoulos

The University of British Columbia, Vancouver, Canada
Email: colind@ece.ubc.ca, panos@ubc.ece.ca

ABSTRACT

Most consumer digital cameras capture color information with a single light sensor and a color filter array (CFA). In these cameras, only one color sample (red, green or blue) is captured at each pixel location. This paper presents a modified H.264 intra prediction scheme for compressing image and video data captured with a color filter array. The H.264 intra prediction modes are modified for the green channel to account for the fact that the green data is not sampled in a rectangular manner in the Bayer pattern, the most popular CFA design. The proposed method increases the compression efficiency of I frames on the green channel by up to 1.2 dB.

Index Terms— intra prediction, video compression, Bayer pattern, color filter array (CFA), single sensor camera

1. INTRODUCTION

Most consumer digital cameras capture color images using a single light sensor and a Color Filter Array (CFA). Instead of using three sensors to capture a red (R), green (G) and blue (B) sample at every pixel location, these cameras use a single light sensor and only capture one color at each location. A color interpolation stage, known as demosaicking, is required to generate the two missing colors at each pixel location. The most commonly used CFA design is the Bayer Pattern (Fig. 1).

The traditional approach to compressing image or video data captured with a CFA is to first perform demosaicking and then compress the resulting full color data with standard methods. This approach is sub-optimal because demosaicking introduces redundancy into the data that the compression stage must remove. Also, demosaicking expands the number of samples by a factor of three, which increases the compression processing time. Consequently, some methods have been proposed for compressing the raw CFA data prior to demosaicking [1]-[4]. Empirical and theoretical studies have shown that the compression-first approach provides better compression efficiency than the traditional demosaick-first approach at high bit rates [5].

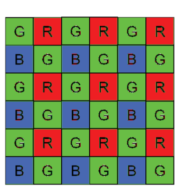

Fig. 1. Bayer Pattern CFA.

In [1], Lee and Ortega propose a method for CFA image compression involving a modified YCbCr conversion followed by rotating the Y samples 45° to form a compact shape. Their YCbCr conversion is performed on each 2x2 cell in the Bayer pattern and produces a Y sample at each green location and one Cb and Cr sample for each cell (Fig. 2). Because the Y values are not sampled in a rectangular manner, they are rotated 45° to form a compact rhombus shape.

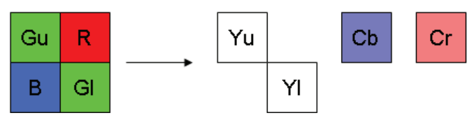

Fig. 2. Modified YCbCr conversion performed on the Bayer unit cell.

Another method for compressing CFA images is the structure conversion method proposed by Koh et al. in [2]. This structure conversion method involves separating the CFA data into three separate arrays of the red, green and blue samples. To fit the green data into a rectangular array, the two green samples in every Bayer pattern unit cell are merged into a single column. These arrays can then be compressed with standard image compression methods such as JPEG. The structure conversion method can be done either in RGB space or with the modified YCbCr conversion proposed in [1].

A simple compression method for CFA video is presented in [3]. They use the structure conversion method to form separate arrays of green, blue and red samples which are compressed with a custom video codec similar to MPEG-2.

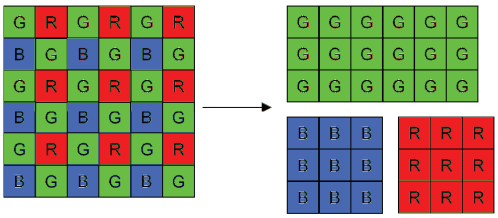

Fig. 3. Structure conversion method for converting mosaic data into separate R, G and B arrays

A CFA video compression method based on the H.264/AVC video coding standard is proposed in [4]. In that paper, the structure conversion method (Fig. 3) is used for generating separate green, blue and red arrays which are compressed together with H.264 in 4:2:2 sampling mode. The motion estimation and motion compensation methods are modified to exploit inter-color correlation during sub-pixel interpolation.

Note that existing CFA video compression methods operate in RGB colour space rather than the modified YCbCr space [3][4]. This is because the modified YCbCr conversion interferes with motion compensation. The modified YCbCr acts on a Bayer cell consisting of four samples. If there is motion, different RGB samples will be grouped together to generate the Y, Cb and Cr samples. Therefore, even if there is simple translational motion, the Y, Cb and Cr samples will change between frames and motion compensation will be less effective.

The H.264 video coding standard uses intra frame prediction to increase coding efficiency in I frames. Intra prediction involves predicting the samples for the block being coded based on surrounding, previously coded, pixels in the same frame. Previous work on CFA video compression does not modify intra frame coding based on the CFA structure [3][4].

In this paper, we propose a modified H.264 intra prediction scheme for image/video data captured with a Bayer CFA. The prediction modes are modified to account for the fact that the green data in the Bayer pattern is not sampled in a rectangular manner. The rest of this paper is organized as follows. Standard H.264 intra prediction is reviewed in Section 2. The proposed modified intra prediction scheme is described in section 3. Experimental results are presented in Section 4 and conclusions are given in Section 5.

2. STANDARD H.264 INTRA PREDICTION

H.264 supports intra prediction on blocks of 4x4 pixels and 16x16 pixels. There are nine different 4x4 prediction modes. Referring to Figure 4, the 16 samples of a 4x4 block labeled a-p are predicted based on their spatial neighbors labeled A-Q.

One 4x4 prediction mode is DC, in which all 16 samples in the 4x4 block are predicted with the average of the eight surrounding samples A-D and I-L. The rest of the 4x4

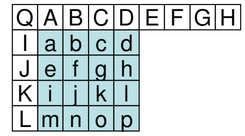

Fig. 4. Intra prediction on a 4x4 block. Lower case letters (a-p) represent the pixels in the block being predicted and upper case letters (A-Q) represent the surrounding pixels used for prediction

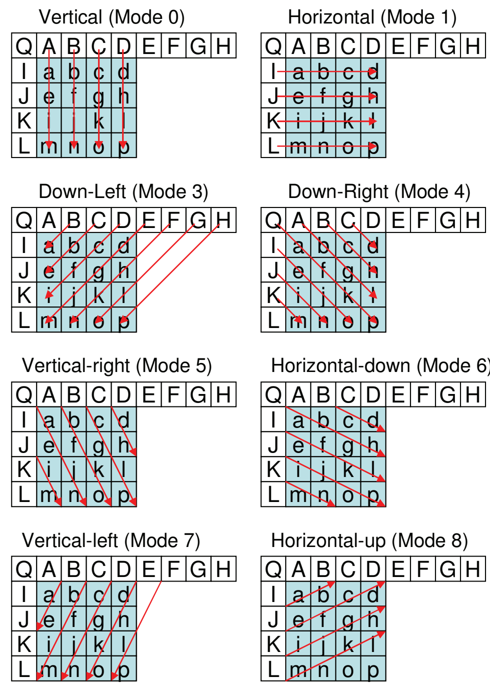

Fig. 5. Standard H.264 directional 4x4 intra prediction modes

prediction modes are directional, and are illustrated in Fig. 5.

In the vertical and horizontal prediction modes, each sample is predicted from the sample directly above or to the left of it (Fig. 5). In the down-left and down-right modes, each sample is a weighted average of three samples in that direction, with weights 0.25, 0.5 and 0.25. For example, in the down-right mode, samples 'i' and 'n' are predicted as:

$$i = n = 0.25I + 0.5J + 0.5K \qquad (1)$$

In the four remaining modes (5 through 8) each sample is also predicted as a weighted average of the samples at the base of the line going through that sample (either two or three source samples are used to form the average).

H.264 also supports luma intra prediction on 16x16 blocks. There are four 16x16 prediction modes: vertical, horizontal, DC and plane. The first three are equivalent to 4x4 modes. Plane prediction fits a smooth linear function to the surrounding luma samples.

3. PROPOSED MODIFIED INTRA PREDICTION

The intra prediction modes in H.264 have been designed assuming the data is sampled in a rectangular fashion. For

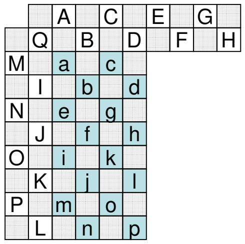

Fig. 6. Sample labeling for modified intra prediction of the green channel

Bayer CFA data, the green channel is not sampled on a rectangular grid, so we propose modified prediction modes to be applied to the green channel.

When an array of the green data is formed through the structure conversion method (Fig. 3), the true spatial sampling of the pixels in the 4x4 block is as shown in Figure 6. The gray squares in Fig. 6 represent locations where either a red or blue sample is captured in the Bayer pattern. If standard H.264 intra prediction is used, most of the directional prediction modes will not work correctly. For example, since samples ‘a’ and ‘c’ are not horizontally aligned with sample ‘I’, the horizontal prediction will not work as intended.

In our method, we use four additional surrounding samples for intra prediction, labeled M-P in Fig. 6.

We modify the DC prediction mode to use a twelve point average of samples A-D, I-L and M-P. The directional prediction modes are modified as shown in Fig. 7. In the horizontal and vertical modes, each sample is predicted based on the sample from which the arrow originates in Fig. 7. For the remaining directional modes (3-8), each sample is predicted as a weighted average of the three samples surrounding the base of its arrow with weights 0.25, 0.5 and 0.25. For example, in down-right prediction mode (mode 4), samples ‘c’ and ‘d’ are predicted as:

$$c = d = 0.25A + 0.5B + 0.25C \tag{2}$$

As another example, in vertical-right mode (mode 5), samples ‘e’, ‘j’ and ‘o’ are predicted as:

$$e = j = o = 0.25M + 0.5I + 0.25N \tag{3}$$

Exceptions occur in modes 3, 7 and 8 due to lack of source samples in the required direction. In these cases, the closest source sample is simply copied to form the prediction. In mode 3, samples ‘h’, ‘k’, ‘j’, ‘l’, ‘m’, ‘o’, ‘n’ and ‘p’ are predicted with H. In mode 7, sample ‘p’ is predicted with H. Finally, in mode 8, samples ‘l’, ‘m’, ‘n’, ‘o’ and ‘p’ are predicted with L.

Note that all 4x4 prediction modes have been modified in the proposed method except the vertical mode (mode 0), since it works as intended without modification.

The 16x16 prediction modes 1 (horizontal) and 2 (DC) are modified in our proposed method in the same way that the corresponding 4x4 modes have been modified.

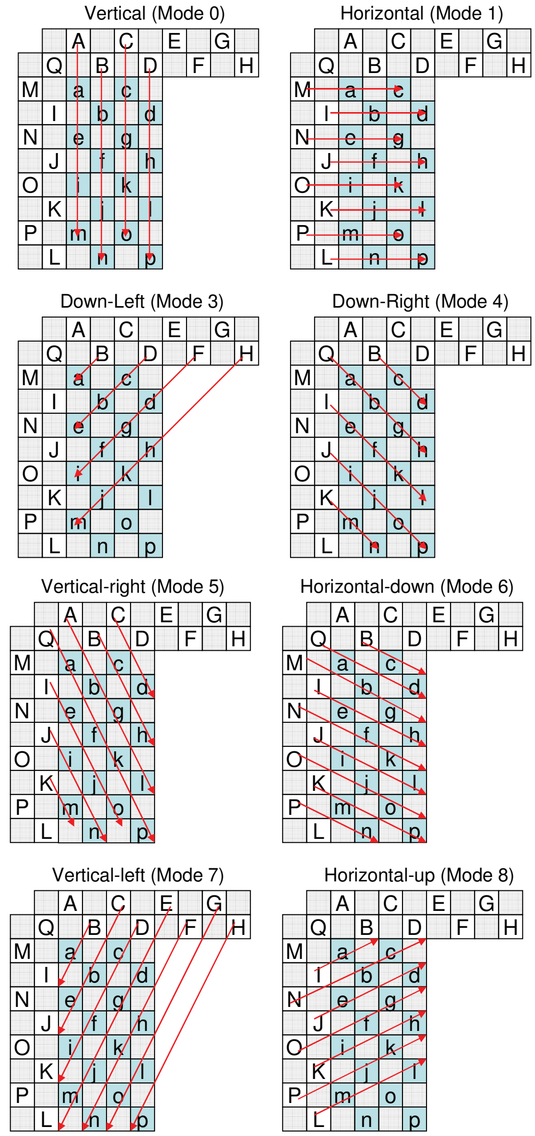

Fig. 7. Modified directional 4x4 intra prediction modes

Our method can be applied either to the green samples in the Bayer pattern, or the luma samples obtained through the modified YCbCr conversion in [1] (Fig. 2), as the sampling of the Y samples in the modified conversion is the same as the sampling of the green samples.

The proposed method can easily be modified for intra prediction on other block sizes or other CFA designs.

4. EXPERIMENTAL RESULTS

We have implemented the proposed modified intra prediction scheme into the H.264 Reference software (JM 12.4). Fig. 8 shows plots of PSNR vs. bitrate of the green channel of four test videos, Foreman, Tempete, Paris and Mother and Daughter. The test videos were obtained by converting the videos to RGB color space, sampling with the

Bayer pattern, and applying the structure conversion method (Fig. 3) to the green channel to form a compact array of green samples.

60 frames of each video were compressed, with every frame coded as an I frame. CABAC was used for entropy coding, and high complexity rate-distortion optimization was used (RDOptimization=1).

We compare our proposed method to compressing with standard H.264 intra prediction, which is the way I frames are coded in [4], and equivalent to the method in [2] with H.264 intra coding used instead of JPEG. For the same bit-rate, the proposed method increases the quality of I frames by about 1 dB in PSNR. The BD-PSNR gains calculated with the Bjontegaard metric [6] are as follows: 1.12 dB on Foreman, 0.96 dB on Tempete, 1.16 dB on Paris, and 0.95 dB on Mother and Daughter.

5. CONCLUSIONS

In this paper a modified H.264 intra prediction scheme is proposed for compressing image/video data captured with a color filter array. Modified intra prediction modes are applied to the green channel to account for the fact that it is not sampled in a rectangular fashion in the Bayer pattern. The proposed method increases I frame quality on the green channel by about 1 dB in PSNR on several standard test videos.

6. REFERENCES

[1] S.Y. Lee and A. Ortega, "A novel approach of image compression in digital cameras with a Bayer color filter array," IEEE Int. Conf. Image Processing 2001, vol. 3, pp. 482-485, Oct 2001.

[2] C.C. Koh, J. Mukherjee, S.K. Mitra. "New efficient methods of image compression in digital cameras with color filter array". IEEE Transactions on Consumer Electronics 2003; 49(4):1448–56.

[3] F. Gastaldi, C. C. Koh, M. Carli, A. Neri and S. K. Mitra, "Compression of videos captured via bayer patterned color filter arrays," Proc. 13th European Signal Processing Conference (EUSIPCO), Antalya, Turkey, Sep. 2005.

[4] C. Doutre, P. Nasiopoulos, and K.N. Plataniotis, "H.264 Based Compression of Bayer Pattern Video Sequences." IEEE Transactions on Circuits and Systems for Video Technology, vol. 18, no. 6, pp. 725 – 734, Jun. 2008

[5] N.X. Lian, L. Chang, V. Zagorodnov, Y. P. Tan, "Reversing Demosaicking and Compression in Color Filter Array Image Processing: Performance Analysis and Modeling," IEEE Transactions on Image Processing, vol. 15, no. 11, pp. 3261–3278, Nov. 2006.

[6] G. Bjontegaard, "Calculation of average PSNR differences between RD curves," VCEG-M33, ITU-T VCEG Meeting, Austin, TX, 2-4 April, 2001.

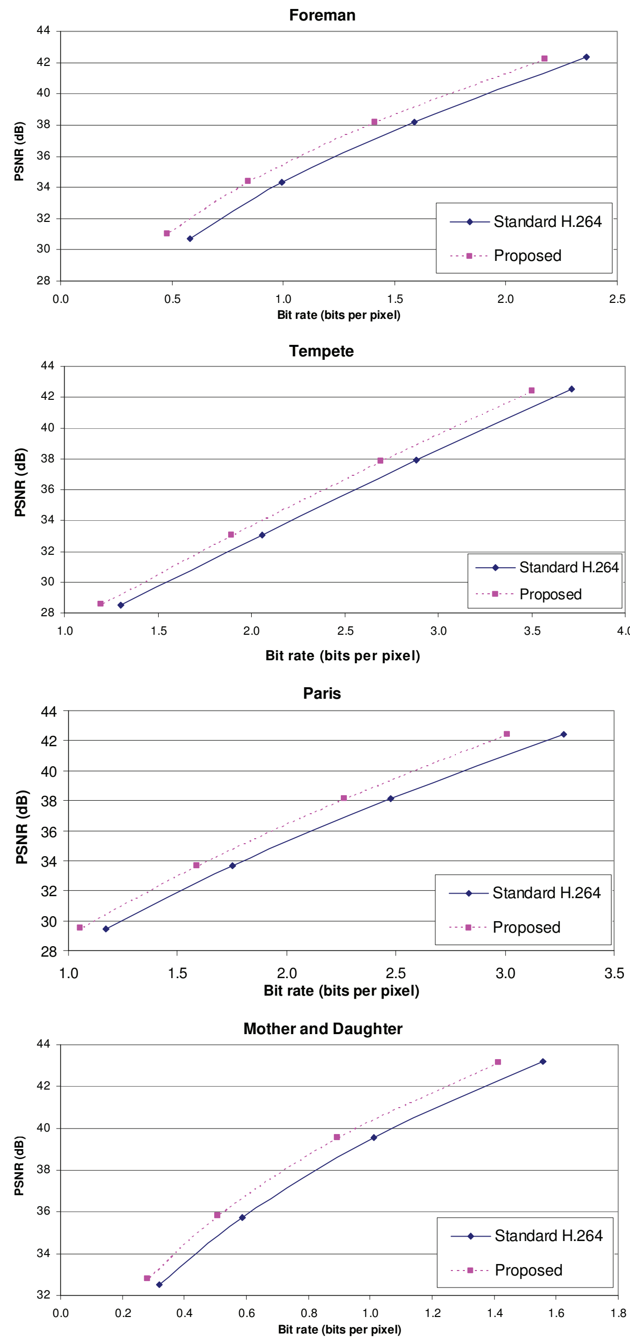

Fig. 8. Rate-distortion curves of the green channel for four test videos.