COMPRESSION OF VIDEOS CAPTURED VIA BAYER PATTERNED COLOR FILTER ARRAYS

F. Gastaldi⋆*, C. C. Koh*†*, M. Carli*⋆*, A. Neri*⋆ *and S. K. Mitra*†

⋆ AE Dept., University of 'Roma TRE', Rome, Italy
† ECE Dept., UCSB, CA, USA

ABSTRACT

In this paper we address the problem of the compression of a video sequence acquired by most inexpensive single sensor video cameras. For each pixel in a frame only one chrominance component is available and an interpolation is used to obtain the full color frame. Our goal is to compress the video directly from the Bayer *color filter array* (CFA) data. We propose a new method for the reduction of temporal redundancy in video sequences. Our approach consists of a pre- and post-processing phases in combination with a standard motion prediction scheme. Simulation results confirm the effectiveness of the proposed method. Compared to standard methods, the improvement in quality is achieved at low and high compression rates. The proposed method offers bandwidth reduction where videos are transmitted over a communications link at low bit-rates while maintaining the same quality produced in the conventional method.

1. INTRODUCTION

Most inexpensive digital video cameras use a CFA with each pixel element of the sensor recording the intensity information of one color component, typically red, green, or blue. Although several different CFAs have been proposed [1], the Bayer CFA [2] shown in Figure 1 is widely used.

Here the green filters are in a quincunx (interlaced) pattern with the red and blue filters filling up the empty locations. There are twice as many green pixels compared to the other two colors and is the result of the higher luminance information captured in the color green

G	B	G	B	G	B	G
R	G	R	G	R	G	R
G	B	G	B	G	B	G
R	G	R	G	R	G	R
G	B	G	B	G	B	G
R	G	R	G	R	G	R
G	B	G	B	G	B	G

Figure 1: Typical Bayer pattern used in single-CCD digital cameras.

In video compression, the reduction of temporal redundancy, obtained by motion estimation algorithm, is used to achieve high compression rates with little reduction in quality. In the conventional system, the CFA data is first interpolated into a full color image and then it is used as the input of the encoder. Obviously, this step increase the amount of data to be precessed.

As well known in literature, image data sparsely sampled through a color filter array are very sensitive to compression errors especially due to the propagation of errors caused by the interpolation. In fact, a single error on the raw data may cause a spatial propagation of the error pattern depending on the CFA pattern position. Moreover, the interpolation process is time consuming and it increases the dimension of the data without increasing the information content (entropy) of the original image. An alternative way to deal with CFA data is to directly compress it and perform the full color interpolation at the decoder.

The paper is divided as follows. In Section 2, the proposed pre-processing algorithm is described. In Section 3 a brief summary of the MPEG2-like compression method is reported. The experimental results are reported in Section 4. Finally, in Section 5, we present our conclusions.

2. THE PROPOSED APPROACH

The core of the proposed method is the reduction of redundancy created by the compression of the full color interpolated Bayer patterned frames. Figure 2 illustrates this concept.

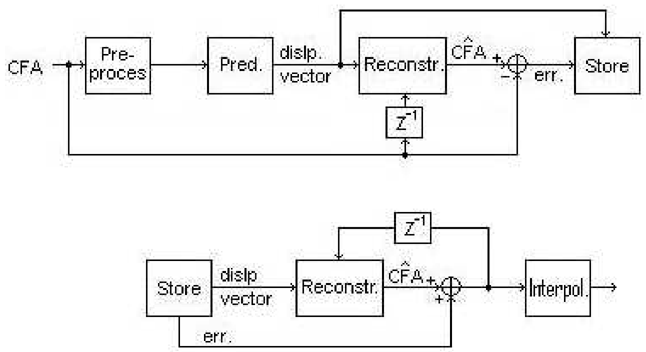

Figure 2: Proposed compression scheme.

2.1 Pre-processing

A pre-processing phase is performed to re-arrange the data for the successive compression phase. Several methods have been investigated. More in detail, we used the Bayer pattern directly, only the green component, the modification of the luminance and chrominance components as suggested by Lee et al. [3] and a modified version of *structure conversion* proposed in [4] for both the green and the luminance components.

Preliminary simulation results indicate that a modified version of the *structure conversion* method applied for green component provides the best results. Here, it is applied directly on the green component; in the following it will be denoted as SC-G. Let us consider the generic CFA frame of dimension $N \times N$. First, we select only the green quincunx component. Then it is collapsed into a compact format as illustrated in Figure 3. The SC-G method allows for the reduction of the image size from $N \times N$ to $N \times (N/2)$. The modified data, used as input to the prediction method, yields half the original motion vectors.

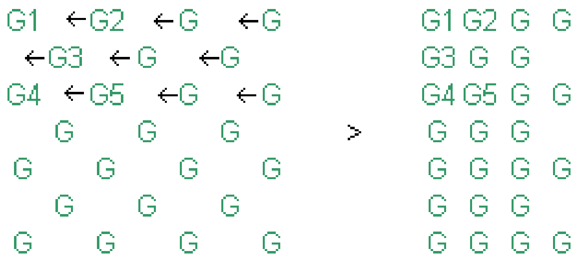

Figure 3: SC method applied for green pixels arranged in a quincunx format.

2.2 Prediction

The motion estimation method adopted is the classical block-matching [5] (method with 8×8 blocks and a 16×16 search window). We used only the green component in SC format as input resulting in the displacement vectors in SC format as shown in Figure 4.

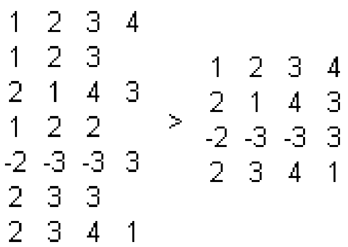

Figure 4: Example of displacement vectors in SC format and sub-sampled version.

2.3 Reconstruction

The green, blue and red component for each pixel is computed separately. However, the reconstruction of the green component is different from the blue and red components because of the non-uniform structure of the bayer pattern. In the case of the green component, reconstruction is effected using all the available displacement vectors.In the case of the blue and red components, a sub-sampled version of the displacement vectors are used, as illustrated in Figure 4. Sub-sampling is required since the displacement vectors in the odd rows do not provide valid motion information for the blue and red components. The predicted color components are then *expanded* into their original dimensions and merged to form the CFA array again.

2.4 Interpolation

In the literature, several types of interpolation algorithms have been proposed for the Bayer pattern [6]. By using the bilinear method, the three color planes are independently processed using linear averaging of the nearest neighbors of the same color. As expected, bilinear interpolation generates significant blurring artifacts. To reduce the presence of such artifacts, more complex but better performing schemes can be employed. Examples of such schemes are those proposed by Hamilton-Adams [7], Freeman [8], Laroche-Prescott [9], Chang et.al. [10], Pei-Tam [11], Kimmel [12], Cai et.al. [13] and Malvar [14]. The interpolation algorithm used in all our simulation is the algorithm proposed by Hamilton-Adams [7], which we denote as Laplacian.

3. VIDEO PROCESSING

For video processing we have used the same algorithms as in MPEG2. In MPEG2, there are three different type of frames.

- I-frames (intra) are compressed frames which contain all of the spatial information of a video frame.
- P-frames (predicted) are computed based on the nearest previous I-frame or P-frame. P-frames are more highly compressed than I-frames and provide a reference for the calculation of the B-frames (bi-directional).
- B-frames use both past and subsequent frames as a reference to calculate the compressed frame data.

We have used just one group-of-pictures (GOP), where a GOP is composed ***I-B-B-P-B-B-P-B-B-P-B-B-I*** of frames where the last I-frame belongs to the next GOP.

In the conventional method, the full color video is obtained by the application of full color interpolation on the CFA data. The next step is a color-space transformation into the luminance and chrominance components. The luminance is then used for motion compensation and the chrominance components are sent by using 4:2:0 sub-sampling. To simplify the simulation method, we have used standard JPEG on the error and the I-frame. We have also used a simple linear method to compute the B-frames.

4. SIMULATION RESULTS

4.1 CPSNR computation

The assessment of the performance of the proposed technique is based on the comparison of the decoded sequence with a full color reference video generated by a Laplacian interpolation of each original CFA frame. As videoquality metric the average Composite-Peak-Signal-to-Noise-Ratio (CPSNR), where the difference between the decoded sequence and the reference video is considered as noise,

$$CPSNR = 10\log_{10} \frac{255^2}{\frac{1}{3TMN}\sum_{t=1}^{T}\sum_{k=1}^{3}\sum_{i=1}^{N}\sum_{j=1}^{M}[I_{i,j,k}^{ref}(t) - I_{i,j,k}^{dec}(t)]^2}. \quad (1)$$

Here, $I_{i,j,k}^{dec}(t)$ is the k-th color component of pixel at the location (i, j) at time t of the decoded and then interpolated video sequence, $I_{i,j,k}^{ref}(t)$ is the corresponding value of the reference sequence, M and N are the height and the width of a frame, and T is the video duration.

4.2 Test sequences

We have employed full color QCIF (144 $\times$ 176) images in our simulations. Example frames are shown in Figure. 5).

1. **Foreman** - A talking head with the camera panning from left to right.
2. **News** - News report with two news readers and a video clip of a ballet dancer in the background. Movement is focused on the dancer in the.
3. **Carphone** - A talking head in a moving vehicle. Movement is focused on the passing scenery caused by the moving vehicle.
4. **Suzie** - A talking head where the subject is flicking her hair. Movement is focused on her head and hair.

The Bayer pattern CFA videos have been directly generated by sampling these videos appropriately as illustrated in Figure 6.

4.3 Comparison

We compared both the conventional method and the proposed SC-G method at several levels of compression by varying several levels of quality from 100 to 10. The results of the simulation are provided in Figures 7-10 where the average CPSNR computed on the 12 frames of a single GOP is plotted versus the length, in Kbytes, of the encoded bitstream.

As can be seen, the SC-G method outperforms the conventional method at low and high CPSNR values. In fact, for both low and high CPSNR values the SC-G method reach the lower bitstream (higher compression ratio). At the mid CPSNR values, the conventional method performs better.

Figure 5: Sample frames of the full color videos.

Figure 6: Bayer CFA data generated from the sample frames of the full color videos of Figure 5.

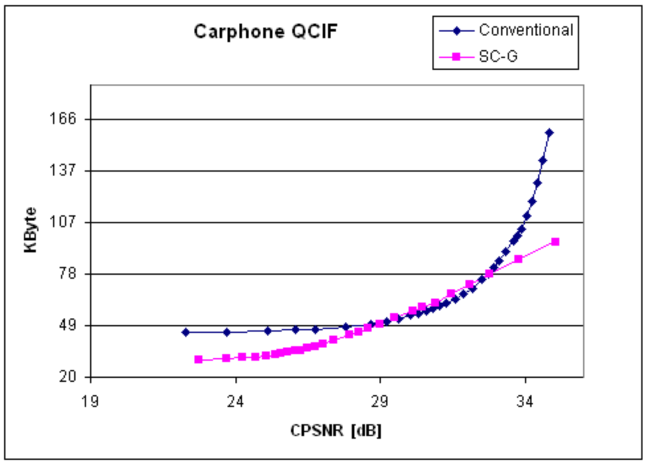

Figure 7: Performance plot for the Carphone video.

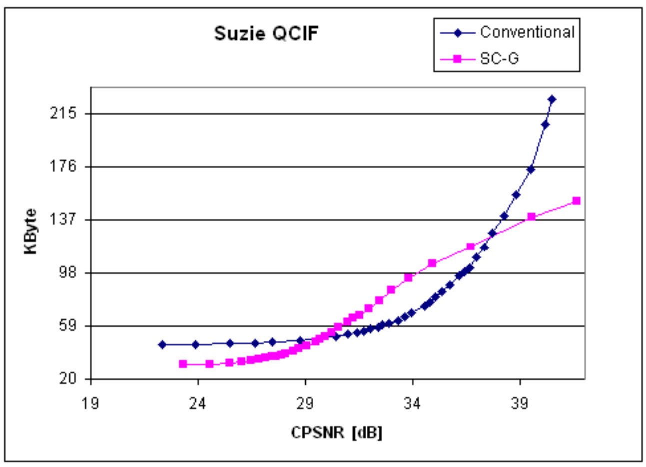

Figure 8: Performance plot for the Suzie video.

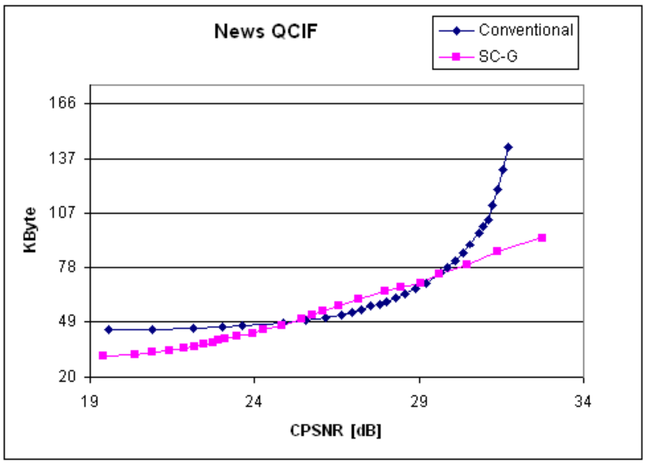

Figure 9: Performance plot for the News video.

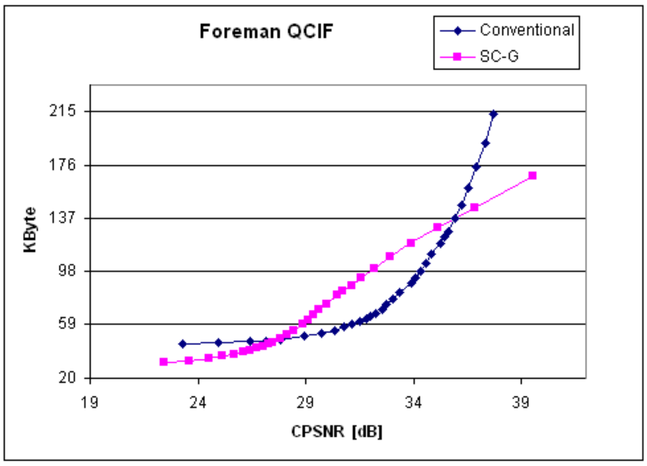

Figure 10: Performance plot for the Foreman video.

5. CONCLUDING REMARKS

In this work we have studied the feasibility of direct coding of the composite CFA video signal obtained without full color interpolation.

We exploited the advantage of working on the reduced size of the original CFA data. Since, the spatial resampling in the full color conversion process does not introduce any new information, compared to conventional systems, higher CPSNR values, can be achieved when the available bitrate allows a fine coding of the interframe prediction residuals.

In fact, although the SC-G technique makes use of only half the number of displacement vectors employed by conventional coders ($N \times M/2$ instead of $N \times M$), the quantization step along the horizontal axis is twice the pixel width. Therefore the energy of the interframe difference is usually high. However, at low and high bit rates, focusing the available resources on just one component (i.e. the Bayer pattern) allows better performance compared to solutions that have to split them among three smaller components.

Further improvements can be expected by a smarter subpixel motion compensation of the Bayer pattern. Nevertheless, the low complexity scheme descibed in this contribution allows effective bandwidth reduction when videos are transmitted over a communications link at low bitrate while maintaining the same quality provided in the conventional method.

REFERENCES

[1] K.A. Parulski, “Color filters and processing alternatives for one-chip cameras,” *IEEE Trans. Electron Devices.* Vol. ED-32, No. 8. Pp. 1381, Aug. 1985.

[2] B.E. Bayer, “Color imaging array,” *U.S. Patent No. 3,971,065*, 1976.

[3] 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*, Vol. 3, Pp. 482-485, Oct 2001.

[4] C.C. Koh, J. Mukherjee, and S.K. Mitra, “New efficient methods of image compression in digital cameras with color filter array,” *IEEE Transactions on Consumer Electronics*, Vol. 49, No. 4, Nov 2003.

[5] A.M. Tekalp, ”*Digital Video Processing*,” Prentice-Hall, Inc., Upper Saddle River, NJ, 1995.

[6] B.K. Gunturk, J. Glotzbach, Y. Altunbasak, R.W. Shafer, and R.M. Mersereau “Demosaicking: Color Filter Array Interpolation,” *IEEE Signal Processing Magazine* Vol. 22, No. 1, Jan. 2005.

[7] J.F. Hamilton and J.E. Adams, “Adaptive color plane interpolation in single sensor color electronic camera,” U.S. Patent No. 5,629,734, 1997.

[8] W.T. Freeman, “Median filter for reconstructing missing color samples,” U.S. Patent No. 4,724,395, 1988.

[9] C.A. Laroche and M.A. Prescott, “Apparatus and method for adaptively interpolating a full color image utilizing chrominance gradients,” U.S. Patent No. 5,373,322, 1994.

[10] E. Chang, S. Cheung, and D.Y. Pan, “Color filter array recovery using a threshold-based variable number of gradients,” *Proc. SPIE*, Vol. 3650, Pp. 3643, Jan 1999.

[11] S.C. Pei and I.K. Tam, “Effective color interpolation in ccd color filter array using signal correlation,” *IEEE Trans. Circuits and Systems for Video Tecnology*, Vol. 13, No. 6, Pp. 503513, Jun 2003.

[12] R. Kimmel, “Demosaicing: Image reconstruction from color ccd samples,” *IEEE Trans. Image Processing*, Vol. 8, Iss. 9. Pp. 1221-1228, Sep 1999.

[13] C. Cai ,Y-H Yu, and S.K. Mitra, “Saturation-based adaptive inverse gradient interpolation for Bayer pattern images,” *IEEE Proceedings - Vision, Image and Signal Processing*, Vol. 148, June 2001, Pp. 202-208.

[14] H.S. Malvar, L.W. He, and R. Cutler, “High-quality linear interpolation for demosaicing of bayer-patterned color images,” *Proc. ICIP*, Vol. 3, Pp. 485-488, May 2004.