Novel Chroma Subsampling Strategy Based on Mathematical Optimization for Compressing Mosaic Videos With Arbitrary RGB Color Filter Arrays in H.264/AVC and HEVC

Chien-Hsiung Lin, Kuo-Liang Chung, *Senior Member, IEEE*, and Chun-Wei Yu

***Abstract*— To save manufacturing cost, most color digital video cameras employ a single-sensor technology with a red–green–blue (RGB) color filter array (CFA) to capture real-world scenes. Due to only one primary color measured at each pixel location, the captured videos are usually referred to as the mosaic videos. For the purposes of economical storage and transmission, it is very important to achieve a good tradeoff between the quality and bitrate when compressing mosaic videos with different RGB-CFA structures. In this paper, based on mathematical optimization technique, a novel chroma subsampling strategy is presented for compressing mosaic videos with arbitrary RGB-CFA structures in H.264/AVC and High Efficiency Video Coding (HEVC). For each 2 × 2 YUV block to be subsampled with 4:2:0 format, the proposed strategy determines the proper sampled *U* and *V* components by minimizing, prior to compression, the quality distortion between the original colocated mosaic block and the mosaic block conversed from the current subsampled YUV block. Through the mathematical optimization formulated in the proposed strategy, the significance of the sampled *U* and *V* components for reconstructing R, G, and B pixels can be simultaneously taken into consideration. The experimental results demonstrate that the proposed chroma subsampling strategy has the best quality and bitrate tradeoff at a similar execution time requirement for compressing mosaic videos with arbitrary RGB-CFA structures in H.264/AVC and HEVC compared with the state-of-the-art ones by Chen *et al.* and Yang *et al.* as well as the three commonly used ones.**

***Index Terms*— Arbitrary red–green–blue (RGB) color filter arrays (CFAs), Bjøntegaard delta peak signal-to-noise ratio (BD-PSNR), chroma subsampling, H.264/AVC, High Efficiency Video Coding (HEVC), mathematical optimization, mosaic videos.**

I. Introduction

A FULL-COLOR video usually consists of three primary red–green–blue (RGB) color channels and accordingly, three independent sensors per pixel location, each for acquiring a specific primary color, are demanded in a color digital video camera. To save manufacturing cost, most color digital video cameras equip only a single sensor covered with an RGB color filter array (CFA) per pixel location to capture real-world scenes. Since each pixel in the image frames of the acquired video has only one of three primary RGB color components, this kind of video is usually referred to as a mosaic video. For a mosaic video, the information about its RGB-CFA structure, which means the arrangement of the RGB color filters in the CFA, is quite important, since with this information, the demosaicking process [5], [6], [10], [14], [17]–[20], [23], [25]–[27] can convert the acquired mosaic video into the full-color video. In general, the RGB-CFA structure information can be obtained from the camera manufacturer or the header of the mosaic video stored in TIFF-EP format. Seven typical types of RGB-CFA structures [21] are shown in Fig. 1, among which the Bayer CFA [1] is currently the most popular one.

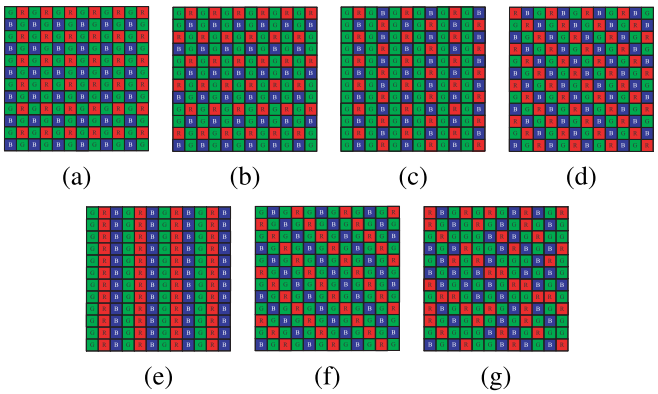


Fig. 1. Seven typical RGB-CFA structures. (a) Bayer CFA. (b) Lukac and Plataniotis CFA. (c) Yamanaka CFA. (d) Diagonal stripe CFA. (e) Vertical stripe CFA. (f) Modified Bayer CFA. (g) HVS-based CFA.

A digital video camera is often a portable recording device driven by a battery power and with limited capacity for data storage and transmission. After captured by the camera, the videos are often conveyed from the camera to another video manipulation environment with greater computing resources and then are enhanced by postprocessing operations such as lens aberration correction, color correction, exposure and white balance adjustment, dead pixel removal, tone mapping, and

Manuscript received October 17, 2014; revised March 18, 2015; accepted April 27, 2015. Date of publication August 24, 2015; date of current version September 1, 2016. This work was supported by the Ministry of Science and Technology, Taiwan, under Contract MOST 101-2221-E-011-139-MY3 and Contract MOST 102-2221-E-011-055-MY3. This paper was recommended by Associate Editor M. Rabbani. *(Corresponding author: Kuo-Liang Chung.)*

The authors are with the Department of Computer Science and Information Engineering, National Taiwan University of Science and Technology, Taipei 10672, Taiwan (e-mail: d9409301@mail.ntust.edu.tw; klchung01@gmail.com; cwyu1057@gmail.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCSVT.2015.2472118

so on. For single-sensor digital video cameras aiming at low manufacturing cost, in order to maximize their utility under the circumstance of their limited storage space and transmission capacity, compressing the captured mosaic videos prior to storing or transmitting the videos while preserving good quality of the reconstructed videos for the subsequent postprocessing operations is necessary and rather desired in practice. During the past decades, several compression schemes for mosaic videos have been proposed and can be roughly divided into two categories: 1) compression-first and 2) demosaicking-first mosaic video compression schemes. Compression-first mosaic video compression schemes, which first convert the CFA structure by color separation, geometric adjustment, or color domain transformation and then directly carry out compression without demosaicking for achieving better compression efficiency, are often applied to mosaic videos with the Bayer CFA since the Bayer CFA has 2:1:1 (G:B:R) color component composition in each 2×2 subimage, which is similar to the 4:2:2 subsampling format used by the encoder. Lee and Ortega [16] exploited the transformation from RGB color domain to YCbCr color domain together with the geometric rotation technique to directly compress mosaic images with the Bayer CFA in the JPEG environment. Malvar and Sullivan [22] proposed a low-complexity integer-reversible spectral–spatial color domain transformation for mosaic images with the Bayer or Bayer-like CFA by which the concerned mosaic images can directly be compressed with JPEG in a lossless, lossy, or progressive-to-lossless manner. Thanks to only intra-mode encoding supported by most video coders, the above two compression schemes originally designed for compressing mosaic images can also be applied to mosaic video compression. Gastaldi *et al.* [11] presented the first structure-conversion-based coder for compressing mosaic videos with the Bayer CFA under the MPEG-2 environment [12]. Later, due to better compression efficiency achieved by the H.264/AVC video coder [9], Doutre *et al.* [7] incorporated the structure conversion technique in [11] with the prediction schemes in H.264/AVC to deliver better compression results for mosaic videos with the Bayer CFA. Doutre and Nasiopoulos [8] further improved, according to the characteristic of the Bayer CFA, the intra-prediction scheme in H.264/AVC to enhance the quality of the reconstructed mosaic video. Unfortunately, due to the difficulty in converting non-Bayer-like CFA structures, the compression-first mosaic video compression schemes cannot be applied directly to mosaic videos with non-Bayer-like CFA structures.

As a remedy, demosaicking-first mosaic video compression schemes first perform demosaicking on the acquired mosaic video and then apply a regular compression process consisting of the RGB to YUV color domain transformation, chroma subsampling, and video coding to the recovered full-color video such that they are applicable to mosaic videos with arbitrary RGB-CFA structures. The simplest demosaicking-first mosaic video compression scheme is to adopt one demosaicking technique designed for one specific RGB-CFA structure or one universal demosaicking technique designed for arbitrary RGB-CFA structures together with one of the commonly used chroma subsampling strategies originally designed for full-color videos. However, the commonly used chroma subsampling strategies designed for full-color videos always sample the U and V components from the fixed positions and without considering the characteristics of mosaic videos, resulting in inferior quality of the reconstructed mosaic videos. To alleviate this problem, Chen *et al.* [4] and Yang *et al.* [27] proposed the chroma subsampling strategies specially designed for compressing mosaic videos with the Bayer CFA and with arbitrary RGB-CFA structures in H.264/AVC, respectively. In their chroma subsampling strategies, by considering the significance of the sampled U components for reconstructing B pixels as well as the sampled V components for reconstructing R pixels, the positions of the sampled U and V components are dynamically adjusted according to the corresponding RGB-CFA structure. Specifically, Yang *et al.*'s chroma subsampling strategy has the same compression performance as Chen *et al.*'s one for mosaic videos with the Bayer CFA, and yields better compression performance over the commonly used ones for mosaic videos with arbitrary RGB-CFA structures. However, since usually the majority of the pixel locations in the RGB-CFA structure equip G color filters, not considering the significance of the sampled U and V components for reconstructing G pixels in their chroma subsampling strategies leads to a limited quality improvement for the reconstructed mosaic videos.

With the continuous growth of video resolution, the H.264/AVC standard can no longer deliver satisfied compression results. The emerging video coding standard, called High Efficiency Video Coding (HEVC), aims to not only improve the quality of the reconstructed videos but also double the compression ratio, compared with the H.264/AVC standard, with a limited increase in computational complexity. The HEVC project spent much effort in developing advanced prediction modes for better luma/chroma intra/inter predictions where the luma-predict-chroma mode has been shown to provide significant coding gains, revealing that in HEVC, fully using the YUV format as an intermediary for compression would benefit more than in H.264/AVC and hence this brings a potential advantage for developing a demosaicking-first mosaic video compression scheme in HEVC.

In this paper, based on mathematical optimization, a novel chroma subsampling strategy is proposed for compressing mosaic videos with arbitrary RGB-CFA structures in H.264/AVC and HEVC [3]. In the proposed chroma subsampling strategy, for each 2×2 YUV block to be subsampled with 4:2:0 format, the proper sampled U and V components are determined by minimizing, prior to compression, the quality distortion between the original colocated mosaic block and the mosaic block conversed from the current subsampled YUV block. Different from overfavoring the R and B pixels' reconstruction when determining the sampled U and V components in Chen *et al.*'s and Yang *et al.*'s chroma subsampling strategies, the significance of the sampled U and V components for reconstructing R, G, and B pixels is simultaneously taken into consideration through the mathematical optimization formulated in the proposed chroma subsampling strategy. The experimental results demonstrate that the proposed chroma subsampling strategy has the best

quality and bitrate tradeoff at a similar computational time requirement for compressing mosaic videos with arbitrary RGB-CFA structures in H.264/AVC and HEVC compared with the state-of-the-art ones by Chen *et al.* and Yang *et al.* as well as the commonly used ones.

The remainder of this paper is organized as follows. In Section II, we review the demosaicking-first mosaic video compression schemes and the related chroma subsampling strategies. In Section III, we present the proposed novel chroma subsampling strategy based on mathematical optimization for compressing mosaic videos with arbitrary RGB-CFA structures in H.264/AVC and HEVC. Based on six test videos with the seven typical RGB-CFA structures, Section IV gives the empirical results to justify the quality superiority of the proposed chroma subsampling strategy for compressing mosaic videos. The concluding remarks are addressed in Section V.

II. Demosaicking-First Mosaic Video Compression Schemes and Related Chroma Subsampling Strategies

This section begins with a brief description of the demosaicking-first compression schemes for mosaic videos with arbitrary RGB-CFA structures in H.264/AVC and HEVC and then points out the drawbacks of the related chroma subsampling strategies.

A. Demosaicking-First Mosaic Video Compression Schemes

For each mosaic image I_M in the captured mosaic video with an arbitrary RGB-CFA structure, the demosaicking-first mosaic video compression schemes, at the encoder side, first demosaick I_M by a demosaicking technique designed for the used RGB-CFA structure or a universal demosaicking technique designed for arbitrary RGB-CFA structures to obtain the demosaicked full-color RGB image. Then, the transformation from the RGB color domain to the YUV color domain defined as

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.257 & 0.504 & 0.098 \\ -0.148 & -0.291 & 0.439 \\ 0.439 & -0.368 & -0.071 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} \quad (1)$$

is performed on the demosaicked full-color RGB image so as to obtain the corresponding YUV image I_{YUV}. Next, a 4:2:0 chroma subsampling strategy is applied to the U and V components in I_{YUV} such that for each 2×2 block in I_{YUV}, only one sampled U component and V component are kept, as shown in Fig. 2 where Y_k, U_k, and V_k, $1 \leq k \leq 4$, denote, respectively, the Y, U, and V components of the pixel at position k in which k indicates the row-major index of the current block. Finally, the resultant subsampled YUV image is conveyed to the video encoder, e.g., the H.264/AVC encoder or the HEVC encoder, for compression with 4:2:0 format.

At the decoder side, when reconstructing the compressed mosaic video, the video decoder first reconstructs the subsampled YUV images with the specified 4:2:0 format and then duplicates, respectively, the sampled U and V components as the discarded U and V components for all the pixels of each 2×2 block in the reconstructed subsampled YUV images. Next, the transformation from the YUV domain to the RGB domain defined as

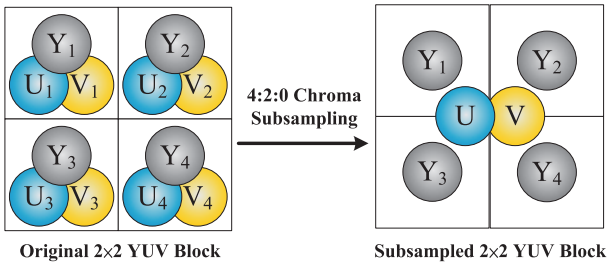


Fig. 2. Results for one 2×2 YUV block before and after the 4:2:0 chroma subsampling.

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1.164 & 0 & 1.596 \\ 1.164 & -0.391 & -0.813 \\ 1.164 & 2.018 & 0 \end{bmatrix} \begin{bmatrix} Y - 16 \\ U - 128 \\ V - 128 \end{bmatrix} \quad (2)$$

is performed on the reconstructed YUV images to obtain the reconstructed RGB images. Finally, the reconstructed mosaic video is delivered by mosaicking these reconstructed RGB images with the associated RGB-CFA structure.

B. Related Chroma Subsampling Strategies and Their Drawbacks

In the existing full-color video compression schemes, there exist three commonly used 4:2:0 chroma subsampling strategies, denoted by 4:2:0(A), 4:2:0(L), and 4:2:0(R), respectively. Although the three chroma subsampling strategies are originally designed for tackling the YUV videos associated with the full-color videos, since in the above mosaic video compression procedure, the input mosaic video has been transformed into a regular YUV video before chroma subsampling, the three chroma subsampling strategies still can be used as the chroma subsampling strategy required for the demosaicking-first mosaic video compression scheme. The 4:2:0(A) determines the one sampled U and V components associated with each 2×2 YUV block, respectively, by averaging the four U and V components in the block. The 4:2:0(L) determines the one sampled U and V components associated with each 2×2 YUV block, respectively, by averaging the U and V components in the left column of the block. As for the 4:2:2(R), its determination process for the one sampled U and V components is similar to that of the 4:2:2(L) but chooses the U and V components in the right column of the block.

However, since the three commonly used chroma subsampling strategies designed for full-color videos always sample the U and V components from the fixed positions and without taking the characteristics of mosaic videos into account, compressing mosaic videos using any of the three chroma subsampling strategies yields inferior quality of the reconstructed mosaic videos. To further enhance the quality of the reconstructed mosaic videos, Chen *et al.* [4] and Yang *et al.* [27] proposed the chroma subsampling strategies specially designed for compressing mosaic videos with the Bayer CFA and with arbitrary RGB-CFA

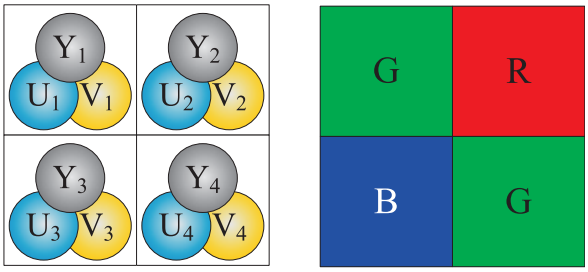


Fig. 3. (a) 2×2 YUV block. (b) Colocated mosaic block.

structures, respectively. From (2), Chen *et al.* and Yang *et al.* observed the following.

1) Only reconstructing the B and G pixels involves the U component and only reconstructing the R and G pixels involves the V component.
2) The coefficients of the U component in (2) show that the U component possesses more impact on reconstructing the B pixel than on reconstructing the G pixel.
3) The coefficients of the V component in (2) reveal that the V component possesses more impact on reconstructing the R pixel than on reconstructing the G pixel.

Accordingly, for better reconstructing the compressed mosaic video, the one sampled U component for each 2×2 YUV block at the encoder side should preferentially be determined as the average of the U components of the pixels that correspond to the B pixels in the colocated mosaic block. Similarly, the one sampled V component for each 2×2 YUV block should preferentially be determined as the average of the V components of the pixels that correspond to the R pixels in the colocated mosaic block. For example, considering a 2×2 YUV block and its colocated mosaic block shown in Fig. 3, Chen *et al.*'s and Yang *et al.*'s chroma subsampling strategies choose U_3 and V_2 to be the sampled U and V components of this YUV block.

Although Chen *et al.*'s and Yang *et al.*'s chroma subsampling strategies exploit the significance of the sampled U components for reconstructing B pixels as well as the sampled V components for reconstructing R pixels so as to yield better quality of the reconstructed mosaic videos, since usually the majority of the pixel locations in the RGB-CFA structure equip G color filters to capture color information, not considering the significance of the sampled U and V components for reconstructing G pixels in their chroma subsampling strategies eventually leads to a limited quality improvement. Thus, in the following section, we propose a novel chroma subsampling strategy based on mathematical optimization for compressing mosaic videos with arbitrary RGB-CFA structures in H.264/AVC and HEVC, which can simultaneously consider the significance of the sampled U and V components for reconstructing R, G, and B pixels and hence superior quality of the reconstructed mosaic videos can be expected.

III. Proposed Chroma Subsampling Strategy Based on Mathematical Optimization

Given an image frame I_M in the input mosaic video with an arbitrary RGB-CFA structure, after the demosaicking and the transformation from the RGB color domain to the YUV color domain as described earlier, we have the associated YUV image frame I_{YUV}. Considering a 2×2 block in I_{YUV}, for the pixel at position k of the current YUV block, where k, $1 \le k \le 4$, indicates the row-major index of the block, let C_k denote the type of RGB color filter used by the pixel at the same position of the colocated 2×2 block in I_M. For example, we have $C_1 = G$, $C_2 = R$, $C_3 = B$, and $C_4 = G$ for the 2×2 YUV block and its colocated mosaic block shown in Fig. 3.

Since the 4:2:0 chroma subsampling operation retains only one sampled U and V components for the current 2×2 YUV block, the information loss due to discarding partial U and V components turns out that the subsampled YUV block cannot be losslessly conversed to the original mosaic block prior to compression, implying that there must exist quality distortion between the original mosaic block and the mosaic block conversed from the subsampled YUV block after the chroma subsampling and before the compression. Intuitively, if this quality distortion is large, then when conveying such a subsampled YUV block for compression, further accompanied by compression distortion, the quality of the reconstructed mosaic video would decay more seriously. Thus, how to determine one proper sampled U and V components such that this quality distortion can be effectively reduced and then the quality of the reconstructed mosaic video can be upgraded provides a direction to develop a better 4:2:0 chroma subsampling strategy for compressing mosaic videos.

Let U and V denote the sampled U and V components to be determined for the current 2×2 YUV block. Since after U and V are determined, U and V will be used as the substitutes for the original U and V components of each pixel of the current YUV block, the mosaic block conversed from the subsampled YUV block must be different from the original mosaic block and the quality distortion between them can be defined as

$$D(U, V) = \sum_{k=1}^{4} d_k^2 \tag{3}$$

with d_k, $1 \le k \le 4$, denoting the difference between the color value of the pixel at position k of the original 2×2 mosaic block and the color value of the pixel at the same position of the conversed mosaic block. From the transformation from the YUV color domain to the RGB color domain shown in (2), we observe that the variation in values of the U and V components have different impacts on the variation in values of the R, G, and B components and such a relationship can be expressed as

$$\begin{bmatrix} \Delta R \\ \Delta G \\ \Delta B \end{bmatrix} = \begin{bmatrix} 0 & 1.596 \\ -0.391 & -0.813 \\ 2.018 & 0 \end{bmatrix} \begin{bmatrix} \Delta U \\ \Delta V \end{bmatrix} \tag{4}$$

where ΔR, ΔG, ΔB, ΔU, and ΔV represent the variations in values of the R, G, B, U, and V components, respectively. For the pixel at position k of the current YUV block, since its original U and V components, U_k and V_k, are replaced by U and V after the chroma subsampling, according to (4), the difference d_k in (3) resulting from the variations

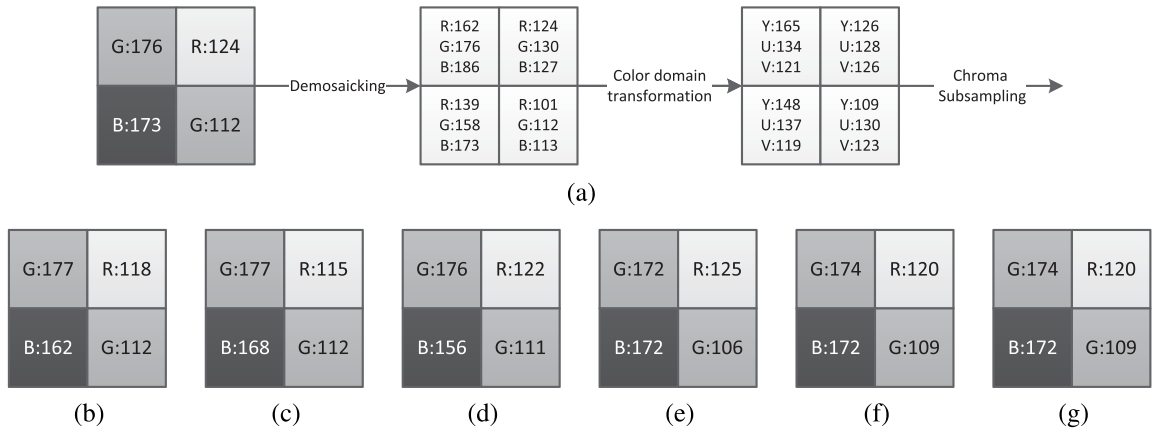


Fig. 4. Example for performing the concerned chroma subsampling strategies on a real 2×2 mosaic block. (a) Original 2×2 mosaic block, its demosaicking, and RGB-to-YUV color transformation. (b)–(f) Conversed mosaic blocks corresponding to the 4:2:0(A) strategy, the 4:2:0(L) strategy, the 4:2:0(R) strategy, Chen *et al.*'s and Yang *et al.*'s strategies, the exhaustive search-based strategy, and the proposed strategy, respectively.

$\Delta U = (U - U_k)$ and $\Delta V = (V - V_k)$ can be given by

$$d_k = a_k(U - U_k) + b_k(V - V_k) \tag{5}$$

where

$$a_k = \begin{cases} 0, & \text{if } C_k = R \\ -0.391, & \text{if } C_k = G \\ 2.018, & \text{if } C_k = B \end{cases} \tag{6}$$

and

$$b_k = \begin{cases} 1.596, & \text{if } C_k = R \\ -0.813, & \text{if } C_k = G \\ 0, & \text{if } C_k = B. \end{cases} \tag{7}$$

According to the sampled U and V, prior to compression, the resulting quality distortion between the original mosaic block and the mosaic block conversed from the subsampled YUV block can be rewritten as

$$D(U, V) = \sum_{k=1}^{4} [a_k(U - U_k) + b_k(V - V_k)]^2. \tag{8}$$

The above equation states that for the current 2×2 YUV block, when the sampled U and V components are determined as U and V, the resulting quality distortion prior to compression will be $D(U, V)$. To upgrade the quality of the reconstructed mosaic video by effectively reducing the above quality distortion, we propose that the determination of U and V should make $D(U, V)$ minimized. Thus, in the proposed chroma subsampling strategy, the way to determine U and V is modeled as the following mathematical optimization problem:

$$\arg \min_{0 \le U, V \le 255} D(U, V). \tag{9}$$

However, since U and V are discrete, obtaining the true solution to the above optimization problem requires an exhaustive search over all possible values of U and V, which will significantly increase the computational complexity. Instead, we first set the first derivative of the above optimization problem with respect to U and V to zero to yield

$$\frac{\partial D(U, V)}{\partial U} = \sum_{k=1}^{4} 2a_k[a_k(U - U_k) + b_k(V - V_k)] = 0 \tag{10}$$

$$\frac{\partial D(U, V)}{\partial V} = \sum_{k=1}^{4} 2b_k[a_k(U - U_k) + b_k(V - V_k)] = 0. \tag{11}$$

By solving this linear system, the approximated solution $(\hat{U}, \hat{V})$ to the above optimization problem can be obtained as (12) and (13), as shown at the bottom of this page. Then, in considerations of the discrete characteristic of U and V as well as the truncation used in the transformation between the RGB and YUV color domains, we thus determine the sampled U and V components for the current YUV block by

$$(U^*, V^*) = \arg \min_{\substack{\lfloor \hat{U} \rfloor \le U \le \lceil \hat{U} \rceil \\ \lfloor \hat{V} \rfloor \le V \le \lceil \hat{V} \rceil}} D(U, V) \tag{14}$$

where $\lfloor x \rfloor$ and $\lceil x \rceil$ represent the floor function of x and the ceil function of x, respectively.

We now take a real 2×2 block from the *Houses* video with the Bayer CFA as an example to demonstrate the quality superiority of the proposed chroma subsampling strategy among the concerned chroma subsampling strategies. Fig. 4(a) shows the original 2×2 mosaic block, the corresponding demosaicking, and the RGB-to-YUV color transformation before the chroma subsampling operation. For the current YUV block to be subsampled, after performing

$$\hat{U} = \frac{\left(\sum_{k=1}^{4} b_k^2\right) \cdot \left(\sum_{k=1}^{4} a_k^2 U_k + a_k b_k V_k\right) - \left(\sum_{k=1}^{4} a_k b_k\right) \cdot \left(\sum_{k=1}^{4} a_k b_k U_k + b_k^2 V_k\right)}{\left(\sum_{k=1}^{4} a_k^2\right) \cdot \left(\sum_{k=1}^{4} b_k^2\right) - \left(\sum_{k=1}^{4} a_k b_k\right)^2} \tag{12}$$

$$\hat{V} = \frac{\left(\sum_{k=1}^{4} a_k b_k\right) \cdot \left(\sum_{k=1}^{4} a_k^2 U_k + a_k b_k V_k\right) - \left(\sum_{k=1}^{4} a_k^2\right) \cdot \left(\sum_{k=1}^{4} a_k b_k U_k + b_k^2 V_k\right)}{\left(\sum_{k=1}^{4} a_k b_k\right)^2 - \left(\sum_{k=1}^{4} a_k^2\right) \cdot \left(\sum_{k=1}^{4} b_k^2\right)} \tag{13}$$

the concerned chroma subsampling strategies as described in Sections II-B and III on it, we have, respectively, $U_s = 132$ and $V_s = 122$ for the 4:2:0(A) strategy, $U_s = 135$ and $V_s = 120$ for the 4:2:0(L) strategy, $U_s = 129$ and $V_s = 124$ for the 4:2:0(R) strategy, $U_s = 137$ and $V_s = 126$ for Chen *et al.*'s and Yang *et al.*'s strategies, and $U_s = 137$ and $V_s = 123$ for the strategy based on an exhaustive search and the proposed strategy, where U_s and V_s denote the sampled U and V components for the current YUV block. After performing the related inverse operations, the conversed mosaic blocks corresponding to the concerned chroma subsampling strategies are shown in Fig. 4(b)–(f). Comparing the six conversed mosaic blocks with the original one, it is clear that prior to compression, the proposed strategy and the exhaustive search-based strategy have the least quality distortion, i.e., $D(U, V) = 30$, whereas Chen *et al.*'s and Yang *et al.*'s strategies, the 4:2:0(A) strategy, the 4:2:0(L) strategy, and the 4:2:0(R) strategy result in more quality distortion, i.e., $D(U, V) = 54$, $D(U, V) = 158$, $D(U, V) = 107$, and $D(U, V) = 294$, respectively. Note that Yang *et al.*'s chroma subsampling strategy has the same quality performance as Chen *et al.*'s one for mosaic videos with the Bayer CFA.

After applying the proposed chroma subsampling strategy to all the 2×2 YUV blocks in the image frames of the YUV video for the determination of their sampled U and V components, the subsampled YUV images are then conveyed to the video encoder for compression in 4:2:0 format. Since the proposed chroma subsampling strategy aims to deliver the subsampled YUV video with minimized quality distortion prior to compression by determining the proper sampled U and V components, the quality of the reconstructed mosaic video using the proposed chroma subsampling strategy can be expected to decay less compared with those using the chroma subsampling strategies described in Section II-B.

IV. Experimental Results

In this section, we present some experimental results to show that the proposed chroma subsampling strategy based on mathematical optimization has better quality and bitrate tradeoff for compressing mosaic videos with arbitrary RGB-CFA structures in H.264/AVC and HEVC compared with the state-of-the-art ones by Chen *et al.* [4] and Yang *et al.* [27] as well as the three commonly used ones, 4:2:0(A), 4:2:0(L), and 4:2:0(R). Although Chen *et al.*'s and Yang *et al.*'s chroma subsampling strategies are at first developed under the H.264/AVC environment, with the release of the HEVC standard, the experiments of applying the two strategies for compressing mosaic videos in HEVC are also performed for comparison. Specifically, since Chen *et al.*'s chroma subsampling strategy has the same compression performance as Yang *et al.*'s one for mosaic videos with the Bayer CFA, only the results of Yang *et al.*'s strategy are reported. In addition, to justify the quality superiority of the proposed strategy, the chroma subsampling strategy based on an exhaustive search over all possible values of the sampled U and V components as described in Section III is implemented for comparison and is named ideal in the comparison results.

Before conducting the experiments, the following three steps are carried out to generate the required test mosaic videos. First, we use the 24 still images in Kodak collection [29] and the four real-world videos, namely, *Foreman*, *Tempete*, *Paris*, and *Mother and Daughter*, to generate the 28 full-color videos of which each is with two hundred 352×288 RGB image frames normalized to 8-bit per color channel. Next, since most cameras in practice employ lens design with a bit of intentional blurring to reduce aliasing due to the RGB-CFA subsampling, in order to fit in the real camera-capturing process, small amounts of Gaussian lowpass filtering are applied to the 28 full-color videos before the RGB-CFA subsampling. Finally, we subsample each of the 28 blurred full-color videos with the seven RGB-CFA structures shown in Fig. 1, respectively, to generate the required test mosaic videos. Here, we classify these mosaic videos into two sets, namely, Set1 and Set2, where the former includes the mosaic videos generated from the full-color videos regarding the 24 still images in Kodak collection and the latter contains those generated from the full-color videos regarding the four real-world videos. We then perform, in H.264/AVC and HEVC, the compression and reconstruction processes accompanied with the six concerned chroma subsampling strategies on all the mosaic videos in Set1 and Set2. Note that in the experiments, we adopt the universal demosaicking technique in [27] to recover the full RGB colors for the mosaic videos.

All the experiments are implemented on a computer with an Intel Core i7-3770 CPU 3.4 GHz and 4-GB RAM. The operating system is Microsoft Windows 7 64-bit operating system. The program development environment is Visual C++ 2010. The H.264/AVC and HEVC reference softwares used for video compression are JM-18.4 and HM-13.0-RExt-6.0, respectively. In addition, the GOP size is set to 8, the GOP structure adopts the random access main profile, and the ten different quantization parameters (QPs) considered in compression are 4, 8, 12, 16, 20, 24, 28, 32, and 36.

A. Comparison in Terms of Quality and Bitrate Tradeoff

In the evaluation of quality and bitrate tradeoff, the quality of the reconstructed mosaic video and the bitrate required for the compressed video are the two main performance measures for comparison. Due to the purpose of display, the mosaic video often needs to be converted into the full-color video, so the quality of the reconstructed full-color video is also important and should be addressed. Here, we adopt the peak signal-to-noise ratio (PSNR) to measure the quality of the reconstructed mosaic video. Let $\mathbb{P} = \{(m, n) | 1 \leq m \leq H, 1 \leq n \leq W\}$ denote the set of pixel coordinates in one image frame of size $W \times H$. The PSNR of the reconstructed mosaic video with N image frames is expressed as

$$\mathrm{PSNR} = \frac{1}{N} \sum_{n=1}^{N} 10 \log_{10} \frac{255^2}{\frac{1}{\mathrm{WH}} \sum_{p \in \mathbb{P}} \left[I_M^n(p) - \widetilde{I}_M^n(p) \right]^2} \quad (15)$$

where $I_M^n(p)$ denotes the color value of the pixel at position p in the nth image frame of the original mosaic video and $\widetilde{I}_M^n(p)$ the reconstructed analogue. To evaluate the quality of

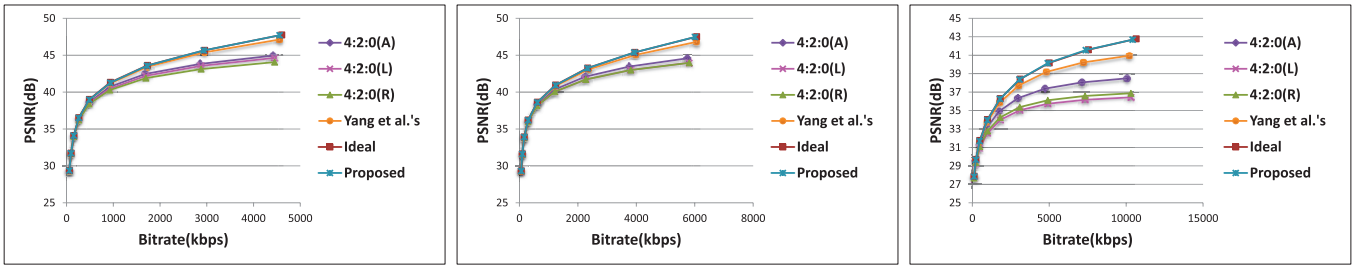


Fig. 5. RD curves of PSNR against bitrate corresponding to the six concerned 4:2:0 chroma subsampling strategies for the mosaic videos with different RGB-CFA structures in H.264/AVC. (a) Bayer CFA. (b) Lukac and Plataniotis CFA. (c) Vertical stripe CFA.

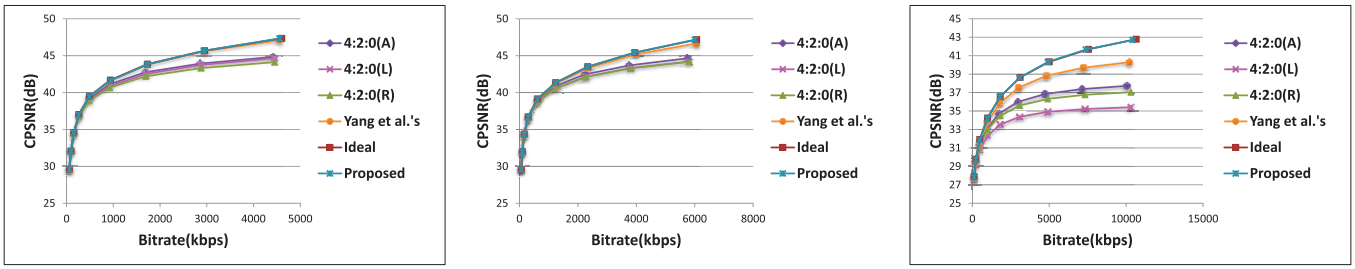


Fig. 6. RD curves of CPSNR against bitrate corresponding to the six concerned 4:2:0 chroma subsampling strategies for the mosaic videos with different RGB-CFA structures in H.264/AVC. (a) Bayer CFA. (b) Lukac and Plataniotis CFA. (c) Vertical stripe CFA.

the reconstructed full-color video, following the same practice in [4], we first demosaick the original mosaic video and the reconstructed mosaic video into full-color videos and then calculate the color PSNR (CPSNR) between them. The CPSNR of the reconstructed full-color video is expressed as

$$\text{CPSNR} = \frac{1}{N}\sum_{n=1}^{N} 10\log_{10}\frac{255^2}{\text{CMSE}} \tag{16}$$

with

$$\text{CMSE} = \frac{1}{3\text{WH}}\sum_{p\in\mathbb{P}}\sum_{C\in\{R,G,B\}}\left[I_{RGB}^{n,C}(p) - \widetilde{I}_{RGB}^{n,C}(p)\right]^2 \tag{17}$$

where $I_{RGB}^{n,C}(p)$ and $\widetilde{I}_{RGB}^{n,C}(p)$ denote the $C \in \{R, G, B\}$ color value of the pixel at position p in the nth image frame of the original and reconstructed demosaicked full-color videos, respectively. Besides the above PSNR and CPSNR, the video quality metric (VQM) proposed by the NTIA [24], which has been standardized in Recommendation ITU-T J.144 [15] as a normative full-reference video quality assessment model, is also adopted for evaluating the quality of the reconstructed full-color video. Note that here we use the NTIA general model and the full reference calibration provided in the batch VQM software 2.0 [28] to measure the values of VQM for all the reconstructed full-color videos. Moreover, given the frame rate F, the bitrate of the compressed video is defined as

$$\text{Bitrate} = \frac{T}{N} \times F \tag{18}$$

where T represents the total number of bits used to compress the mosaic video with N frames. In general, higher values of PSNR and CPSNR indicate, respectively, better quality of the reconstructed mosaic video and the reconstructed full-color video, lower values of VQM imply better quality of the reconstructed full-color video because the VQM measures the video distortion, and lower values of bitrate reflect less storage requirement for the compressed video.

Based on the average values of PSNR, CPSNR, and bitrate over Set1 and Set2 under different QP circumstances, the rate–distortion (RD) curves for different chroma subsampling strategies and different RGB-CFA structures are plotted to illustrate the overall performance of the six concerned chroma subsampling strategies. Due to lack of space, only the RD curves corresponding to compressing the mosaic videos with the Bayer CFA, the Lukac and Plataniotis CFA, and the vertical stripe CFA in H.264/AVC and HEVC are given in Figs. 5–8, respectively. The RD curves corresponding to mosaic videos with the other four RGB-CFA structures can be found in [30]. From Figs. 5–8, it is obvious that in H.264/AVC and HEVC, the proposed novel chroma subsampling strategy based on mathematical optimization has almost the same quality and bitrate tradeoff as the ideal chroma subsampling strategy and outperforms all the other concerned chroma subsampling strategies. Similar results are observed in the RD curves corresponding to mosaic videos with the other four RGB-CFA structures. In particular, for the mosaic videos with the vertical stripe CFA, the improvement of quality and bitrate tradeoff by the proposed chroma subsampling strategy is most significant.

Besides the RD curves, to provide quantitative comparison results, we further employ the Bjøntegaard delta PSNR (BD-PSNR) [2] to indicate the average PSNR

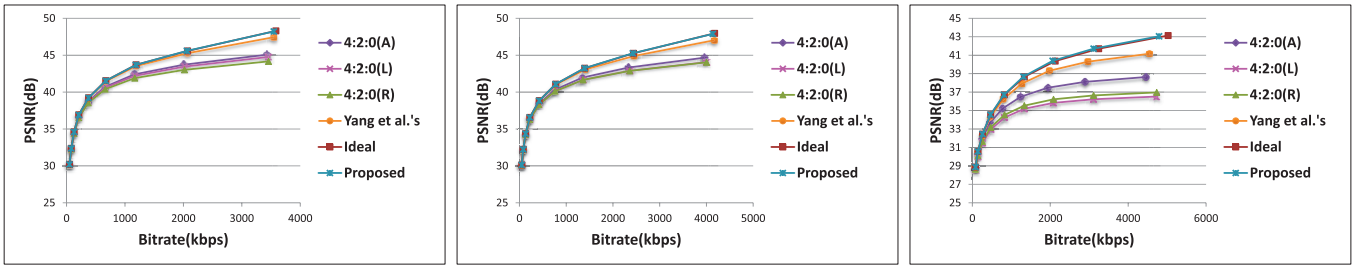


Fig. 7. RD curves of PSNR against bitrate corresponding to the six concerned 4:2:0 chroma subsampling strategies for the mosaic videos with different RGB-CFA structures in HEVC. (a) Bayer CFA. (b) Lukac and Plataniotis CFA. (c) Vertical stripe CFA.

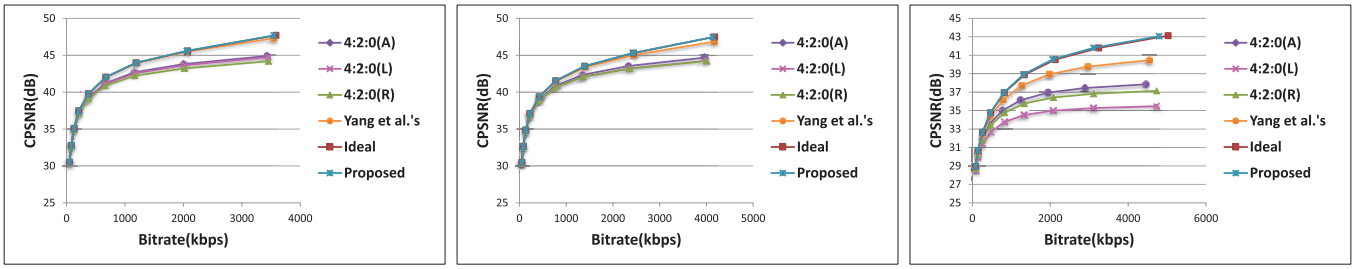


Fig. 8. RD curves of CPSNR against bitrate corresponding to the six concerned 4:2:0 chroma subsampling strategies for the mosaic videos with different RGB-CFA structures in HEVC. (a) Bayer CFA. (b) Lukac and Plataniotis CFA. (c) Vertical stripe CFA.

TABLE I

BD-PSNR RESULTS (dB) OF THE 4:2:0(L), 4:2:0(R), YANG *et al.*'s, IDEAL, AND PROPOSED CHROMA SUBSAMPLING STRATEGIES OVER THE 4:2:0(A) CHROMA SUBSAMPLING STRATEGY FOR THE TWO MOSAIC VIDEO SETS AND DIFFERENT QP INTERVALS IN H.264/AVC

QP interval	Video set	BD-PSNR									
		Reconstructed mosaic video					Reconstructed full-color video				
		4:2:0 (L)	4:2:0 (R)	Yang *et al.*'s	Ideal	Proposed	4:2:0 (L)	4:2:0 (R)	Yang *et al.*'s	Ideal	Proposed
4–16	Set1	-0.7271	-0.5354	0.9234	1.2389	1.2216	-0.7243	-0.3862	0.8747	1.2303	1.2183
	Set2	-0.8223	-0.6348	1.2214	1.5459	1.5192	-0.8269	-0.4811	1.2323	1.5876	1.5613
	Average	-0.7747	-0.5851	1.0724	1.3924	1.3704	-0.7756	-0.4337	1.0535	1.4089	1.3898
12–24	Set1	-0.4121	-0.2976	0.3156	0.4451	0.4168	-0.4554	-0.2389	0.2785	0.4525	0.4253
	Set2	-0.5091	-0.3795	0.5770	0.7330	0.6909	-0.5519	-0.3012	0.5848	0.7827	0.7393
	Average	-0.4606	-0.3386	0.4463	0.5891	0.5538	-0.5036	-0.2700	0.4316	0.6176	0.5823
20–32	Set1	-0.1964	-0.1275	0.0694	0.1086	0.0867	-0.2268	-0.1100	0.0474	0.1079	0.0868
	Set2	-0.2361	-0.1606	0.1941	0.2566	0.2233	-0.2753	-0.1295	0.1916	0.2803	0.2451
	Average	-0.2163	-0.1441	0.1317	0.1826	0.1550	-0.2511	-0.1197	0.1195	0.1941	0.1659
24–36	Set1	-0.1347	-0.0844	0.0212	0.0342	0.0215	-0.1548	-0.0748	0.0073	0.0330	0.0207
	Set2	-0.1548	-0.1018	0.1018	0.1394	0.1144	-0.1835	-0.0814	0.0969	0.1522	0.1259
	Average	-0.1448	-0.0931	0.0615	0.0868	0.0680	-0.1692	-0.0781	0.0521	0.0926	0.0733

difference over the same bitrate interval for any two concerned chroma subsampling strategies. We take the 4:2:0(A) chroma subsampling strategy to be the comparison basis and then tabulate the BD-PSNR results for the 4:2:0(L), 4:2:0(R), Yang *et al.*'s, ideal, and proposed chroma subsampling strategies in Tables I and II. In Tables I and II, a positive value of BD-PSNR, respectively, indicates, compared with the 4:2:0(A) chroma subsampling strategy, the quality superiority of the corresponding chroma subsampling strategy for compressing mosaic videos in H.264/AVC and HEVC. Thus, it is clear that in the low and middle QP circumstances, both the proposed and ideal chroma subsampling strategies outperform Yang *et al.*'s one and dominate the three commonly used ones in terms of the BD-PSNR performance, implying that compressing mosaic videos with either of both the proposed and ideal chroma subsampling strategies can indeed deliver better quality of the reconstructed mosaic videos and the reconstructed full-color videos under the same bitrate. Although the BD-PSNR performances of both the proposed and ideal chroma subsampling strategy decay for the high QP circumstances, it seldom occurs in practice since encoding mosaic videos at a low bitrate often leads to obvious compression artifacts and poor visual perception.

Furthermore, similar to the calculation of BD-PSNR, we also compute the average VQM difference of the reconstructed full-color videos over the same bitrate interval for any two

TABLE II

BD-PSNR RESULTS (dB) OF THE 4:2:0(L), 4:2:0(R), YANG *et al.*'s, IDEAL, AND PROPOSED CHROMA SUBSAMPLING STRATEGIES OVER THE 4:2:0(A) CHROMA SUBSAMPLING STRATEGY FOR THE TWO MOSAIC VIDEO SETS AND DIFFERENT QP INTERVALS IN HEVC

QP interval	Video set	BD-PSNR									
		Reconstructed mosaic video					Reconstructed full-color video				
		4:2:0(L)	4:2:0(R)	Yang *et al.*'s	Ideal	Proposed	4:2:0(L)	4:2:0(R)	Yang *et al.*'s	Ideal	Proposed
4–16	Set1	-0.7258	-0.5271	1.0403	1.3766	1.3922	-0.7006	-0.3530	1.0193	1.3771	1.4030
	Set2	-0.7763	-0.5900	1.3132	1.6467	1.6353	-0.7697	-0.4250	1.3441	1.6940	1.6894
	Average	-0.7511	-0.5585	1.1768	1.5117	1.5138	-0.7352	-0.3890	1.1817	1.5355	1.5462
12–24	Set1	-0.3866	-0.2650	0.4525	0.5579	0.5617	-0.4062	-0.1796	0.4591	0.5995	0.6117
	Set2	-0.4600	-0.3323	0.6549	0.7884	0.7725	-0.4867	-0.2399	0.6888	0.8561	0.8448
	Average	-0.4233	-0.2986	0.5537	0.6731	0.6671	-0.4464	-0.2098	0.5739	0.7278	0.7282
20–32	Set1	-0.1905	-0.1141	0.1611	0.1869	0.1850	-0.2072	-0.0784	0.1680	0.2100	0.2144
	Set2	-0.2257	-0.1453	0.2711	0.3209	0.3079	-0.2500	-0.1013	0.2880	0.3585	0.3481
	Average	-0.2081	-0.1297	0.2161	0.2539	0.2464	-0.2286	-0.0899	0.2280	0.2842	0.2812
24–36	Set1	-0.1341	-0.0733	0.0846	0.0887	0.0883	-0.1455	-0.0515	0.0897	0.1030	0.1077
	Set2	-0.1487	-0.0864	0.1610	0.1890	0.1803	-0.1661	-0.0571	0.1704	0.2133	0.2057
	Average	-0.1414	-0.0798	0.1228	0.1389	0.1343	-0.1558	-0.0543	0.1301	0.1581	0.1567

TABLE III

AVERAGE VQM DIFFERENCE OF THE 4:2:0(L), 4:2:0(R), YANG *et al.*'s, IDEAL, AND PROPOSED CHROMA SUBSAMPLING STRATEGIES OVER THE 4:2:0(A) CHROMA SUBSAMPLING STRATEGY FOR THE TWO MOSAIC VIDEO SETS AND DIFFERENT QP INTERVALS IN H.264/AVC AND HEVC

QP interval	Video set	In the H.264/AVC environment					In the HEVC environment				
		4:2:0(L)	4:2:0(R)	Yang *et al.*'s	Ideal	Proposed	4:2:0(L)	4:2:0(R)	Yang *et al.*'s	Ideal	Proposed
4–16	Set1	0.0043	0.0017	-0.0032	-0.0047	-0.0049	0.0045	0.0022	-0.0030	-0.0045	-0.0047
	Set2	0.0050	0.0018	-0.0036	-0.0048	-0.0049	0.0050	0.0019	-0.0036	-0.0043	-0.0042
	Average	0.0047	0.0018	-0.0034	-0.0048	-0.0049	0.0047	0.0020	-0.0033	-0.0047	-0.0049
12–24	Set1	0.0037	0.0013	-0.0031	-0.0039	-0.0041	0.0041	0.0015	-0.0028	-0.0036	-0.0038
	Set2	0.0052	0.0016	-0.0033	-0.0041	-0.0042	0.0049	0.0016	-0.0032	-0.0043	-0.0042
	Average	0.0044	0.0015	-0.0032	-0.0040	-0.0042	0.0045	0.0015	-0.0030	-0.0039	-0.0040
20–32	Set1	0.0035	0.0010	-0.0038	-0.0035	-0.0042	0.0040	0.0008	-0.0035	-0.0033	-0.0036
	Set2	0.0055	0.0016	-0.0031	-0.0038	-0.0041	0.0044	0.0010	-0.0033	-0.0044	-0.0042
	Average	0.0045	0.0013	-0.0035	-0.0037	-0.0041	0.0042	0.0009	-0.0034	-0.0038	-0.0039
24–36	Set1	0.0035	0.0013	-0.0035	-0.0036	-0.0036	0.0035	0.0007	-0.0027	-0.0029	-0.0029
	Set2	0.0039	0.0010	-0.0029	-0.0031	-0.0031	0.0035	0.0006	-0.0031	-00037	-0.0036
	Average	0.0037	0.0012	-0.0032	-0.0034	-0.0034	0.0035	0.0007	-0.0029	-0.0033	-0.0033

concerned chroma subsampling strategies. Table III gives the average VQM difference results for the 4:2:0(L), 4:2:0(R), Yang *et al.*'s, ideal, and proposed chroma subsampling strategies, in which the 4:2:0(A) is taken as the comparison basis. A negative value of the average VQM difference in Table III indicates, compared with the 4:2:0(A) chroma subsampling strategy, the quality superiority of the corresponding chroma subsampling strategy for compressing mosaic videos. It is worth noting that the conclusions drawn from the average VQM difference results exactly coincide with those from the BD-PSNR results, which confirms again the quality superiority of both the proposed and ideal chroma subsampling strategies for compressing mosaic videos in H.264/AVC and HEVC.

B. Comparison in Terms of Execution Time

To demonstrate the feasibility of one chroma subsampling strategy, the execution time of the concerned chroma subsampling strategy itself and that of the whole compression procedure should be addressed. This section investigates the execution time comparison among the proposed chroma subsampling strategy and the other five concerned ones. For each concerned chroma subsampling strategy, each of the seven RGB-CFA structures, and the H.264/AVC and HEVC environments, we measure, over the cases of the ten considered QPs as well as Set1 and Set2, the average execution time spent, respectively, in the chroma subsampling operation and in the whole compression procedure and then tabulate the results in Tables IV–VI.

From Tables IV–VI, it is obvious that the ideal chroma subsampling strategy leads to extremely high execution time in the chroma subsampling operation and the whole compression procedure compared with the other five concerned ones. This is mainly because the ideal chroma subsampling strategy requires performing an exhaustive search over all possible values of the U and V components so as to determine the

TABLE IV

COMPARISON OF THE EXECUTION TIMES REQUIRED IN THE CHROMA SUBSAMPLING OPERATION FOR THE SIX CONCERNED STRATEGIES UNDER THE H.264/AVC AND HEVC ENVIRONMENTS

CFA	Video set	4:2:0 (A)	4:2:0 (L)	4:2:0 (R)	Yang *et al.*'s	Ideal	Proposed
Bayer	Set1	0.058	0.058	0.068	0.172	2.4×10^4	2.394
	Set2	0.052	0.059	0.102	0.161	2.4×10^4	2.379
Lukac and Plataniotis	Set1	0.057	0.056	0.069	0.174	2.4×10^4	2.386
	Set2	0.050	0.060	0.081	0.197	2.4×10^4	2.414
Yamanaka	Set1	0.059	0.056	0.066	0.173	2.4×10^4	2.533
	Set2	0.050	0.059	0.077	0.176	2.4×10^4	2.480
Diagonal stripe	Set1	0.057	0.056	0.068	0.175	2.4×10^4	2.518
	Set2	0.055	0.062	0.070	0.160	2.4×10^4	2.476
Vertical stripe	Set1	0.057	0.055	0.069	0.179	2.4×10^4	2.277
	Set2	0.049	0.066	0.086	0.165	2.4×10^4	2.244
Modified Bayer	Set1	0.056	0.056	0.068	0.172	2.4×10^4	2.558
	Set2	0.049	0.064	0.088	0.160	2.4×10^4	2.466
HVS-based	Set1	0.055	0.056	0.068	0.184	2.4×10^4	2.495
	Set2	0.050	0.057	0.089	0.165	2.4×10^4	2.417
Average		0.054	0.058	0.076	0.172	2.4×10^4	2.431

TABLE V

COMPARISON OF THE EXECUTION TIMES REQUIRED IN THE WHOLE COMPRESSION PROCEDURE FOR THE SIX CONCERNED STRATEGIES UNDER THE H.264/AVC ENVIRONMENT

CFA	Video set	4:2:0 (A)	4:2:0 (L)	4:2:0 (R)	Yang *et al.*'s	Ideal	Proposed
Bayer	Set1	2034.89	2034.14	2032.91	2033.14	2.6×10^4	2037.00
	Set2	2048.04	2048.00	2049.40	2050.83	2.6×10^4	2051.02
Lukac and Plataniotis	Set1	2049.88	2050.06	2049.82	2049.65	2.6×10^4	2053.76
	Set2	2058.05	2054.00	2054.42	2056.93	2.6×10^4	2060.60
Yamanaka	Set1	2053.19	2049.53	2051.62	2052.46	2.6×10^4	2054.66
	Set2	2058.95	2059.67	2059.89	2056.65	2.6×10^4	2061.03
Diagonal stripe	Set1	2040.47	2039.69	2042.15	2042.75	2.6×10^4	2044.90
	Set2	2043.97	2049.01	2051.01	2043.65	2.6×10^4	2050.59
Vertical stripe	Set1	2089.58	2090.28	2090.36	2092.72	2.6×10^4	2093.53
	Set2	2047.00	2045.14	2045.35	2043.69	2.6×10^4	2047.57
Modified Bayer	Set1	2035.99	2036.62	2035.46	2036.85	2.6×10^4	2040.01
	Set2	2041.65	2042.25	2044.21	2042.95	2.6×10^4	2047.22
HVS-based	Set1	2023.12	2023.52	2023.09	2024.73	2.6×10^4	2026.31
	Set2	2047.71	2050.15	2049.63	2050.57	2.6×10^4	2050.95
Average		2048.04	2048.00	2048.52	2048.40	2.6×10^4	2051.37

TABLE VI

COMPARISON OF THE EXECUTION TIMES REQUIRED IN THE WHOLE COMPRESSION PROCEDURE FOR THE SIX CONCERNED STRATEGIES UNDER THE HEVC ENVIRONMENT

CFA	Video set	4:2:0 (A)	4:2:0 (L)	4:2:0 (R)	Yang *et al.*'s	Ideal	Proposed
Bayer	Set1	809.10	810.08	796.09	811.69	2.5×10^4	813.69
	Set2	1208.02	1222.01	1212.42	1217.52	2.5×10^4	1216.60
Lukac and Plataniotis	Set1	874.03	876.77	853.05	875.14	2.5×10^4	876.30
	Set2	1221.53	1238.38	1224.08	1236.71	2.5×10^4	1233.27
Yamanaka	Set1	874.63	869.16	843.36	877.01	2.5×10^4	879.60
	Set2	1223.96	1238.13	1227.03	1238.91	2.5×10^4	1236.98
Diagonal stripe	Set1	790.21	790.49	763.23	814.62	2.5×10^4	832.54
	Set2	1225.73	1241.94	1230.07	1236.51	2.5×10^4	1239.19
Vertical stripe	Set1	987.56	993.04	990.17	995.92	2.5×10^4	996.56
	Set2	1259.68	1281.67	1270.23	1274.72	2.5×10^4	1272.25
Modified Bayer	Set1	870.38	861.17	866.39	884.21	2.5×10^4	889.69
	Set2	1220.84	1235.46	1223.24	1234.61	2.5×10^4	1233.28
HVS-based	Set1	907.60	900.79	897.46	917.38	2.5×10^4	920.99
	Set2	1251.78	1253.58	1254.45	1253.37	2.5×10^4	1252.75
Average		1051.79	1058.05	1046.52	1062.02	2.5×10^4	1063.83

sampled U and V components for each 2×2 YUV block, and hence, its execution time requirement increases substantially, implying that the ideal chroma subsampling strategy is infeasible in practice even though it has superior quality and bitrate tradeoff. In contrast, although the proposed chroma subsampling strategy needs more execution time in the chroma subsampling operation than the 4:2:0(A), 4:2:0(L), 4:2:0(R), and Yang *et al.*'s ones, it, on average, causes only at most 1.65% [$=(1063.83 - 1046.52)/1046.52$] increase in the execution time of the whole compression procedure, revealing that the proposed chroma subsampling strategy for compressing

mosaic videos in H.264/AVC and HEVC is competitive, in terms of execution time requirement, with Yang *et al.*'s and the three commonly used ones.

In summary, the proposed novel chroma subsampling strategy based on mathematical optimization for compressing mosaic videos with arbitrary RGB-CFA structures in H.264/AVC and HEVC does improve the quality and bitrate tradeoff achieved by Yang *et al.*'s and the three commonly used ones, while being computationally feasible in practice.

V. Conclusion

In this paper, we have presented a novel chroma subsampling strategy based on mathematical optimization for compressing mosaic videos with arbitrary RGB-CFA structures in H.264/AVC and HEVC. For each 2×2 YUV block to be subsampled with 4:2:0 format, the proposed strategy determines the proper sampled U and V components by minimizing, prior to compression, the quality distortion between the original colocated mosaic block and the mosaic block conversed from the current subsampled YUV block. Different from the previous chroma subsampling strategies that either always sample the U and V components from the fixed positions or only overfavor the R and B pixels' reconstruction when determining the sampled U and V components, the proposed chroma subsampling strategy simultaneously considers, through the formulated mathematical optimization, the significance of the sampled U and V components for reconstructing R, G, and B pixels so as to determine the proper sampled U and V components for each 2×2 YUV block. The experimental results on the 28 test videos with the seven RGB-CFA structures demonstrate that the proposed chroma subsampling strategy does achieve the best quality and bitrate tradeoff at a similar execution time requirement for compressing mosaic videos in H.264/AVC and HEVC compared with the state-of-the-art ones by Chen *et al.* and Yang *et al.* as well as the three commonly used ones.

References

[1] B. E. Bayer, "Color imaging array," U.S. Patent 3 971 065, Jul. 20, 1976.

[2] G. Bjontegaard, *Calculation of Average PSNR Difference Between RD-Curves*, Austin, TX, USA, document VECG-M33, Apr. 2001.

[3] B. Bross, W.-J. Han, J.-R. Ohm, G. J. Sullivan, Y.-K. Wang, and T. Wiegand, *High Efficiency Video Coding (HEVC) Text Specification Draft 10*, document JCTVC-L1003, Jan. 2013.

[4] H. Chen, M. Sun, and E. Steinbach, "Compression of Bayer-pattern video sequences using adjusted chroma subsampling," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 19, no. 12, pp. 1891–1896, Dec. 2009.

[5] K.-H. Chung and Y.-H. Chan, "Color demosaicing using variance of color differences," *IEEE Trans. Image Process.*, vol. 15, no. 10, pp. 2944–2955, Oct. 2006.

[6] K.-L. Chung, W.-J. Yang, W.-M. Yan, and C.-S. Fuh, "New joint demosaicing and arbitrary-ratio resizing algorithm for color filter array based on DCT approach," *IEEE Trans. Consum. Electron.*, vol. 56, no. 2, pp. 783–791, May 2010.

[7] C. Doutre, P. Nasiopoulos, and K. N. Plataniotis, "H.264-based compression of Bayer pattern video sequences," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 18, no. 6, pp. 725–734, Jun. 2008.

[8] C. Doutre and P. Nasiopoulos, "Modified H.264 intra prediction for compression of video and images captured with a color filter array," in *Proc. 16th IEEE Int. Conf. Image Process.*, Nov. 2009, pp. 3401–3404.

[9] *Draft ITU-T Recommendation and Final Draft International Standard of Joint Video Specification*, document ITU-T Rec. H.264/ISO/IEC 14496-10 AVC, May 2003.

[10] S. Ferradans, M. Bertalmio, and V. Caselles, "Geometry-based demosaicking," *IEEE Trans. Image Process.*, vol. 18, no. 3, pp. 665–670, Mar. 2009.

[11] F. Gastaldi, C. C. Koh, M. Carli, A. Neri, and S. K. Mitra, "Compression of videos captured via Bayer patterned color filter arrays," in *Proc. 13th Eur. Signal Process. Conf.*, Sep. 2005, pp. 1–4.

[12] *Generic Coding of Moving Pictures and Associated Audio (MPEG-2)*, document ISO/IEC 13818, Jul. 1995.

[13] B. K. Gunturk, J. Glotzbach, Y. Altunbasak, R. W. Schafer, and R. M. Mersereau, "Demosaicking: Color filter array interpolation," *IEEE Signal Process. Mag.*, vol. 22, no. 1, pp. 44–54, Jan. 2005.

[14] K. Hirakawa and P. J. Wolfe, "Spatio-spectral color filter array design for optimal image recovery," *IEEE Trans. Image Process.*, vol. 17, no. 10, pp. 1876–1890, Oct. 2008.

[15] *Objective Perceptual Video Quality Measurement Techniques for Digital Cable Television in the Presence of a Full Reference*, document ITU-T Rec. J.144, Mar. 2004.

[16] S.-Y. Lee and A. Ortega, "A novel approach of image compression in digital cameras with a Bayer color filter array," in *Proc. IEEE Int. Conf. Image Process.*, Oct. 2001, pp. 482–485.

[17] J. S. J. Li and S. Randhawa, "Color filter array demosaicking using high-order interpolation techniques with a weighted median filter for sharp color edge preservation," *IEEE Trans. Image Process.*, vol. 18, no. 9, pp. 1946–1957, Sep. 2009.

[18] W. Lu and Y.-P. Tan, "Color filter array demosaicking: New method and performance measures," *IEEE Trans. Image Process.*, vol. 12, no. 10, pp. 1194–1210, Oct. 2003.

[19] R. Lukac and K. N. Plataniotis, "Normalized color-ratio modeling for CFA interpolation," *IEEE Trans. Consum. Electron.*, vol. 50, no. 2, pp. 737–745, May 2004.

[20] R. Lukac and K. N. Plataniotis, "Universal demosaicking for imaging pipelines with an RGB color filter array," *Pattern Recognit.*, vol. 38, no. 11, pp. 2208–2212, Apr. 2005.

[21] R. Lukac and K. N. Plataniotis, "Color filter arrays: Design and performance analysis," *IEEE Trans. Consum. Electron.*, vol. 51, no. 4, pp. 1260–1267, Nov. 2005.

[22] H. S. Malvar and G. J. Sullivan, "Progressive-to-lossless compression of color-filter-array images using macropixel spectral-spatial transformation," in *Proc. Data Compress. Conf.*, Apr. 2012, pp. 3–12.

[23] S.-C. Pei and I.-K. Tam, "Effective color interpolation in CCD color filter arrays using signal correlation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 6, pp. 503–513, Jun. 2003.

[24] M. H. Pinson and S. Wolf, "A new standardized method for objectively measuring video quality," *IEEE Trans. Broadcast.*, vol. 50, no. 3, pp. 312–322, Sep. 2004.

[25] X. Wu and N. Zhang, "Primary-consistent soft-decision color demosaicking for digital cameras (patent pending)," *IEEE Trans. Image Process.*, vol. 13, no. 9, pp. 1263–1274, Sep. 2004.

[26] X. Wu and X. Zhang, "Joint color decrosstalk and demosaicking for CFA cameras," *IEEE Trans. Image Process.*, vol. 19, no. 12, pp. 3181–3189, Dec. 2010.

[27] W.-J. Yang, K.-L. Chung, W.-N. Yang, and L.-C. Lin, "Universal chroma subsampling strategy for compressing mosaic video sequences with arbitrary RGB color filter arrays in H.264/AVC," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 23, no. 4, pp. 591–606, Apr. 2013.

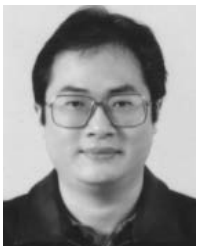
[28] *Batch Video Quality Metric Software 2.0*. [Online]. Available: http://www.its.bldrdoc.gov/resources/video-quality-research/guides-and-tutorials/description-of-vqm-tools.aspx

[29] *Kodak True Color Image Collection*. [Online]. Available: http://r0k.us/graphics/kodak/

[30] [Online]. Available: http://140.118.175.164/CWYu/ChromaSubsampling/

Chien-Hsiung Lin received the B.S., M.S., and Ph.D. degrees in information management from National Taiwan University of Science and Technology (NTUST), Taipei, Taiwan, in 2003, 2005, and 2011, respectively.

He is a Post-Doctoral Researcher with the Department of Computer Science and Information Engineering, NTUST. His research interests include digital image processing, video compression, multiview video coding, data hiding, statistical analysis, and stochastic simulation.



Kuo-Liang Chung (M'91–SM'01) received the B.S., M.S., and Ph.D. degrees from National Taiwan University, Taipei, Taiwan, in 1982, 1984, and 1990, respectively.

He was a Research Assistant with Institute of Information Science, Academia Sinica, Taipei, from 1986 to 1987, after two years of obligatory military service from 1984 to 1986. He was a Visiting Scholar with University of Washington, Seattle, WA, USA, in 1999. From 2003 to 2006, he was the Chair of the Department of Computer Science and Information Engineering with National Taiwan University of Science and Technology (NTUST), Taipei. He has been a Professor with the Department of Computer Science and Information Engineering, NTUST, since 1995, where he has been a University Chair Professor since 2009. His research interests include video compression, multiview video coding, data hiding, camera image processing, and shape analysis.

Dr. Chung is a fellow of the Institute of Engineering and Technology. He was a recipient of the Distinguished Research Award from 2004 to 2007, and the Distinguished Research Project Award from 2009 to 2012, from the National Science Council of Taiwan, the Best Paper Award from the Image Processing and Pattern Recognition Society of Taiwan in 2007, and the Distinguished Teaching Award of NTUST in 2009. In 2000, he was the Program Co-Chair of the Conference on Computer Vision, Graphics, and Image Processing, Taiwan. He had been a Managing Editor of *Journal of Chinese Institute of Engineers* from 1996 to 1998. He is currently an Associate Editor of *Journal of Visual Communication and Image Representation*.

Chun-Wei Yu received the B.S. degree in computer science and information engineering from National Taiwan Ocean University, Keelung, Taiwan, in 2013. He is currently working toward the M.S. degree in computer science and information engineering with National Taiwan University of Science and Technology, Taipei, Taiwan.

His research interests include image processing, video compression, and algorithms.