Journal of Electronic Imaging

SPIEDigitalLibrary.org/jei

Efficient chroma subsampling strategy for compressing digital time delay integration mosaic video sequences in H.264/AVC

Kuo-Liang Chung
Wei-Jen Yang
Chyou-Hwa Chen
Hong-Yuan Mark Liao
Sheng-Mao Zeng

SPIE

IS&T imaging.org

Efficient chroma subsampling strategy for compressing digital time delay integration mosaic video sequences in H.264/AVC

Kuo-Liang Chung
Wei-Jen Yang
Chyou-Hwa Chen
National Taiwan University of Science and Technology
Department of Computer Science and Information Engineering
No. 43, Section 4, Keelung Road
Taipei, 10672 Taiwan
E-mail: wjyang@mail.ntust.edu.tw

Hong-Yuan Mark Liao
Academia Sinica
Institute of Information Science
No. 128, Section 2, Academia Road
Taipei, 11529 Taiwan

Sheng-Mao Zeng
National Taiwan University of Science and Technology
Department of Computer Science and Information Engineering
No. 43, Section 4, Keelung Road
Taipei, 10672 Taiwan

Abstract. *Digital time delay and integration (DTDI) mosaic video sequences captured by high-speed DTDI line-scan cameras are commonly used in industrial print inspection and high-speed capture applications. To reduce the memory requirement for saving these video sequences, it is necessary to compress them. In this paper, we present an efficient chroma subsampling strategy for compressing DTDI mosaic video sequences in H.264/AVC. Based on the color domain transform between the RGB domain and the YUV domain, a position-selection strategy is proposed to determine the two subsampling chroma components, U and V, according to the DTDI mosaic structure. The quality of reconstructed DTDI video sequences is better than those reconstructed by conventional methods. By experimenting on some popular test DTDI mosaic video sequences, the results turned out to be superior than conventional ones that adopt H.264/AVC as compression standard.* © 2011 SPIE and IS&T.
[DOI: 10.1117/1.3586799]

1 Introduction

Recently, area-scan digital cameras are getting more popular in the consumer electronics market. To cut down the cost, instead of using three charge-coupled device/complementary metal-oxide-semiconductor (CCD/CMOS) sensors for each pixel, most manufacturers use a single sensor with the Bayer color filter array (CFA) structure[1–4] to capture the color information. In the Bayer CFA structure depicted in Fig. 1, each pixel in the captured image frame of a video sequence records only one of the three primary color components. Because the green (G) component is the most important factor to determine the luminance of a color image frame, half of the pixels are assigned to the G component. The red (R) and blue (B) components which share the rest of the pixels are considered the chrominance of an image frame.

Besides an area-scan camera, a line-scan camera is another type of digital camera. The line-scan camera is mostly used in industrial and high-speed capture applications, e.g., the print inspection community.[5] Because of the short exposure time and high-speed transport restrictions, in many cases it is very difficult to increase the intensity of illumination. To overcome this problem, some special techniques, such as high dynamic range imaging[6] or time delay and integration (TDI),[7] are demanded to maintain image frame quality while the quantity of available light decreases. Nowadays, digital time delay and integration (DTDI) line-scan cameras[5,8,9] based CMOS sensors with the Bayer CFA structure and field programmable gate arrays (FPGAs) have been developed. Because the sensors used in DTDI line-scan cameras are realized by the CMOS technique, DTDI line-scan cameras have

Paper 10192R received Nov. 5, 2010; revised manuscript received Mar. 30, 2011; accepted for publication Apr. 13, 2011; published online May 24, 2011.

1017-9909/2011/20(2)/023011/15/$25.00 © 2011 SPIE and IS&T

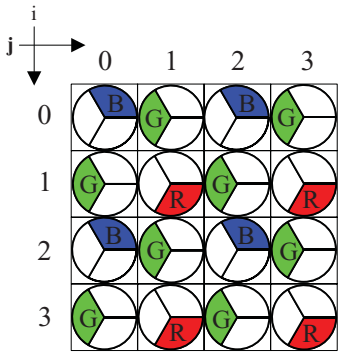


Fig. 1 Bayer CFA structure.

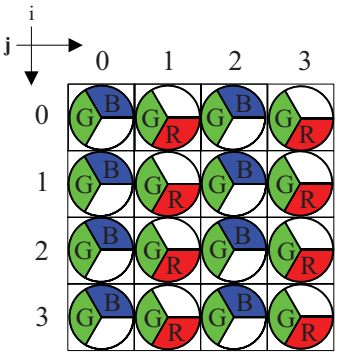


Fig. 3 DTDI structure.

the advantages of low power consumption, low production cost, and high-speed. For DTDI line-scan cameras, instead of scanning moving objects row by row, a large number of moving object rows are exposed in parallel, and then they are sequentially accumulated in the responses. Since the object is moved by one row after each exposure, the number of times with which each object pixel is captured is as many as the DTDI stages. Finally, the output of each pixel is collected by the responses of the partial exposures.

Figure 2 illustrates the principle of a DTDI line-scan camera. From Fig. 2 it is obvious that the principle of a DTDI line-scan camera is constructed from a Bayer CFA and some delay stages, denoted as z^{-1}. The delay stages, implemented by FPGAs, are used to integrate two consecutive rows captured at different time instants. For each exposure, either BG-row or GR-row would be captured and delivered. This indicates that each object pixel will be captured by either the G and R components or the G and B components alternately. The capturing and delivering process would continue until the number of required DTDI stages is reached, and then the output of each pixel could be computed. For this example with four DTDI stages, the resultant output of each pixel can be computed by the following rule:

$$
\begin{aligned}
B'_{2m} &= B_{0,2m} + B_{2,2m}z^{-2}, \\
G'_{2m} &= G_{1,2m}z^{-1} + G_{3,2m}z^{-3}, \\
G'_{2m+1} &= G_{0,2m+1} + G_{2,2m+1}z^{-2}, \\
R'_{2m+1} &= R_{1,2m+1}z^{-1} + R_{3,2m+1}z^{-3}.
\end{aligned}
$$

Video sequences captured by a DTDI line-scan camera are called DTDI mosaic video sequences. For a DTDI mosaic video sequence, Fig. 3 depicts the structure of a DTDI mosaic image frame. In Fig. 3, it is observed that the G components are fully populated in the image frame, whereas the B and R components are only available at even and odd column positions, respectively. Consequently, each pixel in the DTDI mosaic image frame comprises two color components, G and R or G and B.

When transmitting DTDI mosaic video sequences through networks or saving them in a storage device, it is necessary to compress them using compression standards, such as H.264/AVC video coding standard.[10–13] In the past, a number of compression algorithms for Bayer CFA images and video sequences captured by area-scan cameras were developed.[14–22] However, these existing compression algorithms cannot be directly used in compressing DTDI mosaic video sequences because the structure of a DTDI mosaic image frame is quite different from that of a Bayer CFA image frame. Due to the important application of a line-scan camera which can be used to capture DTDI video sequences, we want to develop a modified chroma subsampling strategy for compressing DTDI mosaic video sequences.

In the proposed chroma subsampling strategy, exploiting the color space transform between the RGB domain and the YUV domain, we appropriately select the positions to obtain the two subsampling chroma components, U and V, according to the DTDI mosaic structure. It would reconstruct DTDI video sequences with good quality after decompression. Based on some real DTDI mosaic video sequences, experimental results demonstrate that the proposed chroma subsampling strategy is superior to conventional H.264/AVC-based approaches.

The remainder of this paper is organized as follows. In Sec. 2, we introduce the traditional chroma subsampling strategy in H.264/AVC and point out its shortcoming when

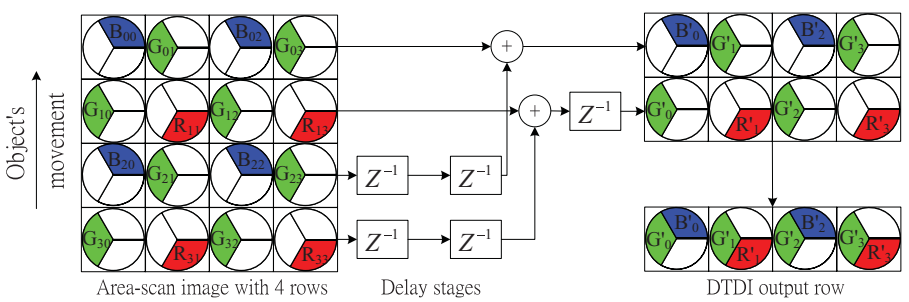


Fig. 2 The principle of a DTDI line-scan camera.

reconstructing a compressed DTDI mosaic video sequence. In Sec. 3, the proposed modified chroma subsampling strategy for compressing DTDI mosaic video sequences is presented. In Sec. 4, we report the experimental results to demonstrate the quality advantage of the proposed chroma subsampling strategy. Section 5 addresses some concluding remarks.

2 Traditional Chroma Subsampling Strategy in H.264/AVC

In this section, we first introduce the traditional chroma subsampling strategy in H.264/AVC, and then point out its shortcoming when reconstructing a compressed DTDI mosaic video sequence. Let a DTDI mosaic video sequence with n image frames be denoted as $F_{\mathrm{DTDI}} = \{F^k_{\mathrm{DTDI}} | \forall k \in \{1, 2, \ldots, n\}\}$, where F^k_{DTDI} denotes the k'th DTDI mosaic image frame.

Given a DTDI mosaic video sequence, F_{DTDI}, we first demosaic it to obtain a demosaiced full color RGB video sequence $F_{\mathrm{RGB}} = \{F^k_{\mathrm{RGB}} | \forall k \in \{1, 2, \ldots, n\}\}$ by using a DTDI demosaicing method. Any existing DTDI demosaicing method[9] can be used in this step. For each image frame in F_{RGB}, we convert it from the RGB color domain into the YUV color domain, which is the most commonly used input color model in H.264/AVC. Thus, we have a YUV video sequence $F_{\mathrm{YUV}} = \{F^k_{\mathrm{YUV}} | \forall k \in \{1, 2, \ldots, n\}\}$. The color domain transform between the RGB color domain and the YUV color domain is expressed by the following two equations:[23]

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

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

To compress YUV video sequences by using H.264/AVC, the chroma subsampling strategy should be applied to the two chroma components, U and V. In the traditional subsampling strategy in H.264/AVC, one 4:2:2 format and two types of 4:2:0 formats, i.e., the 4:2:0(I) format and the 4:2:0(II) format, are three commonly used subsampling formats. Given a 2×2 YUV block in a image frame, F^k_{YUV}, as shown in Fig. 4, Figs. 5(a)–5(c) illustrate the depictions of the 4:2:2 format, the 4:2:0(I) format, and the 4:2:0(II) format, respectively. From Fig. 5(a), it is clear that based on the 4:2:2 format, only the U and V components in the left column of a 2×2 block

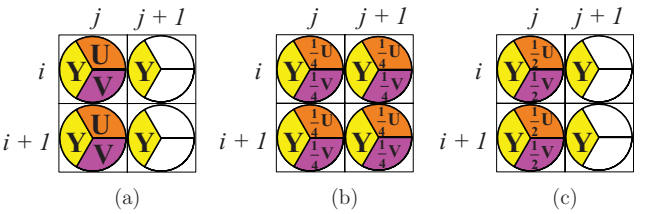


Fig. 5 Three subsampling formats of traditional chroma subsampling strategy in H.264/AVC. (a) The 4:2:2 format. (b) The 4:2:0(I) format. (c) The 4:2:0(II) format.

are retained, whereas those in the right column are discarded. In the two types of 4:2:0 formats, a 2×2 block holds only one average U component and one average V component. For the 4:2:0(I) format shown in Fig. 5(b), the average U and V components are produced by averaging four U and V components in the block, respectively. For the 4:2:0(II) format shown in Fig. 5(c), the average U and V components are produced by averaging two U and V components in the left column of a 2×2 block, respectively.

Next, we describe the quality disadvantage, i.e., the bias problem, in the traditional chroma subsampling strategy when reconstructing a compressed DTDI mosaic video sequence. From Eq. (2), it is obvious that the R component can be reconstructed by $R = 1.164(Y - 16) + 1.596(V - 128)$, in which only Y and V components are required. Similarly, only Y and U components are needed to reconstruct the B component. Therefore, the V component is more significant than the U component when reconstructing the R component. Conversely, the U component is more significant than the V component when reconstructing the B component. For the YUV block shown in Fig. 4, suppose in the corresponding original DTDI mosaic image frame, F^k_{DTDI}, as shown in Fig. 6, the two color components of the pixels at positions (i, j) and $(i + 1, j)$ are G and B, and thus those of the pixels at positions $(i, j + 1)$ and $(i + 1, j + 1)$ are G and R. When reconstructing a compressed DTDI mosaic video sequence, based on the 4:2:2 format shown in Fig. 5, we have complete Y and U components to reconstruct the two corresponding B components of the pixels at positions (i, j) and $(i + 1, j)$. However, due to the lack of V components at positions $(i, j + 1)$ and $(i + 1, j + 1)$, the reconstructed R components of the pixels at positions $(i, j + 1)$ and $(i + 1, j + 1)$ will be biased. With respect to the 4:2:0(I) and 4:2:0(II) formats, the similar bias problem also exists. This bias problem would result in quality degradation of the reconstructed DTDI mosaic video sequences.

To overcome such a bias problem in the traditional chroma subsampling strategy, in Sec. 3, we will present a new

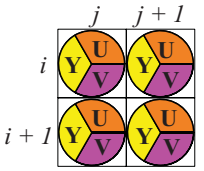


Fig. 4 A 2×2 YUV mosaic block in F^k_{YUV}.

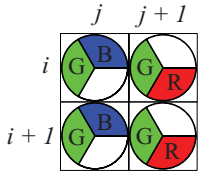


Fig. 6 A 2×2 DTDI mosaic block in F^k_{DTDI}.

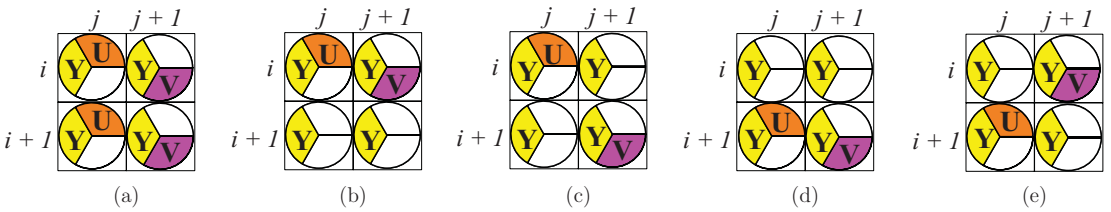


Fig. 7 Five subsampling formats of the proposed modified chroma subsampling strategy. (a) The 4:2:2 format. (b) The 4:2:0(I) format. (c) The 4:2:0(II) format. (d) The 4:2:0(III) format. (e) The 4:2:0(IV) format.

modified chroma subsampling strategy for compressing DTDI mosaic video sequences.

3 Proposed Modified Chroma Subsampling Strategy

In this section, we present a modified chroma subsampling strategy for compressing DTDI mosaic video sequences by considering the correlation between the DTDI structure and the chroma subsampling positions. The proposed modified chroma subsampling strategy could overcome the bias problem in the traditional chroma subsampling strategy and it would result in better quality of reconstructed DTDI video sequences.

Given a DTDI video sequence, $F_{\mathrm{DTDI}} = \{F^k_{\mathrm{DTDI}} | \forall k \in \{1, 2, \ldots, n\}\}$, after the demosaicing and color domain transform processes, we have a YUV video sequence $F_{\mathrm{YUV}} = \{F^k_{\mathrm{YUV}} | \forall k \in \{1, 2, \ldots, n\}\}$. For each 2×2 block in an image frame, F^k_{YUV}, the chroma subsampling strategy should be applied to sample the U and V components. Based on the analysis in Sec. 2, it is known that the V component is more significant than the U component when reconstructing the R component; the U component is more significant than the V component when reconstructing the B component. Thus, we can take the pixels whose corresponding pixels comprise G and B components to obtain the subsampling U components. On the other hand, the subsampling V components can be obtained from the pixels whose corresponding pixels comprise G and R components. Given a 2×2 block shown in Fig. 4, suppose in the original DTDI mosaic image frame, F^k_{DTDI}, as shown in Fig. 6, the corresponding pixels at positions (i, j) and $(i + 1, j)$ comprise G and B components, and those at positions $(i, j + 1)$ and $(i + 1, j + 1)$ comprise G and R components. Figures 7(a)–7(e) illustrate,

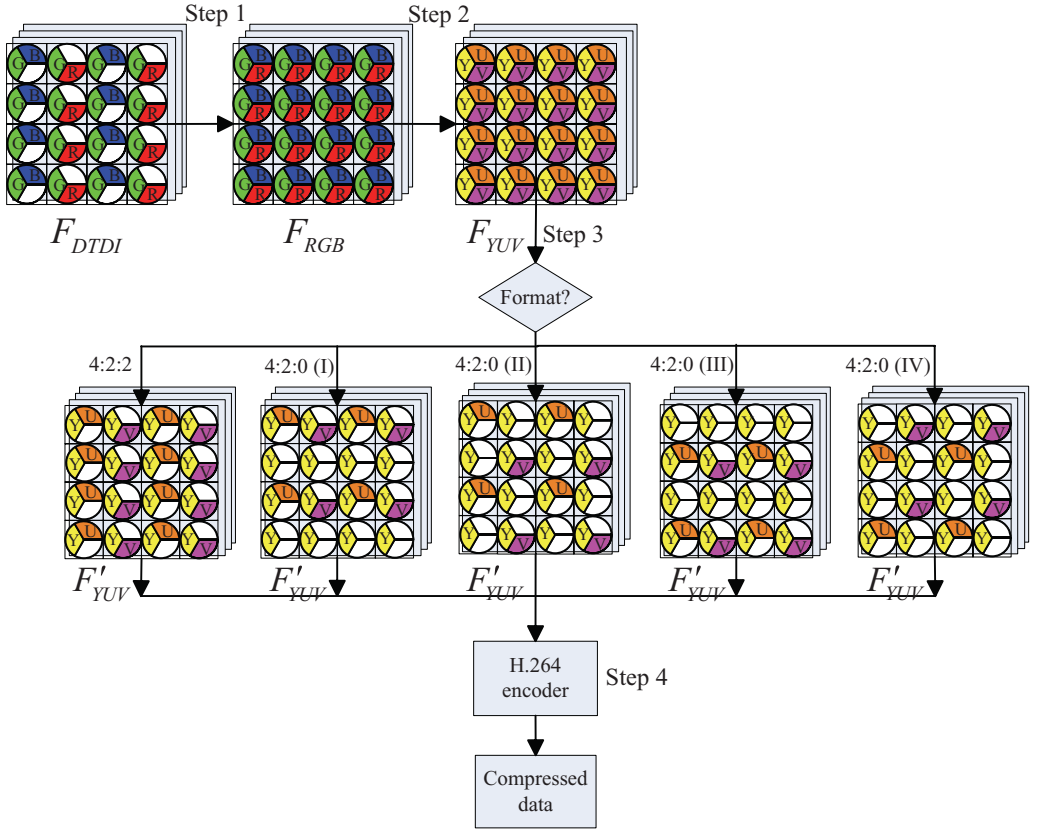


Fig. 8 The flowchart of the compression process.

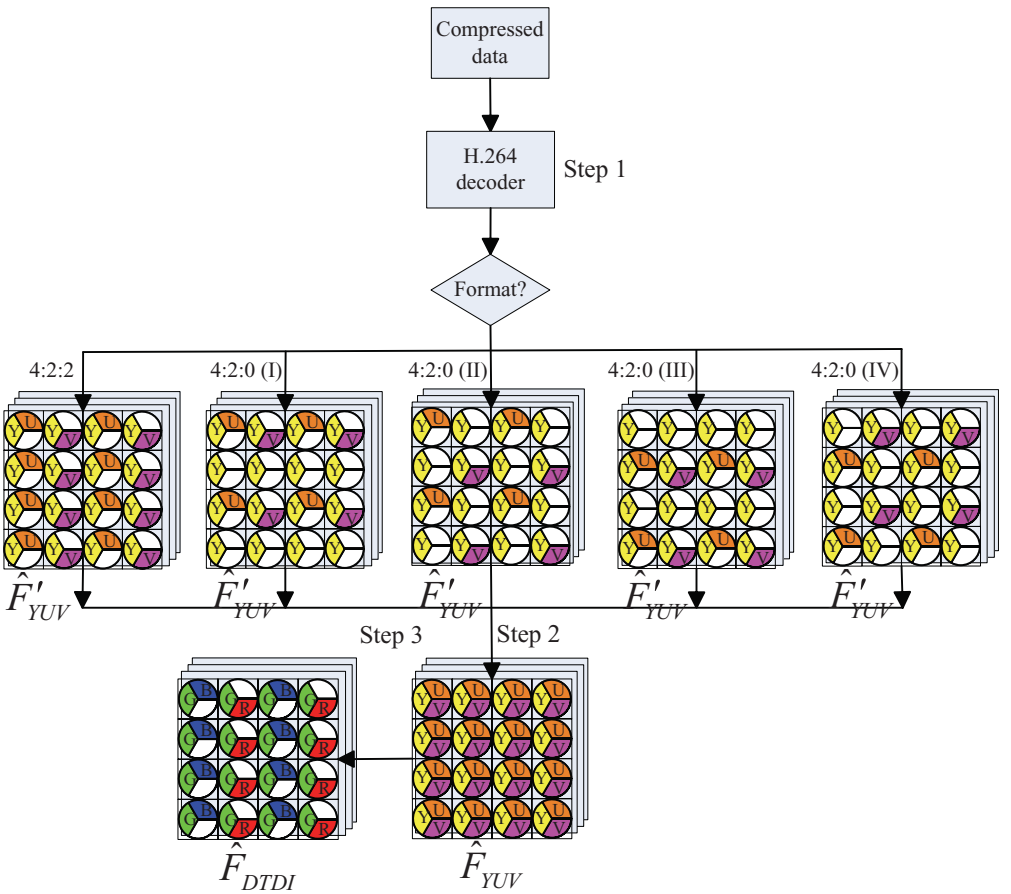


Fig. 9 The flowchart of the reconstruction process.

respectively, the depictions of a 4:2:2 format and four types of 4:2:0 formats, i.e., the 4:2:0(I)–(IV) formats, in the proposed modified chroma subsampling strategy.

From Fig. 7(a), it is obvious that for the 4:2:2 format, we take the two subsampling U components from the pixel at positions (i, j) and $(i+1, j)$ where the corresponding pixels in Fig. 6 comprise G and B components. Conversely, we take the two subsampling V components from the pixel at positions $(i, j+1)$ and $(i+1, j+1)$ where the corresponding pixels in Fig. 6 comprise G and R components. Therefore, in the reconstruction process, we have complete Y and U components to reconstruct the two corresponding B components of the pixels at positions (i, j) and $(i+1, j)$. Similarly, the two R components of the pixels at positions $(i, j+1)$ and $(i+1, j+1)$ can be reconstructed by the corresponding complete Y and V components. Because reconstructing R and B components will not create the bias problem, it would lead into constructing more accurate R and B components and achieving better quality of a decompressed DTDI video sequence.

For the four types of 4:2:0 formats shown in Figs. 7(b)–7(e), respectively, each 2×2 block in F^k_{YUV} comprises only one subsampling U component and one subsampling V component. Instead of using the average U and V components in the traditional chroma subsampling strategy, according to the structure of DTDI block shown in Fig. 6, we directly take one subsampling U component from the pixel at position (i, j) or $(i+1, j)$. On the other hand, we also take a subsampling V component from the pixel at position $(i, j+1)$ or $(i+1, j+1)$. Considering the case of the 4:2:0(I) format, we have complete Y and U components to reconstruct the corresponding B component of the pixel at positions (i, j) and the R component of the pixel at position $(i, j+1)$ can be reconstructed by complete Y and V components. This would lead into alleviating the bias problem and constructing more accurate R and B components. By the same argument, for the 4:2:0(II) format, the 4:2:0(III) format, and the 4:2:0(IV) format, the bias problem can also be alleviated and we can reconstruct more accurate R and B components. Therefore, the proposed modified chroma subsampling strategy would achieve better quality of a decompressed DTDI video sequence.

At present, we describe the complete compression and reconstruction processes. For ease of explanation, Figs. 8 and 9 illustrate the flowcharts of the compression and reconstruction processes, respectively. Given a DTDI video sequence, $F_{\mathrm{DTDI}} = \{F^k_{\mathrm{DTDI}} | \forall k \in \{1, 2, \ldots, n\}\}$, as shown in Fig. 8, the compression process involves four steps:

Step 1: Preform a DTDI demosaicing method[9] to obtain the demosaiced full color RGB video sequence, $F_{\mathrm{RGB}} = \{F^k_{\mathrm{RGB}} | \forall k \in \{1, 2, \ldots, n\}\}$.

Step 2: For each image frame in F_{RGB}, convert each pixel from the RGB color domain into the YUV color domain. Thus, we have a YUV video sequence, $F_{\mathrm{YUV}} = \{F^k_{\mathrm{YUV}} | \forall k \in \{1, 2, \ldots, n\}\}$.

(a) (b) (c) (d)

Fig. 10 Four test video sequences. (a) Ducks takeoff video sequence. (b) Old town cross video sequence. (c) Park joy video sequence. (d) Magazine content video sequence.

Step 3: For each image frame in F_{YUV}, apply the proposed modified chroma subsampling strategy to sample the U and V components based on the specified format. Then, we have a sampled YUV video sequence, $F'_{\mathrm{YUV}} = \{F'^{k}_{\mathrm{YUV}} | \forall k \in \{1, 2, \ldots, n\}\}$

Step 4: Compress the sampled YUV video sequence F'_{YUV} to obtain compressed data by using an H.264/AVC encoder.

As shown in Fig. 9, given a compressed data, the reconstruction process is comprised of three steps:

Step 1: Decompress the input compressed data to obtain the sampled YUV video sequence, $\hat{F}'_{\mathrm{YUV}} = \{\hat{F}'^{k}_{\mathrm{YUV}} | \forall k \in \{1, 2, \ldots, n\}\}$ by using an H.264/AVC decoder.

Step 2: For each image frame in $\hat{F}'_{\mathrm{YUV}}$, apply the interpolation technique to obtain the fully populated YUV video sequence, $\hat{F}_{\mathrm{YUV}} = \{\hat{F}^{k}_{\mathrm{YUV}} | \forall k \in \{1, 2, \ldots, n\}\}$.

Step 3: For each image frame in $\hat{F}_{\mathrm{YUV}}$, convert the necessary color information in accordance with DTDI structure from the YUV color domain into the RGB color

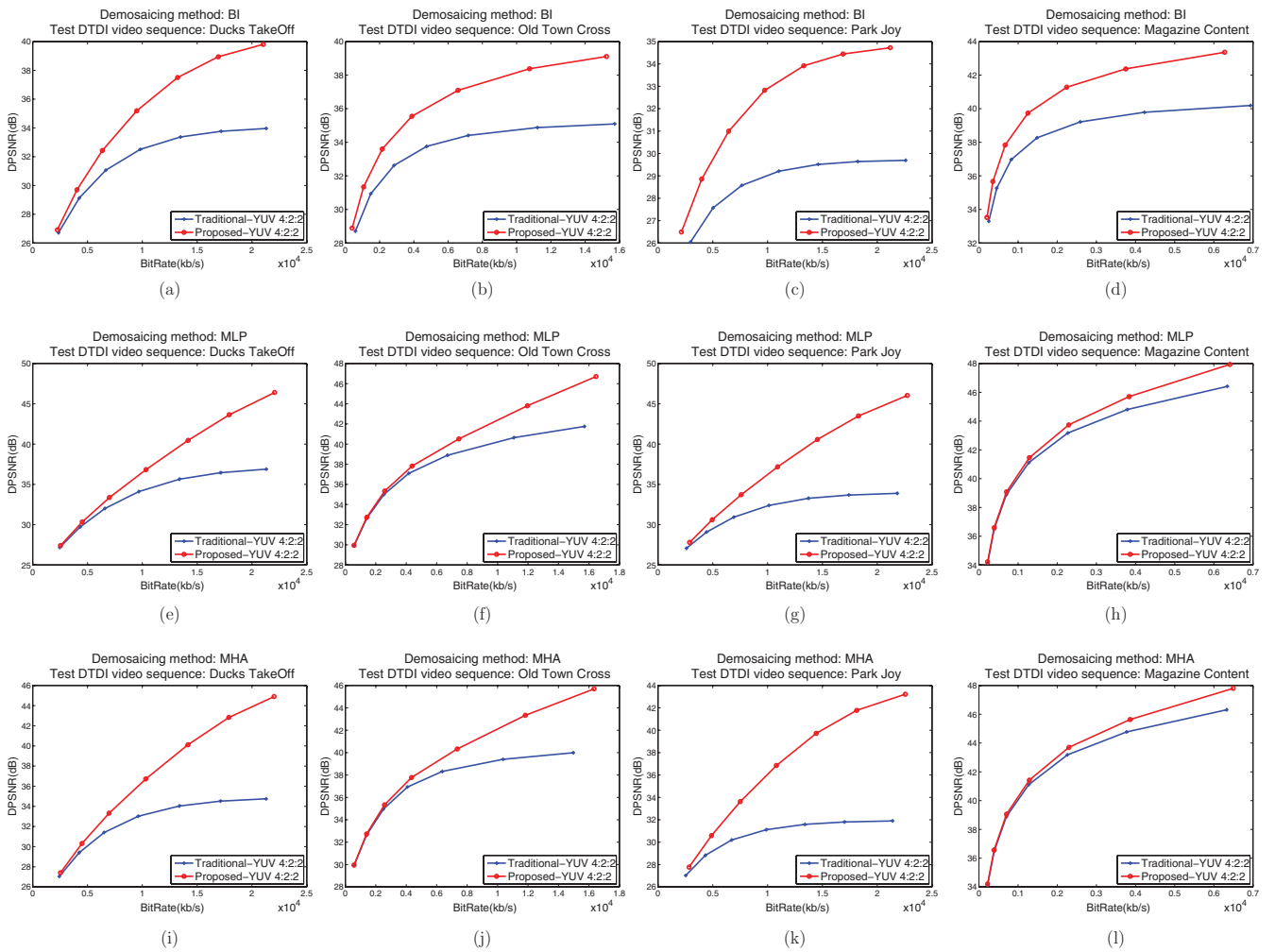


Fig. 11 Using the three concerned demosaicing methods in the compression process, the distortion versus the bit rate curves of the two concerned chroma subsampling strategies based on the 4:2:2 format for the four test video sequences.

Table 1 Using the three concerned demosaicing methods in the compression process, the average DPSNR comparison results between the two concerned chroma subsampling strategies based on the 4:2:2 format under seven different QPs.

	QP						
Format	8	12	16	20	24	28	32
Demosaicing method: BI							
Traditional 4:2:2	34.7344	34.5174	34.1267	33.4360	32.3053	30.7222	28.6871
Proposed 4:2:2	39.2501	38.5357	37.4477	35.8233	33.7160	31.3918	28.9533
Demosaicing method: MLP							
Traditional 4:2:2	39.7363	38.8913	37.7465	36.1778	34.2074	31.9689	29.5738
Proposed 4:2:2	46.7699	44.1542	41.3278	38.3209	35.3756	32.5732	29.8470
Demosaicing method: MHA							
Traditional 4:2:2	38.2388	37.6234	36.7835	35.5447	33.8647	31.8321	29.5318
Proposed 4:2:2	45.4031	43.3921	40.9659	38.1906	35.3325	32.5536	29.8318

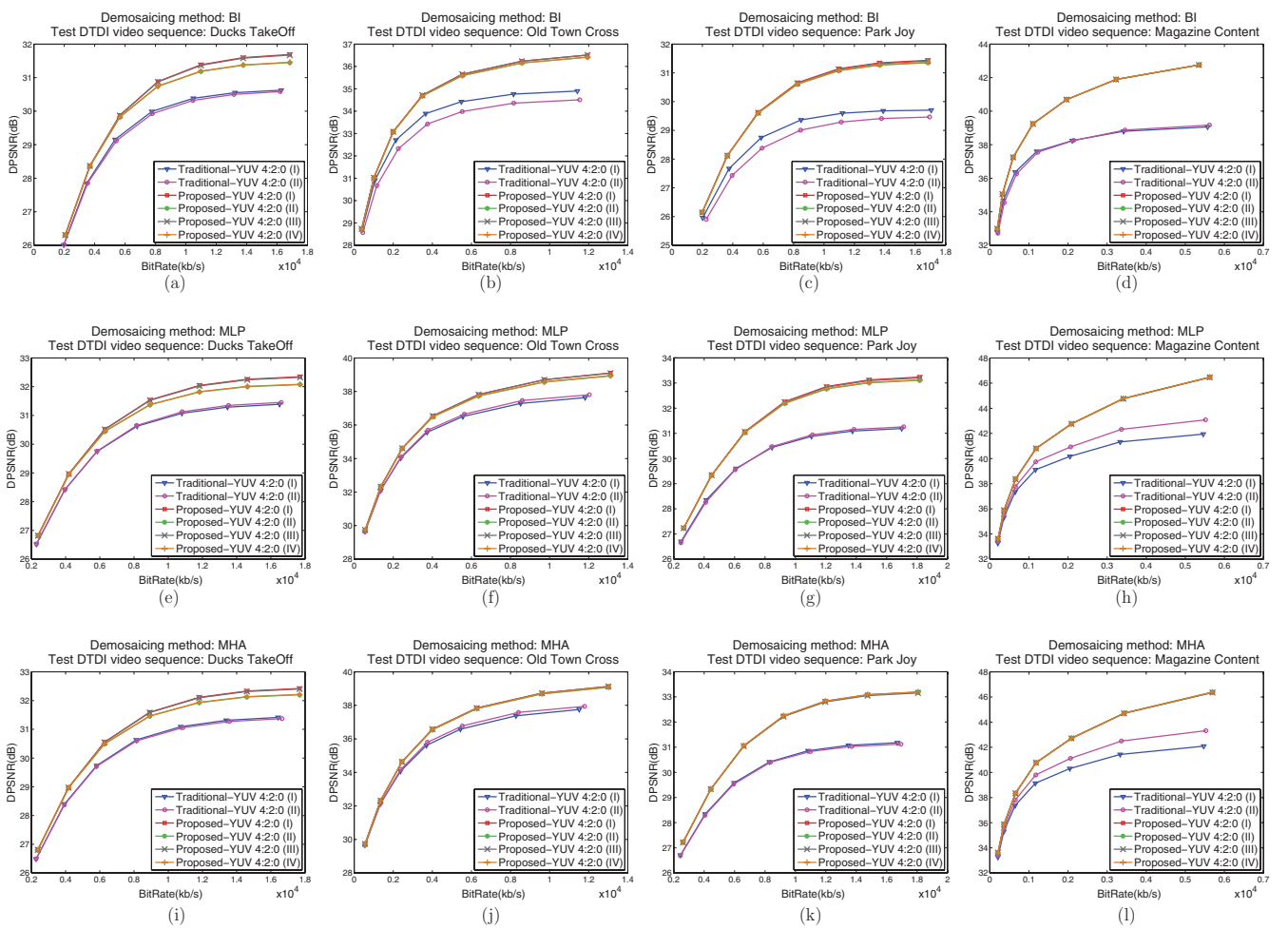


Fig. 12 Using the three concerned demosaicing methods in the compression process, the distortion versus the bit rate curves of the two concerned chroma subsampling strategies based on the 4:2:0 format for the four test video sequences.

Table 2 Using the three concerned demosaicing methods in the compression process, the average DPSNR comparison results between the two concerned chroma subsampling strategies based on the 4:2:0 format under seven different QPs.

	QP						
Format	8	12	16	20	24	28	32
Demosaicing method: BI							
Traditional 4:2:0(I)	34.5542	33.4485	33.1616	32.7108	31.7429	30.2719	28.3651
Traditional 4:2:0(II)	34.3805	33.2859	32.9601	32.4785	31.5171	30.1242	28.3011
Proposed 4:2:0(I)	36.9030	35.2682	34.7163	33.8744	32.4623	30.6408	28.5408
Proposed 4:2:0(II)	36.8411	35.1721	34.6345	33.8175	32.4347	30.6294	28.5380
Proposed 4:2:0(III)	36.8907	35.2567	34.7070	33.8665	32.4526	30.6389	28.5411
Proposed 4:2:0(IV)	36.8484	35.1790	34.6409	33.8221	32.4407	30.6385	28.5343
Demosaicing method: MLP							
Traditional 4:2:0(I)	36.9236	35.2478	34.6591	33.9349	32.6848	31.0416	29.0359
Traditional 4:2:0(II)	37.3809	35.5725	34.9113	34.1417	32.7926	31.0777	29.0599
Proposed 4:2:0(I)	39.6018	37.2122	36.3716	35.2841	33.6444	31.6309	29.3502
Proposed 4:2:0(II)	39.5035	37.0873	36.2705	35.2146	33.6074	31.6196	29.3534
Proposed 4:2:0(III)	39.5840	37.1978	36.3593	35.2756	33.6369	31.6246	29.3503
Proposed 4:2:0(IV)	39.5132	37.0954	36.2748	35.2211	33.6183	31.6217	29.3516
Demosaicing method: MHA							
Traditional 4:2:0(I)	37.0064	35.2977	34.7091	33.9490	32.6888	31.0410	29.0306
Traditional 4:2:0(II)	37.4597	35.5953	34.9451	34.1490	32.8061	31.0941	29.0649
Proposed 4:2:0(I)	39.5608	37.2124	36.3736	35.2954	33.6488	31.6238	29.3396
Proposed 4:2:0(II)	39.5496	37.1504	36.3194	35.2537	33.6259	31.6229	29.3383
Proposed 4:2:0(III)	39.5449	37.1976	36.3601	35.2887	33.6401	31.6211	29.3432
Proposed 4:2:0(IV)	39.5585	37.1590	36.3253	35.2627	33.6344	31.6216	29.3408

domain. Thus, we have the reconstructed DTDI video sequence, $\hat{F}_{\mathrm{DTDI}} = \{\hat{F}^k_{\mathrm{DTDI}} | \forall k \in \{1, 2, \ldots, n\}\}$.

Referring to Fig. 9, it is observed that we can obtain the full color video sequence from the fully populated YUV video sequence, $\hat{F}_{\mathrm{YUV}}$. In general, the full color video sequence generated by $\hat{F}_{\mathrm{YUV}}$ has good quality performance, but we adopt the DTDI video sequences to be the output format of the reconstruction process. The benefits are twofold. First, we can keep the video format consistency between the compression process and the reconstruction process. Second, in our experiments, the demosaiced full color image reconstructed from $\hat{F}_{\mathrm{DTDI}}$ usually has better quality performance than the one generated by $\hat{F}_{\mathrm{YUV}}$.

After describing the whole compression and reconstruction processes, in Sec. 4, we report the results of experiments to demonstrate the quality advantage of the proposed chroma subsampling strategy.

4 Experimental Results

To test the effectiveness of the proposed chroma subsampling strategy, we conducted experiments on four test video sequences shown in Fig. 10. For the test video sequences, Figs. 10(a)–10(c), adopted from the website,[24] are taken to show the human vision application and Fig. 10(d) with some noticeable texts is used to clarify the application of the print inspection. The four test video sequences were first downsampled to obtain the DTDI mosaic video sequences. Then, we ran the compression and reconstruction processes with the traditional chroma subsampling strategy and those with the proposed modified chroma subsampling strategy on the four test DTDI mosaic video sequences. For the four test DTDI mosaic video sequences, the original size of each image frame was 360×288. In our experiments, a group of pictures consisted of one intraframe and nine interframes. In addition, seven different quantization parameters (QPs), i.e., 8, 12, 16, 20, 24, 28, and 32, were considered in the compres-

sion process. All the concerned methods in the experiments were implemented on the IBM compatible computer with Intel Core 2 Duo E7500 CPU 2.93 GHz and 2GB RAM. The operating system used was Microsoft Windows 7. The demosaicing, subsampling, and upsampling processes were implemented in Borland C++ Builder 6.0. The compression platform was JM14.2,[25] which is realized in Visual C++ 2005.

The three DTDI demosaicing methods used in our experiments were bilinear interpolation (BI), modified Laroche and Prescott's (MLP) method,[26] and modified Hamilton and Adams' (MHA) method.[27] The three demosaicing method can be found in Ref. 9. For completeness, a brief summary of the three DTDI demosaicing methods is introduced. The interpolation of R pixel values and B pixel values is the same; thus, we only consider the case of R pixel values. BI is the simplest demosaicing algorithm in which the unknown R pixel value, $F^{k,r}_{\mathrm{RGB}}(i,j)$, is obtained by

$$F^{k,r}_{\mathrm{RGB}}(i,j)=\frac{1}{2}\sum_{k=\{\pm 1\}}F^{k,r}_{\mathrm{DTDI}}(i,j+k).$$

Based on the MLP method, the unknown R pixel value can be calculated by

$$F^{k,r}_{\mathrm{RGB}}(i,j)=F^{k,g}_{\mathrm{DTDI}}(i,j)+\frac{1}{2}\left[\sum_{k=\{\pm 1\}}F^{k,r}_{\mathrm{DTDI}}(i,j+k)-F^{k,g}_{\mathrm{DTDI}}(i,j+k)\right].$$

According to the concept of the MHA method, we can estimate the unknown R pixel value by the following rule:

$$F^{k,r}_{\mathrm{RGB}}(i,j)=\begin{cases}F^{k,r}_{H}(i,j) & \text{if } \Delta D_H(i,j)+\xi<\min\{\Delta D_{\pi/4}(i,j),\Delta D_{-\pi/4}(i,j)\}\\ F^{k,r}_{\pi/4}(i,j) & \text{if } \Delta D_{\pi/4}(i,j)+\xi<\min\{\Delta D_H(i,j),\Delta D_{-\pi/4}(i,j)\}\\ F^{k,r}_{-\pi/4}(i,j) & \text{if } \Delta D_{-\pi/4}(i,j)+\xi<\min\{\Delta D_H(i,j),\Delta D_{\pi/4}(i,j)\}\\ F^{k,r}_{O}(i,j) & \text{otherwise,}\end{cases}$$

where

$$F^{k,r}_{H}(i,j)=F^{k,g}_{\mathrm{DTDI}}(i,j)+\frac{1}{2}\left[\sum_{k=\{\pm 1\}}F^{k,r}_{\mathrm{DTDI}}(i,j+k)-F^{k,g}_{\mathrm{DTDI}}(i,j+k)\right],$$

$$F^{k,r}_{\frac{\pi}{4}}(i,j)=F^{k,g}_{\mathrm{DTDI}}(i,j)+\frac{1}{2}\left[\sum_{k=\{\pm 1\}}F^{k,r}_{\mathrm{DTDI}}(i+k,j-k)-F^{k,g}_{\mathrm{DTDI}}(i+k,j-k)\right],$$

$$F^{k,r}_{\frac{-\pi}{4}}(i,j)=F^{k,g}_{\mathrm{DTDI}}(i,j)+\frac{1}{2}\left[\sum_{k=\{\pm 1\}}F^{k,r}_{\mathrm{DTDI}}(i+k,j+k)-F^{k,g}_{\mathrm{DTDI}}(i+k,j+k)\right],$$

$$F^{k,r}_{O}(i,j)=\frac{1}{3}\sum_{d\in\{H,\pi/4,-\pi/4\}}F^{k,r}_{d}(i,j),$$

$$\Delta D_H(i,j)=\left|2F^{k,g}_{\mathrm{DTDI}}(i,j)-\sum_{k=\{\pm 1\}}F^{k,g}_{\mathrm{DTDI}}(i,j+k)\right|+\left|\sum_{k=\{\pm 1\}}kF^{k,r}_{\mathrm{DTDI}}(i,j+k)\right|,$$

$$\Delta D_{\frac{\pi}{4}}(i,j)=\left|2F^{k,g}_{\mathrm{DTDI}}(i,j)-\sum_{k=\{\pm 1\}}F^{k,g}_{\mathrm{DTDI}}(i+k,j-k)\right|+\left|\sum_{k=\{\pm 1\}}kF^{k,r}_{\mathrm{DTDI}}(i+k,j-k)\right|,$$

$$\Delta D_{\frac{-\pi}{4}}(i,j)=\left|2F^{k,g}_{\mathrm{DTDI}}(i,j)-\sum_{k=\{\pm 1\}}F^{k,g}_{\mathrm{DTDI}}(i+k,j+k)\right|+\left|\sum_{k=\{\pm 1\}}kF^{k,r}_{\mathrm{DTDI}}(i+k,j+k)\right|,$$

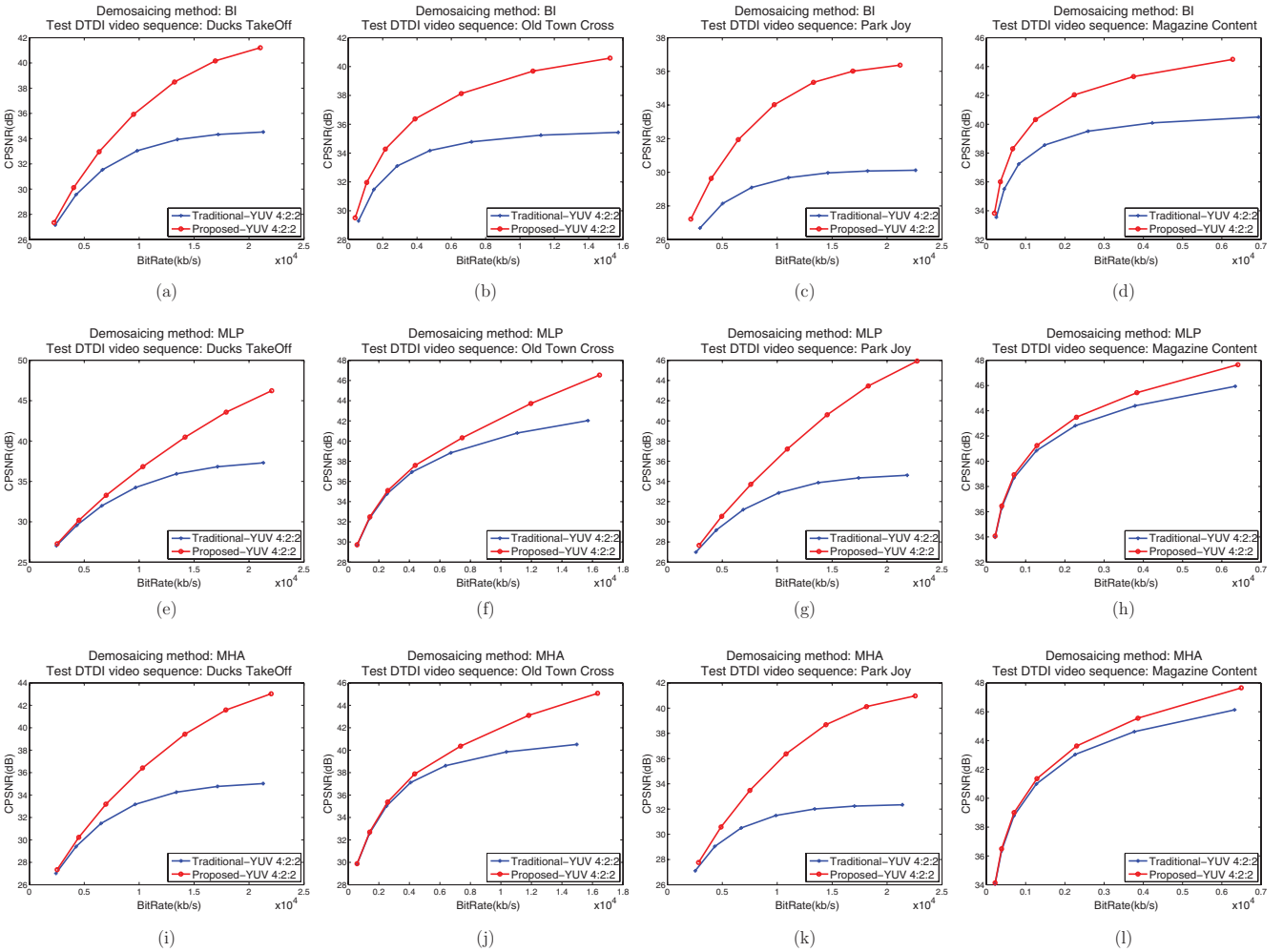


Fig. 13 After demosaicing the reconstructed DTDI video sequences by the three concerned demosaicing methods, the distortion versus the bit rate curves of the two concerned chroma subsampling strategies based on the 4:2:2 format for the four test video sequences.

Table 3 After demosaicing the reconstructed DTDI video sequences by the three concerned demosaicing methods, the average CPSNR comparison results between the two concerned chroma subsampling strategies based on the 4:2:2 format under seven different QPs.

	QP						
Format	8	12	16	20	24	28	32
Demosaicing method: BI							
Traditional 4:2:2	35.1461	34.9368	34.5475	33.8621	32.7439	31.1696	29.1694
Proposed 4:2:2	40.6643	39.7916	38.5008	36.6581	34.3705	31.9291	29.4756
Demosaicing method: MLP							
Traditional 4:2:2	39.9691	39.0901	37.8677	36.2193	34.1593	31.8415	29.4160
Proposed 4:2:2	46.5878	44.0439	41.2231	38.2179	35.2539	32.4190	29.6883
Demosaicing method: MHA							
Traditional 4:2:2	38.5015	37.8663	36.9782	35.6924	33.9460	31.8479	29.5008
Proposed 4:2:2	44.1830	42.5902	40.5222	38.0066	35.2583	32.4988	29.7781

Table 4 After demosaicing the reconstructed DTDI video sequences by the three concerned demosaicing methods, the average CPSNR comparison results between the two concerned chroma subsampling strategies based on the 4:2:0 format under seven different QPs.

	QP						
Format	8	12	16	20	24	28	32
Demosaicing method: BI							
Traditional 4:2:0(I)	33.8090	33.6956	33.4083	32.9805	32.0306	30.5774	28.7313
Traditional 4:2:0(II)	33.5643	33.4364	33.1206	32.6764	31.7541	30.4046	28.6581
Proposed 4:2:0(I)	36.1269	35.7628	35.1639	34.2948	32.8468	31.0061	28.9504
Proposed 4:2:0(II)	36.0465	35.6911	35.1024	34.2516	32.8240	30.9939	28.9472
Proposed 4:2:0(III)	36.1097	35.7488	35.1527	34.2846	32.8353	31.0020	28.9490
Proposed 4:2:0(IV)	36.0559	35.6995	35.1101	34.2563	32.8300	31.0032	28.9420
Demosaicing method: MLP							
Traditional 4:2:0(I)	35.3416	35.0541	34.4545	33.7322	32.4659	30.8042	28.8047
Traditional 4:2:0(II)	35.6557	35.3443	34.6798	33.9235	32.5646	30.8344	28.8247
Proposed 4:2:0(I)	37.5697	37.0022	36.1572	35.0847	33.4464	31.4180	29.1433
Proposed 4:2:0(II)	37.5258	36.9633	36.1243	35.0591	33.4259	31.4105	29.1483
Proposed 4:2:0(III)	37.5490	36.9853	36.1436	35.0756	33.4373	31.4107	29.1413
Proposed 4:2:0(IV)	37.5362	36.9717	36.1287	35.0649	33.4357	31.4100	29.1441
Demosaicing method: MHA							
Traditional 4:2:0(I)	35.6507	35.3424	34.7320	33.9595	32.6569	30.9640	28.9307
Traditional 4:2:0(II)	36.0066	35.6672	34.9889	34.1788	32.7872	31.0231	28.9660
Proposed 4:2:0(I)	37.7988	37.2344	36.3755	35.2925	33.6180	31.5560	29.2551
Proposed 4:2:0(II)	37.7571	37.1915	36.3360	35.2626	33.5998	31.5578	29.2551
Proposed 4:2:0(III)	37.7885	37.2250	36.3653	35.2885	33.6114	31.5546	29.2598
Proposed 4:2:0(IV)	37.7683	37.2015	36.3437	35.2718	33.6085	31.5552	29.2572

ξ is a specific threshold. In general, the MLP method and the MHA method can yield better quality performance of demosaiced images than BI.

To compare the quality performance between the traditional chroma subsampling strategy and the proposed modified chroma subsampling strategy, we used DTDI peak signal-to-noise ratio (DPSNR) to test the quality of the reconstructed DTDI video sequences. The DPSNR of a DTDI mosaic video sequence with N frames, each with size $X \times Y$, is defined as follows:

$$\mathrm{DPSNR} = 10 \log_{10} \frac{255^2}{\frac{1}{N}\sum_{k=1}^{N} \mathrm{DFMSE}^{\mathrm{k}}}$$

$$\mathrm{DFMSE}^{\mathrm{k}} = \frac{1}{2XY} \sum_{i=0}^{X-1} \sum_{j=0}^{Y-1} \sum_{c \in \Psi} \delta_j^c \left[F_{\mathrm{DTDI}}^{k,c}(i,j) - \hat{F}_{\mathrm{DTDI}}^{k,c}(i,j) \right]^2$$

$$\Psi = \{r, g, b\}; \quad \delta_j^g = 1; \quad \delta_j^r = j \bmod 2; \quad \delta_j^b = (j+1) \bmod 2,$$

where $F_{\mathrm{DTDI}}^{k,r}(i,j)$, $F_{\mathrm{DTDI}}^{k,g}(i,j)$, and $F_{\mathrm{DTDI}}^{k,b}(i,j)$ denote, respectively, the R, G, and B color values of the pixel at position (i,j) in the k'th image frame of the original DTDI mosaic video sequence; $\hat{F}_{\mathrm{DTDI}}^{k,r}(i,j)$, $\hat{F}_{\mathrm{DTDI}}^{k,g}(i,j)$, and $\hat{F}_{\mathrm{DTDI}}^{k,b}(i,j)$ denote, respectively, the R, G, and B color values of the pixel at position (i,j) in the k'th image frame of the reconstructed DTDI mosaic video sequence; DFMSE is the DTDI frame mean square error. The larger the DPSNR, the better will be the video sequence quality.

Based on the four test DTDI mosaic video sequences, Fig. 11 illustrates the distortion versus the bit rate curves generated by the traditional chroma subsampling strategy and the proposed modified chroma subsampling strategy. In Fig. 11, the x-axis denotes the total bit rate size at a rate of 25 frames/s and the y-axis denotes the DPSNR value. Based on the 4:2:2 format and seven different QPs, Fig. 11 demonstrates the quality performance comparisons between two concerned chroma subsampling strategies using three concerned DTDI

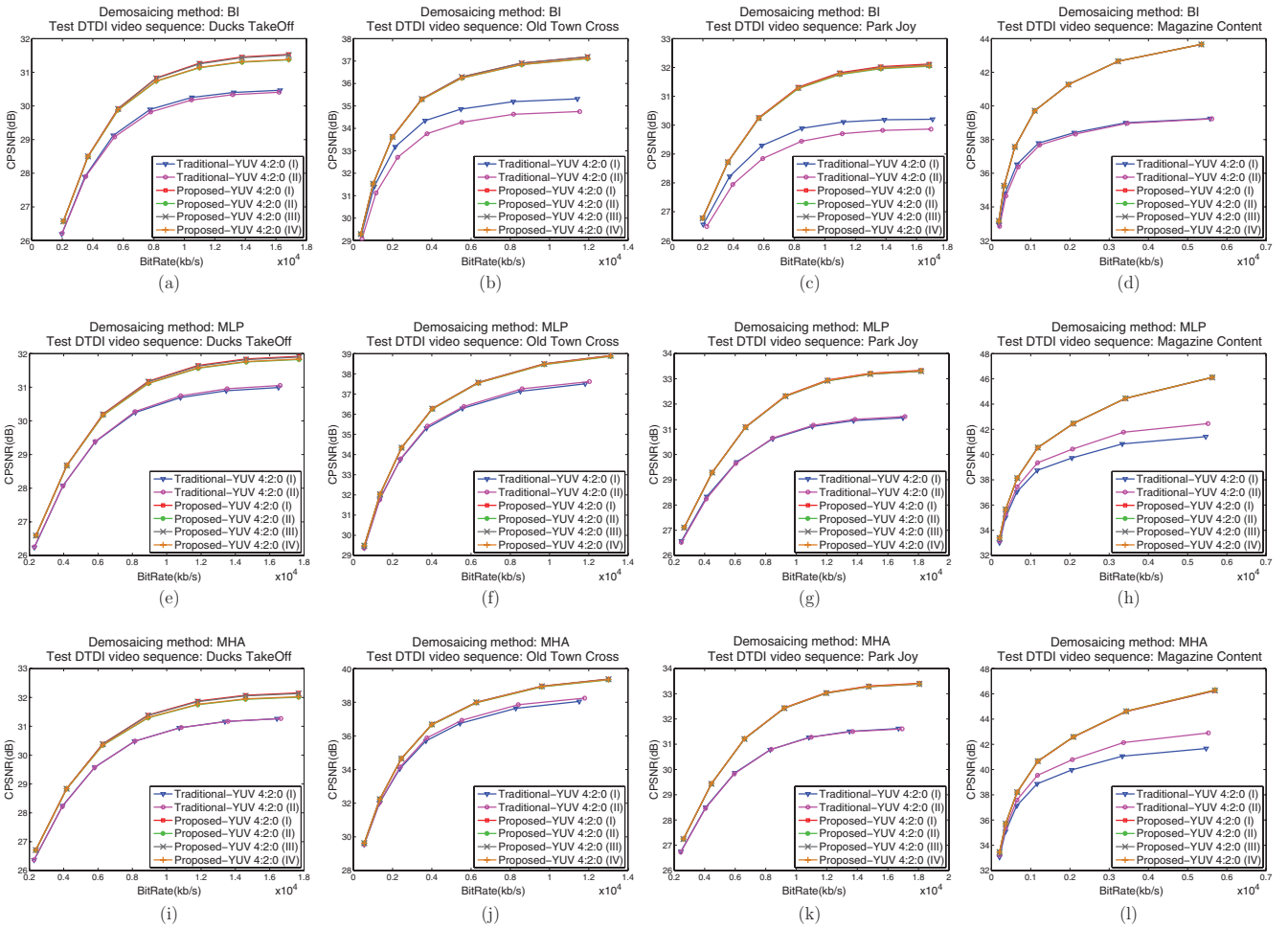


Fig. 14 After demosaicing the reconstructed DTDI video sequences by the three concerned demosaicing methods, the distortion versus the bit rate curves of the two concerned chroma subsampling strategies based on the 4:2:0 format for the four test video sequences.

demosaicing methods in the compression process. From Fig. 11, it is clear that under constraints of equal bit rate and the same demosaicing method, the proposed modified chroma subsampling strategy produced consistently better video sequence quality in terms of DPSNR when compared with the traditional chroma subsampling strategy. Then, based on the four test DTDI mosaic sequences, Table 1 demonstrates the average quality comparisons in terms of DPSNR under seven different QPs. From these tables, it indicates that under the same QP constraint, the proposed modified chroma subsampling strategy yields better average quality performance than the traditional one. In addition, using the MLP method in the compression process, the proposed modified chroma subsampling strategy can produce the best average video sequence quality performance. Based on the 4:2:0 format, similar comparisons between the two concerned chroma subsampling strategies are demonstrated in Fig. 12 and Table 2. From the figures, it is clear that under constraints of equal bit rate and the same demosaicing method, the four types of the 4:2:0 formats of the proposed modified chroma subsampling strategy yields similar video sequence quality performance, but they produce a better video sequence quality performance than the two types of the 4:2:0 formats of the traditional chroma subsampling strategy. From Table 2, it also reveals that under the same QP constraint, the proposed modified chroma subsampling strategy has better average quality performance than the traditional chroma subsampling strategy. In addition, referring to Tables 1 and 2, it is obvious that for the 4:2:2 format, using the MLP method in the compression process can usually yield better quality performance of a reconstructed DTDI video sequence than the other two demosaicing methods; for the 4:2:0 format, using the MHA method in the compression process can usually produce the best DTDI video sequence quality. In addition to evaluating the video sequence quality performance under the DTDI mosaic domain, we further evaluated the video sequence quality performance under the demosaiced full color domain. In our experiments, we considered the real application of users. In actuality, the demosaiced video sequence obtained from the DTDI sequence, i.e., F_{RGB} in Fig. 8, is the full color video sequence which can be obtained by users. Further, this work principally focuses on the quality degradation caused by the compression process, so the distortion induced by Step 1 in Fig. 8 is ignored. According to the above reasons, we adopted F_{RGB} in Fig. 8 to be the reference video sequence.

We still adopted the BI, the MLP method, and the MHA to demosaic the reconstructed DTDI video sequences. To fit the

(a) (b) (c)

Fig. 15 The magnified subimage frames cut from the ducks takeoff video sequence. (a) The one generated by demosaicing the original DTDI video sequence. The ones generated by demosaicing the reconstructed DTDI video sequences based on (b) the traditional chroma subsampling strategy and (c) the proposed chroma subsampling strategy. [Format: 4:2:0(I); demosaicing method: MHA; QP: 28.]

demosaiced full color domain, we used color peak signal-to-noise ratio (CPSNR) to justify the better quality performance of the proposed modified chroma subsampling strategy in the full color domain. The CPSNR of a full color video sequence with N frames, each with size $X \times Y$, is defined as follows:

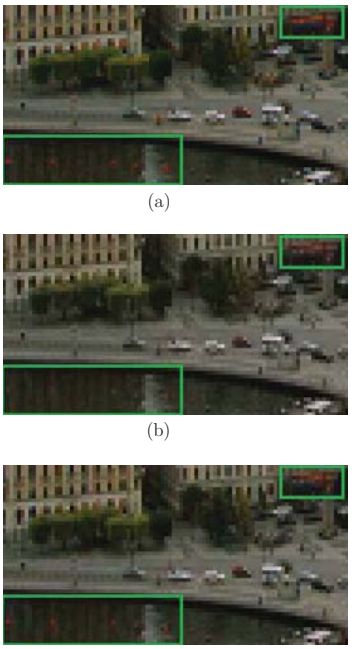
$$\mathrm{CPSNR} = 10 \log_{10} \frac{255^2}{\frac{1}{N}\sum_{k=1}^{N} \mathrm{CFMSE}^{\mathrm{k}}}$$

$$\mathrm{CFMSE}^{\mathrm{k}} = \frac{1}{3XY} \sum_{i=0}^{X-1} \sum_{j=0}^{Y-1} \sum_{c \in \Psi} [F_{\mathrm{RGB}}^{k,c}(i,j) - \hat{F}_{\mathrm{RGB}}^{k,c}(i,j)]^2,$$

$$\Psi = \{r, g, b\},$$

where $F_{\mathrm{RGB}}^{k,r}(i,j)$, $F_{\mathrm{RGB}}^{k,g}(i,j)$, and $F_{\mathrm{RGB}}^{k,b}(i,j)$ denote, respectively, the R, G, and B color values of the pixel at position (i,j) in the k'th image frame of the reference full color video sequence; $\hat{F}_{\mathrm{RGB}}^{k,r}(i,j)$, $\hat{F}_{\mathrm{RGB}}^{k,g}(i,j)$, and $\hat{F}_{\mathrm{RGB}}^{k,b}(i,j)$ denote, respectively, the R, G, and B color values of the pixel at position (i,j) in the k'th image frame of the demosaiced full color video sequence; CFMSE is the color frame mean square error. The larger the CPSNR, the better will be the video sequence quality.

Based on the same test videos, among the six schemes, which combine one of the two concerned chroma subsampling strategies and one of the three existing demosaicing methods, Figs. 13 and 14 illustrate the distortion versus the bit rate curves. In the figures, the x-axis denotes the total bit rate size at a rate of 25 frames/s and the y-axis denotes the CPSNR value. Similar to the quality performance comparison for DTDI mosaic sequences, Figs. 13 and 14 convincingly indicate that under the constraints of equal bit rate and the same demosaicing method, the proposed modified chroma subsampling strategy yields consistently better demosaiced full color video sequence quality in terms of CPSNR than the traditional chroma subsampling strategy. Then, Table 3 demonstrates the average CPSNR comparison results between the two concerned chroma subsampling strategies under seven different QPs. From the tables, it is obvious that under the same QP constraint, the proposed modified chroma subsampling strategy produces better average quality performance than the traditional one. Further, from Tables 3 and 4, it indicates that for the 4:2:2 format, demosaicing the reconstructed DTDI video sequences by the MLP method can usually yield better quality performance of a demosaiced full color video sequence; for the 4:2:0 format, using the MHA method to demosaic the reconstructed DTDI video sequences can usually produce the best demosaiced video sequence quality.



(a) (b) (c)

Fig. 16 The magnified subimage frames cut from the old town cross video sequence. (a) The one generated by demosaicing the original DTDI video sequence. The ones generated by demosaicing the reconstructed DTDI video sequences based on (b) the traditional chroma subsampling strategy and (c) the proposed chroma subsampling strategy. [Format: 4:2:0(I); demosaicing method: MHA; QP: 20.]

Finally, we used the subjective visual measure to demonstrate the visual quality advantage of the proposed chroma subsampling strategy. After performing subsampling and compression processes, some texture loss may arise. We first took the magnified subimage frames cut from the ducks takeoff video sequence to compare the human visual effect. Figure 15 illustrates the magnified subimage frame cut from the full color image frame generated by demosaicing the original DTDI video sequence and the ones generated by demosaicing the reconstructed DTDI video sequences based on two concerned chroma subsampling strategies. From the duck's leg part surrounded by the green box in Fig. 15(a) and the ones in Figs. 15(b) and 15(c), we observe that the subimage frame based on the proposed chroma subsampling strategy yields less texture loss when compared with the one based on the traditional chroma subsampling strategy. Then, we took the magnified subimages frame cut from old town cross video sequence to depict the visual comparison. Figure 16 illustrates the magnified subimage frame cut from the demosaiced image frame generated by demosaicing the original DTDI video sequence and the ones generated by demosaicing the reconstructed DTDI video sequences based on two concerned chroma subsampling strategies. From the floating lamps and the signboard surrounded by green boxes, respectively, in Fig. 16, it is observed that the proposed chroma subsampling strategy has less texture loss and produces a better visual effect. More visual results of

the concerned chroma subsampling strategies are available in Ref. 28.

5 Conclusions

In this paper, we have presented a modified chroma subsampling strategy for compressing DTDI mosaic video sequences in H.264/AVC. The difference between the proposed strategy and the traditional strategy in H.264/AVC is that the proposed strategy uses the color domain transform between the RGB domain and the YUV domain to select the proper positions to obtain the two subsampling U and V components. It would lead into achieving better quality of decompressed DTDI video sequences. Based on some popular test DTDI mosaic video sequences, experimental results demonstrate convincingly the quality advantage of the proposed chroma subsampling strategy. In addition, under the demosaiced full-color domain, the proposed chroma subsampling strategy still yields better quality performance of video sequences in terms of CPSNR when compared with the traditional one in H.264/AVC. It is an interesting research topic to combine demosaicing, subsampling, and compression into a single step to improve the compression performance even more.

References

1. B. E. Bayer, "Color imaging array," US Patent No. 3,971,065 (1976).
2. R. Lukac and K. N. Plataniotis, "Color filter arrays: design and performance analysis," *IEEE Trans. Consum. Electron.* **51**(4), 1260–1267 (2005).
3. S. C. Pei and I. K. Tam, "Effective color interpolation in CCD color filter arrays using signal correlation," *IEEE Trans. Circuits Syst. Video Technol.* **13**(6), 503–513 (2003).
4. R. Ramanath, W. E. Snyder, G. L. Bilbro, and W. A. Sander III, "Demosaicking methods for Bayer color arrays," *J. Electron. Imaging* **11**(3), 306–315 (2002).
5. E. Bodenstorfer, J. Fürtler, J. Brodersen, K. J. Mayer, C. Eckel, K. Gravogel, and H. Nachtnebel, "High speed line-scan camera with digital time delay integration," *Proc. SPIE* **6496**, 64960I (2007).
6. P. E. Debevec and J. Malik, "Recovering high dynamic range radiance maps from photographs," in *Proc. of SIGGRAPH*, pp. 369–378 (1997).
7. H. S. Wong, Y. L. Yao, and E. S. Schlig, "TDI charge-coupled devices: design and applications," *IBM J. Res. Dev.* **36**(1), 83–105 (1992).
8. J. Fürtler, E. Bodenstorfer, K. J. Mayer, J. Brodersen, D. Heiss, H. Penz, C. Eckel, K. Gravogel, and H. Nachtnebel, "High performance camera module for fast quality inspection in industrial printing applications," *Proc. SPIE* **6503**, 65030J (2007).
9. D. Heiss-Czedik, R. Huber-Mörk, D. Soukup, H. Penz, and B. L. García, "Demosaicing algorithms for area- and line-scan cameras in print inspection," *J. Visual Commun. Image Represent.*, **20**(6), 389–398 (2009).
10. Draft ITU-T recommendation and final draft international standard of joint video specification (ITU-T Rec. H.264/ISO/IEC 14 496-10 AVC,) *Joint Video Team of ISO/IEC and ITU-T* (2003).
11. J. Ostermann, J. Bormans, P. List, D. Marpe, M. Narroschke, F. Pereira, T. Stockhammer, and T. Wedi, "Video coding with H.264/AVC: tools, performance, and complexity," *IEEE Circuits Syst. Mag.* **4**(1), 7–28 (2004).
12. I. E. Richardson, *H.264 and MPEG-4 Video Compression: Video Coding for Next-generation Multimedia*, Wiley, New York (2003).
13. T. Wiegand, G. J. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Trans. Circuits Syst. Video Technol.* **13**(7), 560–576 (2003).
14. A. Bazhyna and K. Egiazarian, "Lossless and near lossless compression of real color filter array data," *IEEE Trans. Consum. Electron.*, **54**(4), 1492–1500 (2008).
15. H. Chen, M. Sun, and E. Steinbach, "Compression of Bayer-pattern video sequences using adjusted chroma subsampling," *IEEE Trans. Circuits Syst. Video Technol.*, **19**(12), 1891–1896 (2009).
16. K. H. Chung and Y. H. Chan, "A lossless compression scheme for Bayer color filter array images," *IEEE Trans. Image Process.*, **17**(2), 134–144 (2008).
17. C. Doutre, P. Nasiopoulos, and K. N. Plataniotis, "H.264-based compression of Bayer pattern video sequences," *IEEE Trans. Circuits Syst. Video Technol.*, **18**(6), 725–734 (2008).
18. 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. of 13th European Signal Processing Conf.* (2005).
19. C. C. Koh, J. Mukherjee, and S. K. Mitra, "New efficient methods of image compression in digital cameras with color filter array," *IEEE Trans. Consum. Electron.*, **49**(4), 1448–1456 (2003).
20. R. Lukac and K. N. Plataniotis, "Single-sensor camera image compression," *IEEE Trans. Consum. Electron.* **52**(2), 299–307 (2006).
21. X. Xiang, G. L. Li, and Z. H. Wang, "Low-complexity and high-efficiency image compression algorithm for wireless endoscopy system," *J. Electron. Imaging* **15**(2), 023017 (2006).
22. N. Zhang and X. Wu, "Lossless compression of color mosaic images," *IEEE Trans. Image Process.* **15**(6), 1379–1388 (2006).
23. K. Jack, *Video Demystified: A Handbook for The Digital Engineer*, 5th ed., Eagle Rock, LLH Technology Publishing, VA (2005).
24. Available: ftp://vqeg.its.bldrdoc.gov/HDTV/SVT_MultiFormat/.
25. Available: http://iphome.hhi.de/suehring/tml/download/.
26. C. Laroche and M. Prescott, "Apparatus and method for adaptively interpolating a full color image utilizing chrominance gradients," US Patent No. 5,373,322 (1994).
27. J. Hamilton and J. Adams, "Adaptive color plane interpolation in single sensor color electronic camera," US Patent# No. 5,629,734 (1997).
28. Available: http://140.118.175.164/WJYang/paper/DTDIDownsampling/.

Kuo-Liang Chung received his BS, MS, and PhD degrees in computer science and information engineering from National Taiwan University in 1982, 1984, and 1990, respectively. In 1990, he joined the Department of Computer Science and Information Engineering of National Taiwan University of Science and Technology, Taiwan, as an associate professor. He was promoted to professor and then chair professor in 1995 and 2009, respectively. During 2003 to 2006, he served as the head of the Department of Computer Science and Information Engineering of the National Taiwan University of Science and Technology. He has published numerous articles in well-known international journals. His current research interests include image processing, video coding, and data hiding. During 1996 to 1998, he served as the executive editor of *Journal of the Chinese Institute of Engineers*. In 2000, he served as co-chair of the 13th IPPR Conference on Computer Vision, Graphics, and Image Processing in Taipei. He was a recipient of the Distinguished Engineering Professor Award from Chinese Institute of Engineers in 2001. In 2004, he received the Distinguished Research Award from the National Science Council and in 2007, received the best paper award from the Society of Computer Vision, Graphics, and Image Processing in Taiwan. He is a senior member of IEEE and a fellow of IET.

Wei-Jen Yang received his BS degree in computer science and information engineering from National Taiwan University of Science and Technology in 2004 and his PhD degree in computer science and information engineering from National Taiwan University in 2009. He is now a postdoctoral researcher in the Department of Computer Science and Information Engineering at National Taiwan University of Science and Technology. His current research interests include color image processing, digital camera image processing, data hiding, image/video compression, computer vision, pattern recognition, and algorithms. In 2007, he received the Best Paper Award from the Society of Computer Vision, Graphics, and Image Processing in Taiwan; and in 2010, he received the PhD Dissertation Award from Institute of Information and Computing Machinery in Taiwan.

Chyou-Hwa Chen received his PhD degree from the State University of New York, Stony Brook. He is an associate professor in the Department of Computer Science and Information Engineering, National Taiwan University of Science and Technology, Taiwan. His research interests include H.264/AVC compression, multimedia communications, and computer network.

Hong-Yuan Mark Liao received his BS degree in physics from National Tsing-Hua University, Hsin-Chu, Taiwan, in 1981, and his MS and PhD degrees in electrical engineering from Northwestern University, Evanston, Illinois. in 1985 and 1990, respectively. In 1991, he joined the Institute of Information Science, Academia, Sinica, Taiwan, as an assistant research fellow. He was promoted to associate research fellow and then research fellow in 1995 and 1998, respectively. He is now the division chair of the computer science and information engineering division II, National Science Council of Taiwan. He is also jointly appointed as a professor of the Computer Science and Information Engineering Department of National Chiao-Tung University. Since 2009, he has been jointly appointed as the multimedia information chair professor of National Chung Hsing University. From 2010, he has been appointed as an adjunct chair professor of Chung Yuan Christian University. His current research interests include multimedia signal processing, video-based surveillance systems, content-based multimedia retrieval, and multimedia protection. He was a recipient of the Young Investigators' Award from Academia Sinica in 1998. He received the Distinguished Research Award from the National Science Council of Taiwan in 2003 and the National Invention Award of Taiwan in 2004. In 2008, he received a Distinguished Scholar Research Project Award from National Science Council of Taiwan. In 2010, he received Academia Sinica Investigator Award. In June 2004, he served as the conference co-chair of the 5th International Conference on Multimedia and Exposition (ICME) and technical co-chair of the 8th ICME held at Beijing. In 2011, he will serve as general co-chair of the 17th International Conference on Multimedia Modeling. From 2006 to 2008, he served as the president of the Image Processing and Pattern Recognition Society of Taiwan.

Sheng-Mao Zeng received his BS degree from Computer Science and Information Engineering at Tung-Hai University, Taiwan, in June 2009. He is now pursuing an MS degree in computer science and information engineering at National Taiwan University of Science and Technology, Taiwan. His current research interests include H.264/AVC compression and color image processing.