Adaptive Chroma Subsampling-Binding and Luma-Guided Chroma Reconstruction Method for Screen Content Images

Kuo-Liang Chung, *Senior Member, IEEE*, Chi-Chao Huang, and Tsu-Chun Hsu

***Abstract*— In this paper, we propose a novel adaptive chroma subsampling-binding and luma-guided (ASBLG) chroma reconstruction method for screen content images (SCIs). After receiving the decoded luma and subsampled chroma image from the decoder, a fast winner-first voting strategy is proposed to identify the used chroma subsampling scheme prior to compression. Then, the decoded luma image is subsampled as the identified subsampling scheme was performed on the chroma image such that we are able to conclude an accurate correlation between the subsampled decoded luma image and the decoded subsampled chroma image. Accordingly, an adaptive sliding window-based and luma-guided chroma reconstruction method is proposed. The related computational complexity analysis is also provided. We take two quality metrics, the color peak signal-to-noise ratio (CPSNR) of the reconstructed chroma images and SCIs and the gradient-based structure similarity index (CGSS) of the reconstructed SCIs to evaluate the quality performance. Let the proposed chroma reconstruction method be denoted as 'ASBLG'. Based on 26 typical test SCIs and 6 JCT-VC test screen content video sequences (SCVs), several experiments show that on average, the CPSNR gains of all the reconstructed UV images by 4:2:0(A)-ASBLG, SCIs by 4:2:0(MPEG-B)-ASBLG, and SCVs by 4:2:0(A)-ASBLG are 2.1, 1.87, and 1.87 dB, respectively, when compared with that of the other combinations. Specifically, in terms of CPSNR and CGSS, $CS_{BILINEAR}$-ASBLG for the test SCIs and $CS_{BICUBIC}$-ASBLG for the test SCVs outperform the existing state-of-the-art comparative combinations, where $CS_{BILINEAR}$ and $CS_{BICUBIC}$ denote the luma-aware based chroma subsampling schemes by Wang *et al*.**

***Index Terms*— Chroma reconstruction, chroma subsampling, color gradient-based structure similarity index (CGSS), color peak signal-to-noise ratio (CPSNR), high efficiency video coding (HEVC), screen content images (SCIs), winner-first voting strategy.**

I. Introduction

SCREEN content images (SCIs) are computer generated images for screen display and one SCI usually contains a few different objects, such as symbols, characters, graphics, and/or natural subimage(s). The former three objects are often rendered with few colors, sharp edges, and thin lines while the natural subimage often has continuous-tone colors, blurred edges, and thick lines. Because mobile phones have gained great popularity, SCIs have been widely used in many applications [9], [11], such as video conferencing, cloud computing, screen sharing, remote desktops, and so on. Accordingly, the High Efficiency Video Coding (HEVC) standard includes the Screen Content Coding (SCC) extension [15], [22] in its extended version. In the HEVC-SCC extension, researchers have contributed achievements in various aspects, including intra line copying [1], block vector prediction [23], palette index coding [17], quality assessment metrics [13], [20], the just noticeable difference (JND) model [21], the segmentation algorithm [12], and chroma reconstruction [19].

In the traditional coding system, as shown in Fig. 1, prior to compression, the input RGB image is first transformed into a YUV image, and then the chroma UV image is subsampled by the adopted subsampling scheme. Afterwards, the subsequent processes, such as encoding the subsampled YUV image, transmitting the encoded subsampled YUV image to the decoder, decoding the compressed subsampled YUV image, reconstructing the chroma image, and transforming the YUV image to an RGB image, follow. Before introducing the existing chroma reconstruction methods and highlighting their weaknesses, we introduce the existing chroma 4:2:0 subsampling schemes, although all the discussions in this paper are also applicable to the chroma 4:2:2 subsampling schemes. For convenience, the three terms "chroma 4:2:0 subsampling", "chroma subsampling", and "chroma downsampling", are used interchangeably. The two terms, "chroma reconstruction" and "chroma upsampling", are also used interchangeably.

We first introduce eight existing chroma subsampling schemes. The 4:2:0(A) determines the one sampled U and V components, U_s and V_s, by averaging all the U and V components of the 2×2 UV block. The 4:2:0(L) and 4:2:0(R) determine U_s and V_s by averaging the U and V components in the left and right columns of the block, respectively. The 4:2:0(DIRECT) selects the upper-left U and V components of the block as the U_s and V_s, respectively. The 4:2:0(MPEG-B) performs the 13-tap filter with mask $[2, 0, -4, -3, 5, 19, 26, 19, 5, -3, -4, 0, 2]/64$ on the upper-left U and V components of the block to estimate U_s and V_s, respectively. Zhang *et al.* [26] presented $IDID_{NEDI}$

Manuscript received February 7, 2017; revised June 20, 2017 and August 22, 2017; accepted August 30, 2017. Date of publication September 4, 2017; date of current version October 3, 2017. The work of K. L. Chung was supported by the Ministry of Science and Technology of China under Contract MOST-104-2221-E-011-004-MY3 and Contract MOST-105-2221-E-011-118-MY3. The associate editor coordinating the review of this manuscript and approving it for publication was Pierre-Marc Jodoin. *(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: klchung01@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/TIP.2017.2749148

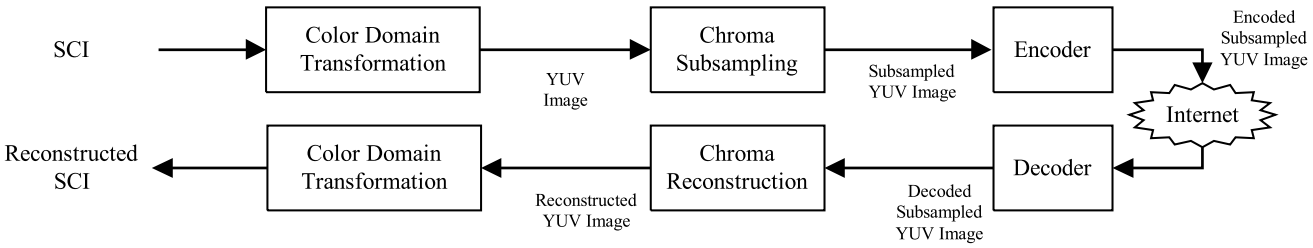


Fig. 1. The traditional coding system.

subsampling scheme which joints the interpolation-dependent image downsampling (IDID) process and the new edge-directed interpolation (NEDI) based upsampling process [10] together such that the error of the reconstructed chroma image could be minimized. Wang *et al.* [19] presented two improved chroma subsampling schemes, called $CS_{BILINEAR}$ and $CS_{BICUBIC}$ which are bundled together with BILINEAR upsampling and BICUBIC upsampling [7], respectively, which should be known in advance such that theoretically, the quality distortion can be minimized. In addition, due to employing the major color and index map (BCIM) representation, i.e. the Palette mode [14], in the luma and chroma images, the reconstructed chroma image has better quality performance. The set of the concerned eight existing subsampling schemes is denoted by CS={4:2:0(A), 4:2:0(L), 4:2:0(R), 4:2:0(DIRECT), 4:2:0(MPEG-B), $IDID_{NEDI}$, $CS_{BILINEAR}$, and $CS_{BICUBIC}$}.

A. Existing Chroma Upsampling Methods and Weaknesses

The existing chroma upsampling methods can be classified into two categories: the luma-independent chroma upsampling category and the luma-dependent chroma upsampling category. As for the luma-independent category, the COPY, 6-TAP [6], BILINEAR, BICUBIC, NEDI, and soft decision adaptive interpolation (SAI) [25] are six popular methods. In the COPY, all the pixels of the 2×2 chroma block receive the subsampled chroma components. In 6-TAP, the filter is realized by running the mask $[1, -5, 20, 20, -5, 1]/32$ centered at each missing chroma pixel location to restore the missing chroma pixel. The BILINEAR and BICUBIC apply the bilinear and bicubic interpolations, respectively, to reconstruct the chroma image from the subsampled chroma image. The NEDI is a modified Wiener filter-based chroma upsampling method and its quality performance is better than that of the edge-directed interpolation (EDI) [3]. The SAI interpolates the missing pixels in each 2×2 chroma block by jointly training the relationship not only between the known pixels and the missing pixels but also between the missing pixels themselves to improve the NEDI. The main weakness existing in the luma-independent category is that each reconstructed chroma pixel is interpolated by only taking its locally neighboring subsampled chroma pixels into account, leading to limited quality improvement of the reconstructed chroma image and SCI.

As for the luma-dependent chroma upsampling category, the luma-assisted chroma upsampling (LAU) method [8], the guided-filter (GF) method [5], the guided chroma reconstruction (GCR) method [18], and the luma aware chroma downsampling and upsampling (LACDU) method [19] are four state-of-the-art chroma upsampling methods. The 4:2:0(A)-LAU combination incorporates the mean squared luma difference between the neighboring chroma pixels of the current missing chroma pixel and the co-located luma pixel into the bilinear interpolation to achieve better quality gain of the reconstructed chroma image when compared to the combination 4:2:0(A)-BILINEAR. The GF applies a linear regression technique to obtain the mapping from each subsampled luma block to the co-located subsampled chroma block, and then the corresponding chroma pixel is reconstructed. The main weakness of the GF is the lack of employing the adopted chroma subsampling information in the luma-guided chroma reconstruction process. In addition, the GF ignores the "out of the min-max range problem," which will be described in Sections III-B, in chroma reconstruction. Both weaknesses degrade the quality performance of the GF. GCR reconstructs the chroma block when the co-located luma block is a textural block and it matches one of two assumptions, "assumption 1: the block must have at most four different luma values and under this condition, if two luma pixels in the block have the same luma value, they also have the same chroma value," and "assumption 2: the chroma component has the same (or inverse) texture as the luma component between the minimum and maximum values." Consequently, a partial chroma image consisting of those textural chroma blocks is reconstructed by the linear regression approach while the remainder consisting of those non-textural chroma blocks is reconstructed by the 4-TAP method. The authors state that the chroma subsampling information is employed in some cases when applying assumption 2. However, the concrete procedure to identify the chroma subsampling information from the subsampled YUV image is not provided in [18]. In addition, the GCR ignores the out of the min-max range problem. Experimental results indicate that the GCR has superior quality performance of the reconstructed chroma image compared with the 4-TAP in HEVC. Due to employing the BCIM representation in the luma and chroma images [19], the experimental results showed that the two combinations, $CS_{BILINEAR}$-LACDU and $CS_{BICUBIC}$-LACDU, have superior quality performance when compared with the comparative combinations, including $IDID_{NEDI}$-NEDI, 4:2:0(A)-(BILINEAR/BICUBIC), MPEG-B-(BILINEAR/BICUBIC), (4:2:0(A)/MPEG-B)-GF, and IDID-GF. Here, (4:2:0(A)/MPEG-B)-GF denotes that the chroma subsampling process is 4:2:0(A) (or the MPEG-B) and the chroma reconstruction process is the GF. The constraint

existing in LACDU is that it must bundle with $CS_{BILINEAR}$ (or $CS_{BICUBIC}$) chroma subsampling scheme.

Since in the three comparative combinations, $IDID_{NEDI}$-NEDI, $CS_{BILINEAR}$-LACDU, and $CS_{BICUBIC}$-LACDU, luma-aware based chroma subsampling process and the chroma reconstruction process are bundled together, we partition the set CS into two subsets, $CS_1 = \{4:2:0(A), 4:2:0(L), 4:2:0(R), 4:2:0(DIRECT), 4:2:0(MPEG\text{-}B)\}$ and $CS_2 = \{IDID_{NEDI}, CS_{BILINEAR}, CS_{BICUBIC}\}$. Similarly, the set CR is partitioned into two subsets, $CR_1 = \{$COPY, 6-TAP, BILINEAR, BICUBIC, NEDI, SAI, GF, GCR, LAU, ASBLG$\}$, in which ASBLG denotes our proposed chroma reconstruction method which will be presented in Sections II and III, and $CR_2 = \{$NEDI, LACDU, ASBLG$\}$. Throughout this paper, the term 'ASBLG' and the term 'the proposed ASBLG chroma reconstruction method' are used interchangeably.

B. Motivation and Contributions

To overcome the weaknesses mentioned in the existing chroma reconstruction methods and to take the special characteristics of the SCIs, mentioned in the first paragraph of the Introduction section, into account, in this paper, we propose an effective adaptive chroma subsampling-binding and luma guided chroma reconstruction method for SCIs. First, we study how to represent the correlation distortion degree (CDD) between the subsampled luma and the co-located subsampled chroma block-pair, and point out that the CDD value under the same subsampling circumstance is often much less than that under different subsampling circumstance. In addition, we show that the lower the value of CDD, the higher the accuracy of the reconstructed chroma component. At the client side, we thus aim at identifying the used chroma subsampling scheme from the decoded UV image without any prior information. To this end, we propose a fast winner-first voting strategy to identify the adopted chroma subsampling scheme used prior to compression. Accordingly, the decoded luma image is subsampled by the identified subsampling scheme. Because the subsampled luma and the co-located subsampled chroma block-pair is under the same subsampling circumstance, it provides an opportunity to derive a more accurate correlation between the two blocks. Further, an adaptive sliding window-based and luma-guided regression approach is proposed to reconstruct the current chroma pixel. However, when the current decoded luma pixel is on the boundary between the foreground and the background, its pixel value usually varies and tends to be out of the min-max range, $[min(Y'_1, Y'_2, \ldots, Y'_n), max(Y'_1, Y'_2, \ldots, Y'_n)]$, where $Y'_1, Y'_2, \ldots,$ and Y'_n denote the subsampled luma pixels in the luma block. The out of the min-max range problem often makes the luma-guided reconstructed chroma value unreliable. As soon as the alarm of the out of the min-max range problem is given, the sliding-window is enlarged to alleviate this problem and increase the chroma reconstruction accuracy. The above-proposed adaptive chroma subsampling-binding and luma-guided (ASBLG) chroma reconstruction method is simply called ASBLG. The related computational complexity analysis of ASBLG is also provided. Based on 26 typical test SCIs collected from the websites in [13] and [19], and 6 SC videos (SCVs) from JCT-VC test video sequences [24], several experiments in HEVC reference software are carried out to show the quality merit of the proposed ASBLG method. Here, the four used quality metrics are the average color peak signal-to-noise ratio (CPSNR) of the reconstructed chroma images, SCIs, and SCVs, the average color gradient-based structure similarity index (CGSS) [13] of the reconstructed SCIs and SCVs, the visual effect of the reconstructed SCIs, and the rate-distortion curve (RD curve) of the reconstructed SCIs and SCVs. For any fixed chroma subsampling scheme c_s in CS_1, c_s-ASBLG has superior quality performance when compared with any comparative combinations $\{c_s\text{-}c_r \mid$ the chroma reconstruction method c_r is in $CR_1\}$. The average CPSNR gain of all the reconstructed UV images by 4:2:0(A)-ASBLG is 2.1 dB when compared with that of the other nine combinations. The average CPSNR gain of the reconstructed SCIs by 4:2:0(MPEG-B)-ASBLG is 1.87 dB when compared with that of the other nine combinations. The average CPSNR gain of the reconstructed SCVs by 4:2:0(A)-ASBLG is 1.87 dB when compared with that of the other nine combinations. Among all the comparative combinations in CS×CR, $CS_{BILINEAR}$-ASBLG for the test SCIs and $CS_{BICUBIC}$-ASBLG for the test SCVs have the best average CPSNR and CGSS performances, in which the chroma subsampling schemes, $CS_{BILINEAR}$ and $CS_{BICUBIC}$, were proposed by Wang *et al.* [19] and ASBLG is our proposed chroma reconstruction method.

The rest of this paper is organized as follows. In Section II, based on the proposed CDD metric for the subsampled luma and chroma block-pair, we study how CDD affects the accuracy of the reconstructed chroma component. Accordingly, in Section III, a fast winner-first voting strategy is proposed to identify the chroma subsampling scheme which had been adopted at the server side. Next, the decoded luma image is subsampled by the identified subsampling scheme. Further, ASBLG is proposed to reconstruct the chroma image at the client side. In Section IV, several experiments are conducted to demonstrate the quality superiority of ASBLG. In Section V, some concluding remarks are addressed.

II. CDD Between the Subsampled Luma and Chroma Block-Pair

Firstly, we propose a correlation point set (CPS) representation to denote the correlation between the subsampled decoded luma block and the co-located decoded subsampled chroma block under the same and different subsampling circumstances. Secondly, based on the CPS representation, a useful metric CDD is proposed to measure the CPS between the subsampled luma and chroma block-pair, and the CDD under the same subsampling scheme could be the lowest when compared with that under different subsampling schemes. Finally, we point out that the lower the CDD is, the higher the chroma reconstruction accuracy is.

A. The CPS Representation Between the Subsampled Luma and Chroma Block-Pair

For easy exposition, a real small example is taken to assist the explanation. Fig. 2(a) depicts a 5×5 chroma

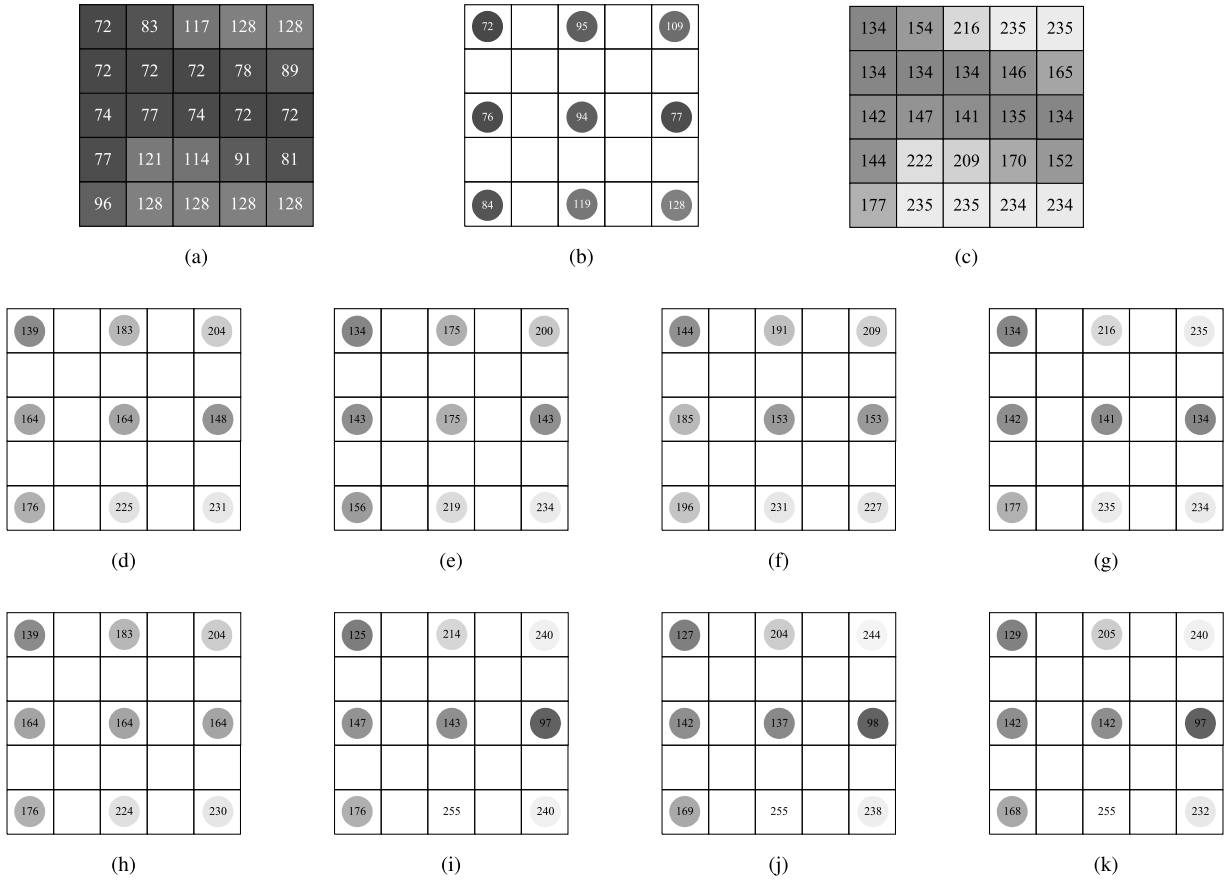


Fig. 2. An example used to illustrate the CDD between the subsampled luma and chroma block-pair. (a) 5×5 chroma block. (b)Subsampled chroma block by 4:2:0(L). (c)Co-located 5×5 luma block. (d) Subsampled luma block by 4:2:0(A). (e) Subsampled luma block by 4:2:0(L). (f) Subsampled luma block by 4:2:0(R). (g) Subsampled luma block by 4:2:0(DIRECT). (h) Subsampled luma block by 4:2:0(MPEG-B). (i) Subsampled luma block by 4:2:0($\text{IDID}_{\text{NEDI}}$). (j) Subsampled luma block by 4:2:0($\text{CS}_{\text{BILINEAR}}$). (k) Subsampled luma block by 4:2:0($\text{CS}_{\text{BICUBIC}}$).

block cut off from the 12th SCI in the test SCI set. Running 4:2:0(L) on Fig. 2(a), Fig. 2(b) illustrates the subsampled chroma block after translating a half pixel to the north. Fig. 2(c) shows the 5×5 luma block. Let the 5×5 subsampled luma and chroma blocks be denoted by $Y'^{(l_s)}(=\{Y'_1, Y'_2, Y'_3, Y'_4, Y'_5, Y'_6, Y'_7, Y'_8, Y'_9\})$ and $U'^{(c_s)}(=\{U'_1, U'_2, U'_3, U'_4, U'_5, U'_6, U'_7, U'_8, U'_9\})$, respectively, where l_s and c_s denote the adopted luma and chroma subsampling schemes in CS.

The CPS between $Y'^{(l_s)}$ and $U'^{(c_s)}$ is expressed as $\{(U'_1, Y'_1), (U'_2, Y'_2), \ldots, (U'_9, Y'_9)\}$. The eight correlations between the 5×5 subsampled chroma block $U'^{(\text{L})}(= U'^{4:2:0(\text{L})})$, as shown in Fig. 2(b), and the eight co-located 5×5 subsampled luma blocks, as shown in Figs. 2(d)-(k), are depicted by the eight CPSs, as shown in Figs. 3(a)-(h), respectively, in which the x-axis denotes the subsampled chroma value and the y-axis denotes the subsampled luma value. From the distributions of the eight CPSs in Fig. 3, we observe that the CDD between $U'^{(\text{L})}$ and $Y'^{(\text{L})}$ is the lowest among the eight CPSs.

B. A Metric CDD to Measure the CPS for the Subsampled Luma and Chroma Block-Pair

In this subsection, a robust metric CDD is proposed to measure the CPS for the subsampled luma and chroma block-pair. The correlation for the block-pair $(U'^{(\text{L})}, Y'^{(l_s)})$ can be determined by solving the following overdetermined system:

$$\begin{bmatrix} Y'_1 & 1 \\ Y'_2 & 1 \\ Y'_3 & 1 \\ \vdots & \vdots \\ Y'_9 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} U'_1 \\ U'_2 \\ U'_3 \\ \vdots \\ U'_9 \end{bmatrix} \tag{1}$$

The correlation parameter-pair (a, b) can be solved by linear regression and it yields

$$\begin{bmatrix} a \\ b \end{bmatrix} = \left(\begin{bmatrix} Y'_1 & 1 \\ Y'_2 & 1 \\ Y'_3 & 1 \\ \vdots & \vdots \\ Y'_9 & 1 \end{bmatrix}^T \begin{bmatrix} Y'_1 & 1 \\ Y'_2 & 1 \\ Y'_3 & 1 \\ \vdots & \vdots \\ Y'_9 & 1 \end{bmatrix} \right)^{-1} \begin{bmatrix} Y'_1 & 1 \\ Y'_2 & 1 \\ Y'_3 & 1 \\ \vdots & \vdots \\ Y'_9 & 1 \end{bmatrix}^T \begin{bmatrix} U'_1 \\ U'_2 \\ U'_3 \\ \vdots \\ U'_9 \end{bmatrix} \tag{2}$$

Geometrically, the two correlation parameters, a and b, could be the slope and intercept of the fitting line:

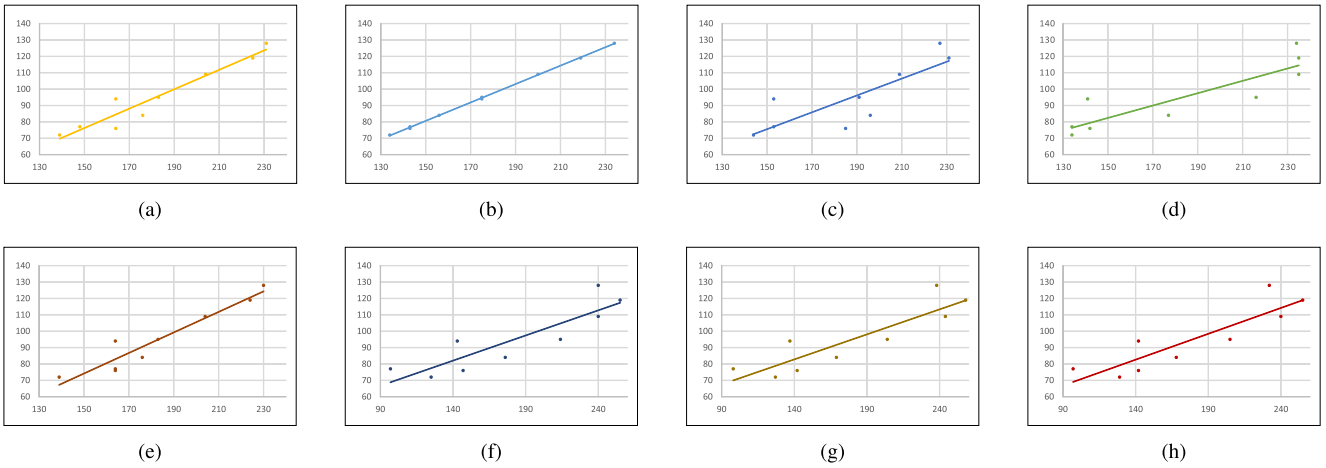


Fig. 3. Eight CPSs for $\{(U'^{(\mathrm{L})}, Y'^{(ls)})|ls \in \mathrm{CS}\}$. (a) CPS for $(U'^{(\mathrm{L})}, Y'^{(\mathrm{A})})$. (b) CPS for $(U'^{(\mathrm{L})}, Y'^{(\mathrm{L})})$. (c) CPS for $(U'^{(\mathrm{L})}, Y'^{(\mathrm{R})})$. (d) CPS for $(U'^{(\mathrm{L})}, Y'^{(\mathrm{DIRECT})})$. (e) CPS for $(U'^{(\mathrm{L})}, Y'^{(\mathrm{MPEG\text{-}B})})$. (f) CPS for $(U'^{(\mathrm{L})}, Y'^{(\mathrm{IDID_{NEDI}})})$. (g) CPS for $(U'^{(\mathrm{L})}, Y'^{(\mathrm{CS_{BILINEAR}})})$. (h) CPS for $(U'^{(\mathrm{L})}, Y'^{(\mathrm{CS_{BICUBIC}})})$.

$U = aY + b$, respectively, where U and Y indicate the variables on the chroma axis and the luma axis. Therefore, the CDD of $(U'^{(\mathrm{L})}, Y'^{(l_s)})$ can be expressed as

$$CDD(U'^{(c_s)}, Y'^{(l_s)}) = \sum_{i=1}^{9} \frac{|a \times Y'_i + b - U'_i|^2}{a^2 + 1} \tag{3}$$

For example, by (2), the correlation parameter-pair for $(U'^{(\mathrm{L})}, Y'^{(\mathrm{DIRECT})})$ corresponding to (Fig. 2(b), Fig. 2(g)) is given by

$$\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 0.38 \\ 25.83 \end{bmatrix} \tag{4}$$

Accordingly, the fitting line is equal to $U' = 0.38 \times Y' + 25.83$, which is denoted by the green line in Fig. 3(d). By (3), the CDD of $(U'^{(\mathrm{L})}, Y'^{(\mathrm{DIRECT})})$ is equal to

$$CDD(U'^{(\mathrm{L})}, Y'^{(\mathrm{DIRECT})}) = 2183.3 \tag{5}$$

Since the CDD of $(U'^{(\mathrm{L})}, Y'^{(\mathrm{DIRECT})})$ is large, it implies that the correlation between the subsampled chroma block by 4:2:0(L) and the subsampled luma block by 4:2:0(DIRECT) is rather dispersed. On the other hand, the probability of 4:2:0(DIRECT) being the candidate chroma subsampling scheme is relatively low, to say the lowest.

By the same argument, the correlation parameter-pair and the CDD of $(U'^{(\mathrm{L})}, Y'^{(\mathrm{L})})$ corresponding to (Fig. 2(b), Fig. 2(e)) are equal to

$$\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 0.56 \\ -4.4 \end{bmatrix} \quad \text{and } CDD(U'^{(\mathrm{L})}, Y'^{(\mathrm{L})}) = 1.16 \tag{6}$$

Clearly the value of $CDD(U'^{(\mathrm{L})}, Y'^{(\mathrm{L})})$ in (6) is rather small and much less than the value of $CDD(U'^{(\mathrm{L})}, Y'^{(\mathrm{DIRECT})})$. As for the candidate chroma subsampling scheme, the probability of 4:2:0(L) is much higher than that of the 4:2:0(DIRECT). Consequently, we take the CDD defined in (3) as the metric to assist the identification of the used chroma subsampling scheme which had been used at the server side. When one candidate chroma subsampling scheme finally turns out to be the true chroma subsampling scheme, it can be used to subsample the decoded luma image in order to obtain an accurate correlation platform between the subsampled luma and chroma image-pair, and further to achieve better chroma reconstruction results.

C. The Relation Between the CDD and the Chroma Reconstruction Accuracy

In this subsection, we investigate how the CDD affects the chroma reconstruction accuracy. It is known that the CDD of $(U'^{(\mathrm{L})}, Y'^{(\mathrm{DIRECT})})$ is 2183.3. By (4) and the central luma pixel value of Fig. 2(c), 141, the central chroma pixel value of Fig. 2(b) can be reconstructed by

$$U'_5 = 0.38 \times 141 + 25.83 = 79 \tag{7}$$

Subtracting the central chroma pixel value of Fig. 2(a), 74, from the reconstructed central chroma value, 79, the chroma reconstruction distortion is 5 $(= 79 - 74)$. We now consider the block-pair, $(U'^{(\mathrm{L})}, Y'^{(\mathrm{L})})$, under the same subsampling circumstance. By (6), it yields

$$U'_5 = 0.56 \times 141 + (-4.4) = 75 \tag{8}$$

Subtracting the central chroma pixel of Fig. 2(a) from the reconstructed central chroma value, 75, the chroma reconstruction distortion is only 1 $(= 75 - 74)$. We conclude that under the same subsampling scheme for luma and chroma block-pair, it leads to higher chroma reconstruction accuracy. It motivates us to develop an adaptive chroma subsampling-binding luma-guided chroma reconstruction method, ASBLG, in order to increase the quality of the reconstructed chroma image and SCI.

III. The Proposed Chroma Reconstruction Method: ASBLG

Due to the correlation between the decoded subsampled luma and chroma block-pair under the same subsampling

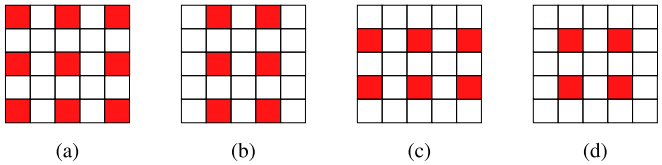


Fig. 4. Four block types. (a) Type A. (b) Type B. (c) Type C. (d) Type D.

scheme and the sharp edge coincidence between the block-pair in SCI, in Subsection III-A, we propose a winner-first voting strategy at the client side to identify the used chroma subsampling scheme at the server side. In addition, the related computational complexity analysis is provided. In Subsection III-B, ASBLG is proposed and the computational complexity analysis is also provided.

A. Winner-First Voting Strategy to Identify the Used Chroma Subsampling Scheme

We begin with the definition of the accumulated CDD (ACDD), and then present the proposed luma-guided winner-first voting strategy to automatically identify the used chroma subsampling scheme.

1) Luma-Guided Winner-First Voting Strategy: After receiving the decoded luma image Y and the subsampled chroma image $U'^{(c_s)}$ from the decoder, our goal is to efficiently identify the used subsampling scheme c_s. We first create eight different 5×5 subsampled luma blocks by running the eight subsampling schemes in CS on the current 5×5 luma block. Since the proposed identification strategy is a block-based approach, the configuration of the subsampled pixels in one 5×5 subsampled luma and chroma block-pair may be the type A block with nine subsampled pixel-pairs in Fig. 4(a), the type B block with six subsampled pixel-pairs in Fig. 4(b), the type C block with six subsampled pixel-pairs in Fig. 4(c), or the type D block with four subsampled pixel-pairs in Fig. 4(d).

In our example, the initial used chroma subsampling scheme is assumed to be 4:2:0(L). According to the row-major scanning order, the first 5×5 subsampled type A chroma block and the eight generated 5×5 subsampled type A luma blocks, as shown in Fig. 2(b) and Figs. 2(d)-(k), respectively, are denoted by $U'^{(\mathrm{L})}$, $Y'^{(\mathrm{A})}$, $Y'^{(\mathrm{L})}$, $Y'^{(\mathrm{R})}$, $Y'^{(\mathrm{DIRECT})}$, $Y'^{(\mathrm{MPEG\text{-}B})}$, $Y'^{(\mathrm{IDID_{NEDI}})}$, $Y'^{(\mathrm{CS_{BILINEAR}})}$, and $Y'^{(\mathrm{CS_{BICUBIC}})}$, respectively. By (3), after the first iteration, we have the eight CDDs: $CDD^{(1)}(U'^{(\mathrm{L})}, Y'^{(\mathrm{A})}) = 183.72$, $CDD^{(1)}(U'^{(\mathrm{L})}, Y'^{(\mathrm{L})}) = 1.16$, $CDD^{(1)}(U'^{(\mathrm{L})}, Y'^{(\mathrm{R})}) = 802.67$, $CDD^{(1)}(U'^{(\mathrm{L})}, Y'^{(\mathrm{DIRECT})}) = 2183.3$, $CDD^{(1)}(U'^{(\mathrm{L})}, Y'^{(\mathrm{MPEG\text{-}B})}) = 205.93$, $CDD^{(1)}(U'^{(\mathrm{L})}, Y'^{(\mathrm{IDID_{NEDI}})}) = 651.13$, $CDD^{(1)}(U'^{(\mathrm{L})}, Y'^{(\mathrm{CS_{BILINEAR}})}) = 616.97$, and $CDD^{(1)}(U'^{(\mathrm{L})}, Y'^{(\mathrm{CS_{BICUBIC}})}) = 634.91$. Among the above eight CDDs, the value of $CDD^{(1)}(U'^{(\mathrm{L})}, Y'^{(\mathrm{L})})$, 1.16, is the smallest one. Thus, we select 4:2:0(L) as the candidate subsampling scheme and 4:2:0(L) gets the second vote. In the second iteration, we only calculate the value of $Error^{(2)}(U'^{(\mathrm{L})}, Y'^{(\mathrm{L})})$. Then, after performing the first two iterations, the value of ACDD is defined to be $ACDD^{(1\le k\le 2)}(U'^{(\mathrm{L})}, Y'^{(\mathrm{L})}) = \sum_{k=1}^{2} CDD^{(k)}(U'^{(\mathrm{L})}, Y'^{(\mathrm{L})})$ which will be used to examine

Fig. 5. The example used in the second iteration of the proposed winner-first voting strategy. (a) Subsampled type B chroma block. (b) Co-located luma block. (c) Subsampled type B luma block by 4:2:0(L).

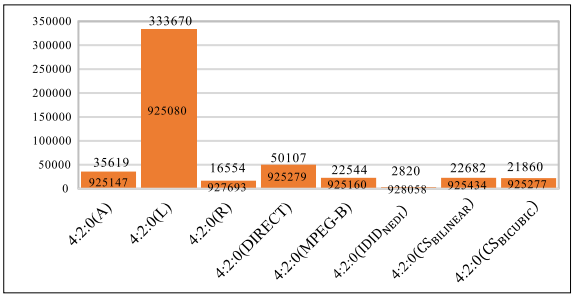


Fig. 6. Final votes and ACDDs of the eight subsampling schemes for the 12th test SCI by the proposed winner-first voting strategy.

whether 4:2:0(L) continues to reserve the candidate subsampling scheme role or not. This is why it is called the winner-first voting strategy. Returning to our example, in the second iteration, we shift one pixel to the right to pick up the next 5×5 subsampled type B chroma block and the co-located 5×5 luma block, as shown in Fig. 5(a) and Fig. 5(b), respectively. After running 4:2:0(L) on the current luma block in Fig. 5(b), the subsampled type B luma block is shown in Fig. 5(c). By (3), it yields $CDD^{(2)}(U'^{(\mathrm{L})}, Y'^{(\mathrm{L})}) = 0.69$ and $ACDD^{(1\le k\le 2)}(U'^{(\mathrm{L})}, Y'^{(\mathrm{L})}) = 1.16 + 0.69 = 1.85$.

After performing the first two iterations, the value of $ACDD^{(1\le k\le 2)}(U'^{(\mathrm{L})}, Y'^{(\mathrm{L})})$, 1.85, is the smallest one among the eight temporary ACDDs. In the third iteration, we continue to select 4:2:0(L) as the candidate subsampling scheme and 4:2:0(L) gets one more vote. We repeat the above process until all the necessary decoded subsampled chroma and luma block-pairs in the decoded subsampled YUV image have been completed. Based on the 12th test SCI, the final votes and ACDDs of the concerned eight subsampling schemes are listed below:

$$\begin{aligned}
ACDD^{(1\le k\le 35619)}(U'^{(\mathrm{L})}, Y'^{(\mathrm{A})}) &= 925147\\
\boldsymbol{ACDD^{(1\le k\le 333670)}(U'^{(\mathrm{L})}, Y'^{(\mathrm{L})})} &\boldsymbol{= 925080}\\
ACDD^{(1\le k\le 16554)}(U'^{(\mathrm{L})}, Y'^{(\mathrm{R})}) &= 927693\\
ACDD^{(1\le k\le 50107)}(U'^{(\mathrm{L})}, Y'^{(\mathrm{DIRECT})}) &= 925279\\
ACDD^{(1\le k\le 22544)}(U'^{(\mathrm{L})}, Y'^{(\mathrm{MPEG\text{-}B})}) &= 925160\\
ACDD^{(1\le k\le 2820)}(U'^{(\mathrm{L})}, Y'^{(\mathrm{IDID_{NEDI}})}) &= 928058\\
ACDD^{(1\le k\le 22682)}(U'^{(\mathrm{L})}, Y'^{(\mathrm{CS_{BILINEAR}})}) &= 925434\\
ACDD^{(1\le k\le 21860)}(U'^{(\mathrm{L})}, Y'^{(\mathrm{CS_{BICUBIC}})}) &= 925277 \qquad (9)
\end{aligned}$$

where all the final ACDDs are truncated to integers. From (9), the final ACDD and votes of 4:2:0(L) in boldface are

925080 and 333670, respectively, which are the smallest and largest in the concerned eight final ACDDs and votes. Accordingly, 4:2:0(L) is the identified true subsampling scheme. Fig. 6 depicts the final ACDD and votes of the eight subsampling schemes. In Fig. 6, except 4:2:0(L), the sum of the final votes of the other seven subsampling schemes is 172186 (= 35619 + 16554 + 50107 + 22544 + 2820 + 22682 + 21860). In total, the sum of final votes of the eight subsampling schemes is 505856 (= 333670 + 172186) which is the size of the 12th decoded luma SCI. It implies that the computational complexity of the proposed subsampling scheme identification is proportional to the image size.

2) Computational Complexity Analysis: In each iteration, the proposed winner-first voting strategy consumes most of the computational efforts in running the candidate subsampling scheme on the current 5×5 luma block, solving the correlation parameter-pair, calculating the ACDD of the corresponding blockpair, and selecting the next candidate subsampling scheme. When there are more than two candidate subsampling schemes with the same ACDD, we randomly select any of them as the next candidate subsampling scheme. Let the number of total pixels of the image be P. Accordingly, the number of total iterations in the proposed voting strategy is bounded by O(P), where the complexity definition in terms of big-O is suggested to refer to [2]. In fact, all the historically solved O(P) correlation parameter-pairs of the final identified subsampling scheme, 4:2:0(L) in our example, can be reserved in one hash table H[] with memory size O(P). Similarly, all of the historically subsampled decoded luma blocks by the 4:2:0(L) can be reserved in one array L[,] with memory size O(P). The built up hash table H[] and array L[,] can be reused in the subsequent adaptive luma-guided chroma reconstruction process which will be described in the next subsection, leading to computation-saving effect.

B. Proposed Adaptive Luma-Guided Chroma Reconstruction Method

In this subsection, we first highlight the out of the min-max range problem, which occurs in the luma-guided chroma pixel reconstruction, and then propose an adaptive sliding-window based approach to alleviate this problem. Furthermore, the related computational complexity is provided. According to the identified subsampling scheme, the 4:2:0(L) in our example, for each 2×2 luma block in the decoded luma image, we determine the subsampled luma value Y_s by averaging the two Y components in the left column of the block.

1) Adaptive Luma-Guided Chroma Reconstruction: As mentioned in the first paragraph of Section III, for the SCIs, the sharp edge between the decoded subsampled luma and chroma block-pair under the same subsampling scheme is often coincident. However, the original luma pixel value on the boundary between the foreground, e.g. text, and the background often varies. If the central pixel value of the 5×5 subsampled luma block, Y_{center}, is out of the min-max range $[min(Y'_1, Y'_2, \ldots, Y'_n), max(Y'_1, Y'_2, \ldots, Y'_n)]$, the solved correlation parameter-pair of the 5×5 subsampled luma and chroma block-pair may be unreliable, and this out of the min-max range problem may cause a biased chroma pixel reconstruction. Practically, a small tolerance value ε is added to make the min-max range bigger. That is, the min-max range becomes $[min(Y'_1, Y'_2, \ldots, Y'_n) - \varepsilon, max(Y'_1, Y'_2, \ldots, Y'_n) + \varepsilon]$. We've tried the value of ε from 0 to 25, and the experimental results indicate that when the tolerance value ε is 5, the average quality of the reconstructed images is the best to all the test images. Figs. 7(a)-(b) illustrate a luma and chroma block-pair example. Figs. 7(c)-(d) illustrate the corresponding central 5×5 subsampled luma and chroma block-pair by 4:2:0(L). We observe that the central luma pixel value of Fig. 7(a) is on the boundary between the text and the background, and the value of Y_{center}, 93, is out of the min-max range [181,204] (= $[min(Y'_1, Y'_2, \ldots, Y'_6) - 5, max(Y'_1, Y'_2, \ldots, Y'_6) + 5] = [min(186, 193, 186, 193, 196, 199) - 5, max(186, 193, 186, 193, 196, 199) + 5]$). By (2), the solved correlation parameter-pair for $(U'^{(L)}, Y'^{(L)})$ corresponding to (Fig. 7(c), Fig. 7(d)) is equal to $(a, b) = (2.28, -318.9)$. The reconstructed chroma value of U'_{center} is estimated by

$$U'_{center} = 93 \times 2.28 - 318.9 = -107 \tag{10}$$

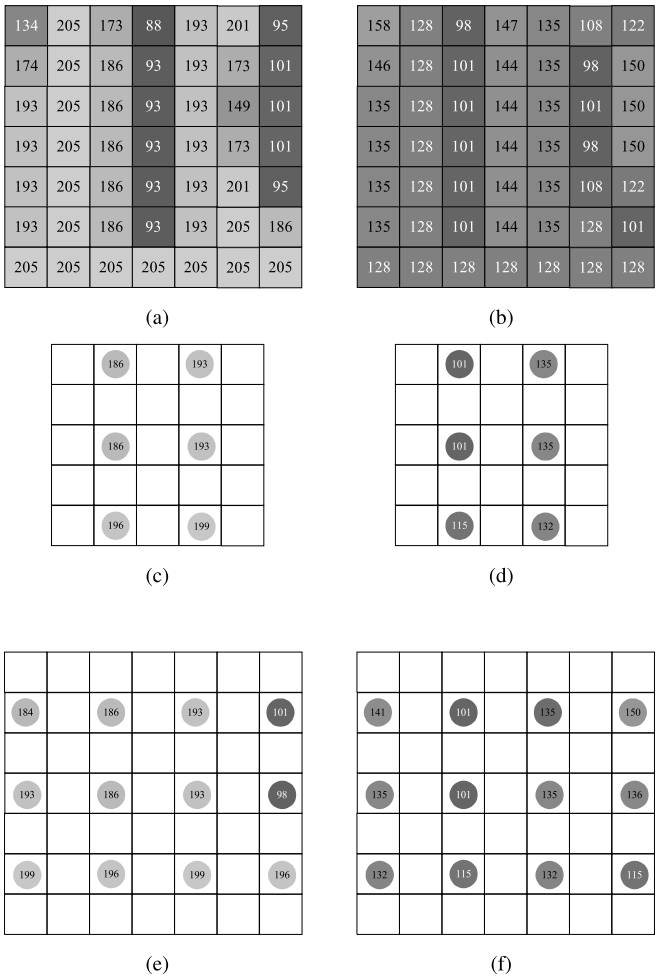


Fig. 7. An example used to illustrate the proposed adaptive sliding-window based chroma reconstruction process. (a) 7×7 luma block. (b) Co-located 7×7 chroma block. (c) Central 5×5 subsampled luma block by 4:2:0(L). (d) Central 5×5 subsampled chroma block by 4:2:0(L). (e) Central 7×7 subsampled luma block by 4:2:0(L). (f) Central 7×7 subsampled chroma block by 4:2:0(L).

Fig. 8. The 26 test SCIs.

It is only to be expected that the inaccurate estimated value of U'_{center} is negative. Heuristically, the reconstructed value of U'_{center} is clipping to 0. On the contrary, when the reconstructed chroma value of U'_{center} is larger than 255, the value of U'_{center} is clipping to 255. Even so, the out of the min-max range problem does degrade the quality performance of the reconstructed chroma image.

When the alarm of the out of the min-max range problem is given, in order to alleviate the quality degradation, we enlarge the sliding window size from 5×5 to 7×7. Figs. 7(e)-(f) illustrate the enlarged 7×7 block-pair. Applying the above chroma reconstruction process to the enlarged block-pair, we have

$$U'_{center} = 93 \times -0.18 + 159.97 = 143 \tag{11}$$

As expected, the reconstructed chroma pixel value, $U'_{center} = 143$, and the value of Y_{center}, 93, is feasible within the min-max range [93,204]. In our implementation, the maximal sliding window size allowable is set to be 11×11.

2) Computational Complexity Analysis: To reconstruct the ith chroma pixel, the solved correlation parameter-pair for the ith subsampled luma and chroma block-pair can be accessed from the hash table H in O(1) time. Next, we check whether the out of the min-max range problem happens or not. If no, it takes O(1) time to compute $U'_{center} = a * Y_{center} + b$ as the reconstructed chroma pixel value; otherwise, we enlarge the sliding-window size from 5×5 to 7×7 and repeat the above chroma pixel reconstruction process. It can be verified that the proposed sliding-window based and luma-guided chroma reconstruction method for reconstructing the whole chroma image can be done in O(P) time.

Combining the computational time required in the proposed winner-first voting strategy, as analyzed in Subsection III-A.2, and the time required in the proposed adaptive luma-guided chroma reconstruction method, it takes O(P) (= O(P) + O(P)) time to complete the whole chroma reconstruction work. Based on the 26 test SCIs, the execution-time improvement ratio is 34.9% (= $((T_{no-reuse} - T_{reuse})/T_{no-reuse})$ in which $T_{no-reuse}$ denotes the execution-time required in the proposed ASBLG method without the hash table reuse and T_{reuse} denotes the execution-time required in the proposed method with the hash table reuse. Note that in [19], the computational time required in Wang et al.'s chroma reconstruction method is also O(P).

IV. Experimental Results

The 26 test SCIs with sizes ranging from 624×624 to 1920×1080, as shown in Fig. 8, and the 6 JCT-VC test SCVs, sc_SlideShow, sc_flyingGraphics, sc_desktop, sc_console, MissionControlClip3, and ChineseEditing, each sequence with 30 image frames with sizes ranging from 1280×720 to 1920×1080, are used to evaluate the quality performance among the concerned chroma reconstruction methods. The quality metrics adopted to compare the quality performance are CPSNR, CGSS, visual effect, and quality-bitrate trade-off in terms of RD curves. All experiments are implemented in the HEVC reference software HM-16.4 on a computer with an Intel Core i7-4790 CPU 3.6 GHz and 8 GB RAM. The operating system is Microsoft Windows 7 64-bit. The program development environment is Visual C++ 2013.

A. CPSNR and CGSS Quality Improvement

In this subsection, the CPSNR and CGSS quality performance comparison among the 50 comparative combinations in $CS_1 \times CR_1$ and the 6 comparative combinations in $CS_2 \times CR_2$, respectively, are reported.

1) CPSNR and CGSS Quality Performance Comparison Among the 50 Comparative Combinations in $CS_1 \times CR_1$: Following the same evaluation way in [19], in order to avoid the compression effect from HM-16.4, the chroma components are just first downsampled and then upsampled. To evaluate the

TABLE I

CPSNR QUALITY COMPARISON AMONG $CS_1 \times CR_1$ FOR THE RECONSTRUCTED 26 UV IMAGES, 26 SCIS, AND 6 SCVS

	COPY	6-TAP	BILINEAR	BICUBIC	NEDI	SAI	GF	GCR	LAU	ASBLG
4:2:0(A)	35.3293	34.4316	34.3607	34.4620	34.2815	33.9051	34.8160	35.4437	35.6193	**36.8404**
	(31.8732)	(30.9862)	(30.8682)	(30.9901)	(30.7715)	(30.5800)	(31.1756)	(31.8339)	(32.0082)	**(33.1201)**
	[29.9758]	[29.0561]	[28.9306]	[29.0511]	[28.9202]	[28.5815]	[29.5328]	[29.9143]	[30.4063]	**[31.2487]**
4:2:0(L)	33.5601	34.4256	34.6972	34.6695	34.6746	34.1068	34.4698	35.0317	34.9399	**36.5816**
	(30.3235)	(31.0724)	(31.2429)	(31.2563)	(31.1887)	(30.8427)	(30.8646)	(31.6736)	(31.3950)	**(32.8743)**
	[28.2629]	[28.8207]	[29.0612]	[29.0272]	[29.1112]	[28.6815]	[29.0590]	[29.1490]	[29.6427]	**[30.8748]**
4:2:0(R)	33.5593	32.2174	32.4253	32.3739	32.3848	32.1240	34.5871	34.9300	35.0030	**36.6582**
	(30.2381)	(28.9376)	(29.0849)	(29.0683)	(29.0228)	(28.8751)	(30.9709)	(30.9149)	(31.4445)	**(32.9707)**
	[28.2409]	[26.9636]	[27.0635]	[27.0628]	[27.0491]	[26.9021]	[29.1241]	[27.6666]	[29.7088]	**[30.7172]**
4:2:0(DIRECT)	32.4071	34.5504	35.0794	34.9762	35.0897	34.5591	34.1919	34.7120	34.5702	**36.5206**
	(29.2823)	(31.2415)	(31.6288)	(31.5865)	(31.5966)	(31.2002)	(30.6120)	(31.4335)	(31.0500)	**(32.8328)**
	[27.2739]	[28.8715]	[29.3376]	[29.2282]	[29.4345]	[28.7486]	[28.6826]	[28.4741]	[29.2290]	**[30.2414]**
4:2:0(MPEG-B)	33.3333	35.2063	35.0475	35.2146	34.9277	34.5841	34.1852	35.3307	34.7104	**36.8211**
	(29.9707)	(31.7706)	(31.5463)	(31.7452)	(31.4065)	(31.2335)	(30.5920)	(31.9062)	(31.1423)	**(33.1223)**
	[28.0644]	[29.7837]	[29.5872]	[29.7610]	[29.5530]	[29.3319]	[28.7779]	[29.2387]	[29.4317]	**[30.5801]**
Average CPSNR	33.6378	34.1663	34.3220	34.3392	34.2707	33.8558	34.4500	35.0896	34.9686	**36.6844**
	(30.3376)	(30.8017)	(30.8778)	(30.9293)	(30.7972)	(30.5463)	(30.8430)	(31.5524)	(31.4080)	**(32.9840)**
	[28.3636]	[28.6991]	[28.7960]	[28.8261]	[28.8136]	[28.4491]	[29.0353]	[28.8886]	[29.6837]	**[30.7324]**
Average CPSNR Gain	3.0466	2.5181	2.3624	2.3452	2.4137	2.8286	2.2344	1.5869	1.7158	
	(2.6464)	(2.1823)	(2.1062)	(2.0547)	(2.1868)	(2.4377)	(2.1410)	(1.4316)	(1.5760)	
	[2.3689]	[2.0333]	[1.9364]	[1.9064]	[1.9189]	[2.2833]	[1.6972]	[1.8439]	[1.0487]	

TABLE II

CGSS QUALITY COMPARISON AMONG $CS_1 \times CR_1$ FOR THE RECONSTRUCTED 26 SCIS AND 6 SCVS

	COPY	6-TAP	BILINEAR	BICUBIC	NEDI	SAI	GF	GCR	LAU	ASBLG
4:2:0(A)	(0.1951)	(0.1983)	(0.1894)	(0.1925)	(0.1936)	(0.1981)	(0.1769)	(0.1849)	(0.1733)	**(0.1714)**
	[0.2252]	[0.2362]	[0.2285]	[0.2310]	[0.2361]	[0.2411]	[0.2115]	[0.2143]	[0.2027]	**[0.1959]**
4:2:0(L)	(0.2007)	(0.2024)	(0.1898)	(0.1941)	(0.1939)	(0.1998)	(0.1779)	(0.1948)	(0.1764)	**(0.1725)**
	[0.2253]	[0.2370]	[0.2249]	[0.2289]	[0.2323]	[0.2409]	[0.2135]	[0.2253]	[0.2060]	**[0.1977]**
4:2:0(R)	(0.2006)	(0.2093)	(0.1990)	(0.2023)	(0.2031)	(0.2085)	(0.1781)	(0.2030)	(0.1767)	**(0.1723)**
	[0.2253]	[0.2433]	[0.2340]	[0.2369]	[0.2392]	[0.2466]	[0.2125]	[0.2339]	[0.2057]	**[0.2000]**
4:2:0(DIRECT)	(0.2064)	(0.2055)	(0.1886)	(0.1943)	(0.1929)	(0.2003)	(0.1802)	(0.1964)	(0.1801)	**0.1717**
	[0.2281]	[0.2379]	[0.2220]	[0.2272]	[0.2275]	[0.2414]	[0.2147]	[0.2248]	[0.2083]	**[0.1988]**
4:2:0(MPEG-B)	(0.2091)	(0.2003)	(0.1936)	(0.1951)	(0.1966)	(0.1995)	(0.1807)	(0.2000)	(0.1889)	**(0.1744)**
	[0.2378]	[0.2361]	[0.2298]	[0.2311]	[0.2342]	[0.2405]	[0.2158]	[0.2314]	[0.2186]	**[0.2058]**
Average CGSS	(0.2024)	(0.2032)	(0.1921)	(0.1957)	(0.1960)	(0.2012)	(0.1788)	(0.1958)	(0.1791)	**(0.1725)**
	[0.2283]	[0.2381]	[0.2278]	[0.2310]	[0.2339]	[0.2421]	[0.2136]	[0.2260]	[0.2082]	**[0.1996]**
Average CGSS Gain	(0.0299)	(0.0307)	(0.0196)	(0.0232)	(0.0236)	(0.0288)	(0.0063)	(0.0233)	(0.0066)	
	[0.0287]	[0.0385]	[0.0282]	[0.0314]	[0.0342]	[0.0425]	[0.0140]	[0.0263]	[0.0086]	

average CPSNR quality of the reconstructed chroma images, the average CPSNR metric is defined by

$$\text{CPSNR} = \frac{1}{N}\sum_{n=1}^{N} 10\log_{10}\frac{255^2}{\text{CMSE}} \tag{12}$$

with $\text{CMSE} = \frac{1}{2WH}\sum_{p\in P}\sum_{C\in\{U,V\}}[I_{n,C}^{ori,UV}(p) - I_{n,C}^{rec,UV}(p)]^2$ in which $\text{P} = \{(l,m)|1 \le l \le H, 1 \le m \le W\}$ denotes the set of pixel coordinates in one W×H image, N (= 26) denotes the number of the test SCIs, $I_{n,C}^{ori,UV}(p)$ denotes the C-color value of the pixel at position p in the nth original UV image and $I_{n,C}^{rec,UV}(p)$ the reconstructed analogue. The CPSNR for the reconstructed SCI can be defined by replacing the two terms, 2 and (U, V), in (12) with the two terms, 3 and (R, G, B), respectively.

Table I indicates the average CPSNR values of the reconstructed 26 UV images, 26 SCIs, and 6 SCVs for each of the 50 comparative combinations in $CS_1 \times CR_1$. For simplicity, the average CPSNR value of the reconstructed SCIs by each combination is listed in the parenthesis '()'. The average CPSNR value of the reconstructed SCVs with respect to the JCT-VC test SCVs for each comparative combination is listed in the parenthesis '[]'. In Table I, for one specific CS scheme, the best CR method is marked in boldface and we clearly observe that ASBLG always has the best CPSNR performance. The average CPSNR gain of all the reconstructed UV images by 4:2:0(A)-ASBLG is 2.1 dB when compared with that of the other nine combinations. The average CPSNR gain of the reconstructed SCIs by 4:2:0(MPEG-B)-ASBLG is 1.87 dB when compared with that of the other nine combinations. The average CPSNR gain of the reconstructed SCVs by 4:2:0(A)-ASBLG is 1.87 dB when compared with that of the other nine combinations. Among the 50 comparative combinations in $CS_1 \times CR_1$, 4:2:0(A)-ASBLG, 4:2:0(MPEG-B)-ASBLG, and 4:2:0(A)-ASBLG have the best average CPSNR performance for all the reconstructed UV images, SCIs, and SCVs, respectively.

The CGSS [13] is a quality assessment metric for evaluating the quality performance of the reconstructed 26 SCIs and 6 SCVs. The CGSS is defined by

$$\text{CGSS} = \frac{1}{N}\sum_{n=1}^{N}\frac{1}{3}(\text{GSS}_\text{R} + \text{GSS}_\text{G} + \text{GSS}_\text{B}) \tag{13}$$

TABLE III

CPSNR QUALITY COMPARISON AMONG $CS_2 \times CR_2$ FOR THE RECONSTRUCTED 26 UV IMAGES, 26 SCIs, AND 6 SCVs; CGSS QUALITY COMPARISON AMONG $CS_2 \times CR_2$ FOR THE RECONSTRUCTED SCIs AND SCVs

	$IDID_{NEDI}$-NEDI	$CS_{BILINEAR}$-LACDU	$CS_{BICUBIC}$-LACDU	$IDID_{NEDI}$-ASBLG	$CS_{BILINEAR}$-ASBLG	$CS_{BICUBIC}$-ASBLG
CPSNR	35.8405 (32.3445) [30.6089]	35.9215 (32.6666) [31.0319]	35.9464 (32.6305) [31.0340]	37.0045 (33.2682) [31.4922]	**37.1053** **(33.4037)** [31.4987]	37.0564 (33.2941) **[31.5172]**
CGSS	(0.2038) [0.2392]	(0.1953) [0.2233]	(0.1956) [0.2233]	(0.1841) [0.2110]	**(0.1804)** [0.2083]	(0.1822) **[0.2081]**
CPSNR Gain	1.2648 (1.0592) [0.9083]	1.1838 (0.7371) [0.4853]	1.1589 (0.7732) [0.4647]	0.1008 (0.1355) [0.0250]	 [0.0185]	0.0489 (0.1096)
CGSS Gain	(0.0234) [0.0311]	(0.0149) [0.0152]	(0.0152) [0.0152]	(0.0037) [0.0029]	 [0.0002]	(0.0018)

(a) (b) (c) (d) (e) (f) (g) (h) (i) (j) (k) (l) (m) (n) (o)

Fig. 9. The visual effect comparison for the 2nd reconstructed SCI. (a) The original SCI (b) The magnified subimage with characters. (c) The magnified subimage with colorful part. (d) After running 4:2:0(R)-LAU on Fig. 9(b). (e) After running 4:2:0(A)-GCR on Fig. 9(b). (f) After running 4:2:0(DIRECT)-SAI on Fig. 9(b). (g) After running $IDID_{NEDI}$-NEDI on Fig. 9(b). (h) After running $CS_{BICUBIC}$-LACDU on Fig. 9(b). (i) After running $CS_{BILINEAR}$-ASBLG on Fig. 9(b). (j) After running 4:2:0(R)-LAU on Fig. 9(c). (k) After running 4:2:0(A)-GCR on Fig. 9(c). (l) After running 4:2:0(DIRECT)-SAI on Fig. 9(c). (m) After running $IDID_{NEDI}$-NEDI on Fig. 9(c). (n) After running $CS_{BICUBIC}$-LACDU on Fig. 9(c). (o) After running $CS_{BILINEAR}$-ASBLG on Fig. 9(c).

where GSS_R, GSS_G and GSS_B indicate the GSS values of the reconstructed R image, G image, and B image, respectively, with

$$\mathrm{GSS} = \left(\frac{1}{WH}\sum_{p\in P}(\mathrm{GS}(p) - m_{GS})^2\right)^{\frac{1}{2}} \tag{14}$$

which $\mathrm{P} = \{(l,m)|1 \le l \le H, 1 \le m \le W\}$ and m_{GS} denotes the mean value of GS, with

$$\mathrm{GS}(p) = [\mathrm{DS}(p)] \cdot [\mathrm{MS}(p)] \tag{15}$$

where DS denotes the gradient direction similarity and MS denotes the gradient magnitude similarity, which are defined by

$$\begin{aligned}\mathrm{DS}(p) &= \frac{2D_o(p)D_r(p) + c_D}{D_o^2(p) + D_r^2(p) + c_D} \\ \mathrm{MS}(p) &= \frac{2G_o(p)G_r(p) + c_M}{G_o^2(p) + G_r^2(p) + c_M}\end{aligned} \tag{16}$$

where D_o and D_r denote the gradient direction maps of the original and reconstructed SCIs, respectively; G_o and G_r denote the gradient magnitude maps of the original and reconstructed SCIs, respectively; c_D and c_M are two small positive values, respectively. According to the definition of

CGSS (13), which is a useful quality metric to evaluate the reconstructed SCIs, the lower the average value of CGSS is, the higher the quality of the reconstructed SCIs is.

Table II indicates the average CGSS value of all the reconstructed SCIs in the parenthesis '()' for each combination in $CS_1 \times CR_1$. Among the 50 comparative combinations in $CS_1 \times CR_1$, 4:2:0(A)-ASBLG has the best average CGSS quality performance for the reconstructed SCIs. In the same table, the average CGSS value of the reconstructed SCVs for each comparative combination is listed in the parenthesis '[]'; we observe that 4:2:0(A)-ASBLG still has the best CGSS quality performance for the reconstructed among the concerned 50 comparative combinations.

2) CPSNR and CGSS Quality Performance Comparison Among the 6 Comparative Combinations in $CS_2 \times CR_2$: We now evaluate the CPSNR and CGSS quality performance in $CS_2 \times CR_2$ ({=$IDID_{NEDI}$-NEDI, $CS_{BILINEAR}$-LACDU, $CS_{BICUBIC}$-LACDU, $IDID_{NEDI}$-ASBLG, $CS_{BILINEAR}$-ASBLG, $CS_{BICUBIC}$-ASBLG}). Table III clearly indicates that as marked in boldface, $IDID_{NEDI}$-ASBLG, $CS_{BILINEAR}$-ASBLG, and $CS_{BICUBIC}$-ASBLG are superior to $IDID_{NEDI}$-NEDI, $CS_{BILINEAR}$-LACDU, and $CS_{BICUBIC}$-LACDU, respectively, in terms of the CPSNR and CGSS. Especially, $CS_{BILINEAR}$-ASBLG has the best CPSNR performance in $CS_2 \times CR_2$ for the reconstructed UV images and has the best CGSS performance for the reconstructed SCIs. Based on the reconstructed SCVs, $CS_{BICUBIC}$-ASBLG has the best CPSNR and CGSS performance among the six combinations.

We conclude that overall, in terms of CPSNR and CGSS, the combinations, $CS_{BILINEAR}$-ASBLG and $CS_{BICUBIC}$-ASBLG, can beat any comparative combinations in $CS_1 \times CR_1$ and $CS_2 \times CR_2$, in which the chroma subsampling schemes, $CS_{BILINEAR}$ and $CS_{BICUBIC}$, were proposed by Wang *et al.* [19] and ASBLG is our proposed chroma reconstruction method.

B. Visual Effect Merit

A subjective quality evaluation is conducted to demonstrate the visual merit of the proposed ASBLG method for the reconstructed SCIs. We take the second original SCI, as shown in Fig. 9(a), as the visual comparison example. For visual effect illustration, two magnified subimages, which are cut off from the characters part and the colorful part in Fig. 9(a), respectively, are shown in Figs. 9(b)-(c). The six comparative combinations are 4:2:0(R)-LAU, 4:2:0(A)-GCR, 4:2:0(DIRECT)-SAI, $IDID_{NEDI}$-NEDI, $CS_{BICUBIC}$-LACDU, and $CS_{BILINEAR}$-ASBLG. After performing the six combinations on Fig. 9(a), the twelve reconstructed magnified subimages for Fig. 9(b) and Fig. 9(c) are shown in Figs. 9(d)-(i) and Figs. 9(j)-(o). As shown in the regions marked by red ellipses, we clearly observe that $CS_{BILINEAR}$-ASBLG displays fewer color-shifting artifacts and has better visual quality than the other five state-of-the-art comparative combinations. It is suggested that the readers refer to the related experiment results in [4].

C. Quality-Bitrate Trade-Off

Eleven different quantization parameters (QPs) considered in the compression are 0, 4, 8, 12, 16, 20, 24, 28, 32, 36, and 40. Based on the eleven different QP values in HEVC, Fig. 10 illustrates the RD curves for the six comparative combinations, 4:2:0(R)-LAU in purple, 4:2:0(A)-GCR in green, 4:2:0(DIRECT)-SAI in orange, $IDID_{NEDI}$-NEDI in yellow, $CS_{BICUBIC}$-LACDU in blue, and $CS_{BILINEAR}$-ASBLG in red. From Fig. 10, we observe that under the same QP value, $CS_{BILINEAR}$-ASBLG always has the best CPSNR performance of the reconstructed SCIs among the six combinations considered.

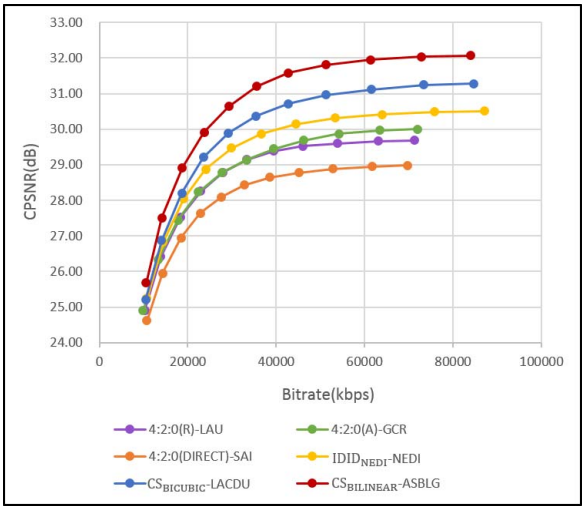


Fig. 10. RD curves for 4:2:0(R)-LAU, 4:2:0(A)-GCR, 4:2:0(DIRECT)-SAI, $IDID_{NEDI}$-NEDI, $CS_{BICUBIC}$-LACDU, and $CS_{BILINEAR}$-ASBLG.

V. Conclusion

In this paper, we have presented the propose ASBLG chroma reconstruction method. The main novelties and contributions of the proposed method include the following aspects. Based on the study of CDD between the subsampled luma and chroma block-pair under the same/different subsampling circumstance, it motivates us to propose a fast winner-first voting strategy to identify the used chroma subsampling scheme at the server side. Then, the decoded luma image is subsampled by the identified subsampling scheme, leading to an accurate correlation between the subsampled decoded luma block and the corresponding decoded subsampled chroma block. Further, we point out the out of the min-max range problem caused in the chroma pixel reconstruction. To alleviate this problem, an adaptive sliding window based approach is proposed. Based on the 26 test SCIs and 6 SCVs, thorough experiments are conducted to show that in terms of CPSNR and CGSS, $CS_{BILINEAR}$-ASBLG has the best quality performance for the reconstructed SCIs; $CS_{BICUBIC}$-ASBLG has the best CPSNR and CGSS for the reconstructed SCVs. Here, the chroma subsampling schemes, $CS_{BILINEAR}$ and $CS_{BICUBIC}$, were proposed by Wang *et al.* [19] and ASBLG is our proposed chroma reconstruction method. However, the natural image often has blurred edges and continuous-tone colors within objects, thus the correlation between the subsampled luma and chroma block-pair is not so high. Therefore, the proposed ASBLG method doesn't show favor to the natural image. For SCIs, one future research work is to classify each block as either the natural image block type or the SCI block type by using the intermediate information from the SCM (screen content

coding test model) encoder, e.g. the block coded by intra block copy or palette mode, and then to develop a hybrid chroma reconstruction method. Another future research work is to incorporate the result of the proposed winner-first voting strategy to identify the used chroma subsampling scheme in the data hiding area. The other future work is to develop an efficient method to robustly solve the case when the chroma subsampling scheme used is not in the list of the proposed fast winner-first voting strategy.

ACKNOWLEDGEMENT

The authors appreciate the source code of [8] provided by Dr. Korhonen, the test screen content video sequences provided by Dr. S. M. Lei and Dr. Y. W. Huang, the proofreading help of Ms. C. Harrington, and the valuable comments of the three anonymous referees to improve the manuscript.

REFERENCES

[1] T.-S. Chang, C.-C. Chen, R.-L. Liao, C.-W. Kuo, and W.-H. Peng, "Intra line copy for HEVC screen content coding," in *Proc. Asia–Pacific Signal Inf. Process. Assoc., Annu. Summit Conf.*, Dec. 2014, pp. 1–8.

[2] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms* (Asymptotic Notation), 3rd ed. London, U.K.: MIT Press, sec. 3.1, 2009.

[3] J. Allebach and P. W. Wong, "Edge-directed Interpolation," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 1996, pp. 707–710.

[4] *Experimental Results*. Accessed: Sep. 14, 2017. [Online]. Available: ftp://140.118.175.164/Experiment

[5] K. He, J. Sun, and X. Tang, "Guided image filtering," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 6, pp. 1397–1409, Jun. 2013.

[6] *Adaptive Basic Unit Layer Rate Control for JVT*, document JVT-G012, ISO/IEC JTC1/SC29/WG11 and ITU-T SG16/Q.6, 7th Meeting, Pattaya, Thailand, 2003.

[7] R. Keys, "Cubic convolution interpolation for digital image processing," *IEEE Trans. Acoust., Speech Signal Process.*, vol. ASSP-29, no. 6, pp. 1153–1160, Dec. 1981.

[8] J. Korhonen, "Improving image fidelity by luma-assisted chroma subsampling," in *Proc. IEEE Int. Conf. Multimedia Expo*, Jun./Jul. 2015, pp. 1–6.

[9] W.-S. Kim *et al.*, "Cross-component prediction in HEVC," *IEEE Trans. Circuits Syst. Video Technol.*, Nov. 2015. [Online]. Available: http://ieeexplore.ieee.org/document/7314885/

[10] X. Li and M. T. Orchard, "New edge-directed interpolation," *IEEE Trans. Image Process.*, vol. 10, no. 10, pp. 1521–1527, Oct. 2001.

[11] Y. Lu, S. Li, and H. Shen, "Virtualized screen: A third element for cloud-mobile convergence," *IEEE Multimedia Mag.*, vol. 18, no. 2, pp. 4–11, Feb. 2011.

[12] S. Minaee and Y. Wang, "Screen content image segmentation using least absolute deviation fitting," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2015, pp. 3295–3299.

[13] Z. Ni, L. Ma, H. Zeng, C. Cai, and K.-K. Ma, "Gradient direction for screen content image quality assessment," *IEEE Signal Process. Lett.*, vol. 23, no. 10, pp. 1394–1398, Oct. 2016.

[14] W. Pu *et al.*, "Palette mode coding in HEVC screen content coding extension," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 6, no. 4, pp. 420–432, Dec. 2016.

[15] *Requirements for an Extension of HEVC for Coding of Screen Content*, document ISO/IEC JTC1/SC29/WG11 MPEG2013/N14174, Jan. 2014.

[16] G. J. Sullivan, J.-R. Ohm, W.-J. Han, and T. Wiegand, "Overview of the high efficiency video coding (HEVC) standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1649–1668, Dec. 2012.

[17] Y.-C. Sun *et al.*, "Improved palette index map coding on HEVC SCC," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2016, pp. 4210–4214.

[18] T. Vermeir *et al.*, "Guided chroma reconstruction for screen content coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 26, no. 10, pp. 1884–1892, Oct. 2016.

[19] S. Wang, K. Gu, S. Ma, and W. Gao, "Joint chroma downsampling and upsampling for screen content image," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 26, no. 9, pp. 1595–1609, Sep. 2016.

[20] S. Wang, K. Gu, X. Zhang, W. Lin, S. Ma, and W. Gao, "Reduced-reference quality assessment of screen content images," *IEEE Trans. Circuits Syst. Video Technol.*, Aug. 2016. [Online]. Available: http://ieeexplore.ieee.org/document/7552436/

[21] S. Wang, L. Ma, Y. Fang, W. Lin, S. Ma, and W. Gao, "Just noticeable difference estimation for screen content images," *IEEE Trans. Image Process.*, vol. 25, no. 8, pp. 3838–3851, Aug. 2016.

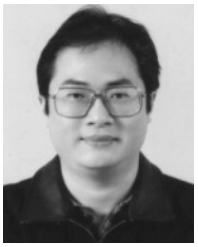
[22] J. Xu, R. Joshi, and R. A. Cohen, "Overview of the emerging HEVC screen content coding extension," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 26, no. 1, pp. 50–62, Jan. 2004.

[23] X. Xu, S. Liu, T. Chuang, and S. Lei, "Block vector prediction for intra block copying in HEVC screen content coding," in *Proc. Data Compress. Conf.*, Apr. 2015, pp. 273–282.

[24] H. Yu, R. Cohen, K. Rapaka, and J. Xu, *Common Test Conditions for Screen Content Coding*, document JCTVC-X1015, Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11 24th Meeting, Geneva, Switzerland, May/Jun. 2016.

[25] X. Zhang and X. Wu, "Image interpolation by adaptive 2-D autoregressive modeling and soft-decision estimation," *IEEE Trans. Image Process.*, vol. 17, no. 6, pp. 887–896, Jun. 2008.

[26] Y. Zhang, D. Zhao, J. Zhang, R. Xiong, and W. Gao, "Interpolation-dependent image downsampling," *IEEE Trans. Image Process.*, vol. 20, no. 11, pp. 3291–3296, Nov. 2011.



Kuo-Liang Chung (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 is currently the Chair Professor with the Department of Computer Science and Information Engineering, National Taiwan University of Science and Technology, Taipei. His research interests include image processing and video compression. He was a recipient of the Distinguished Research Award from 2004 to 2007 and the Distinguished Research Project Award from 2009 to 2010 from the National Science Council of Taiwan. In 2017, he received the Scientific Paper Award from the Far Eastern Y. Z. Hsu Science and Technology Memorial Foundation. He has been an Associate Editor of the *Journal of Visual Communication and Image Representation* since 2011.

Chi-Chao Huang received the B.S. degree in computer science and information engineering from the National ChengChi University, Taipei, Taiwan, in 2015. He is currently pursuing the M.S. degree in computer science and information engineering with the National Taiwan University of Science and Technology, Taipei. His research interests include image processing and video compression.

Tsu-Chun Hsu received the B.S. degree in computer science and information engineering from the National Taiwan Ocean Universi National Taiwan University of Science and ty, Keelung, Taiwan, in 2015. She is currently pursuing the M.S. degree in computer science and information engineering with the National Taiwan University of Science and Technology, Taipei, Taiwan. Her research interests include image processing and video compression.