Novel 2-D MMSE Subpixel-Based Image Down-Sampling

Lu Fang, Oscar C. Au, Ketan Tang, Xing Wen, and Hanli Wang

***Abstract*—Subpixel-based down-sampling is a method that can potentially improve apparent resolution of a down-scaled image on LCD by controlling individual subpixels rather than pixels. However, the increased luminance resolution comes at price of chrominance distortion. A major challenge is to suppress color fringing artifacts while maintaining sharpness. We propose a new subpixel-based down-sampling pattern called diagonal direct subpixel-based down-sampling (DDSD) for which we design a 2-D image reconstruction model. Then, we formulate subpixel-based down-sampling as a MMSE problem and derive the optimal solution called minimum mean square error for subpixel-based down-sampling (MMSE-SD). Unfortunately, straightforward implementation of MMSE-SD is computational intensive. We thus prove that the solution is equivalent to a 2-D linear filter followed by DDSD, which is much simpler. We further reduce computational complexity using a small $k \times k$ filter to approximate the much larger MMSE-SD filter. To compare the performances of pixel and subpixel-based down-sampling methods, we propose two novel objective measures: normalized l_1 high frequency energy for apparent luminance sharpness and $\mathrm{PSNR}_{\mathrm{U(V)}}$ for chrominance distortion. Simulation results show that both MMSE-SD and MMSE-SD(k) can give sharper images compared with conventional down-sampling methods, with little color fringing artifacts.**

***Index Terms*—Color fringing, image down-sampling, subpixel rendering.**

I. Introduction

THIS PAPER is about using a subpixel-based rendering technique to achieve higher apparent resolution of natural images than pixel-based rendering methods, when a down-sampled image is to be displayed on an LCD display. In pixel-based rendering, a pixel is considered as an inseparable region capable of representing a single color out of 2^{24} possible colors. In subpixel-based rendering, a pixel is considered as a region composed of subregions which are called subpixels, and each subpixel is capable of representing a fixed color (such as red, green, or blue) at different intensity.

Manuscript received April 10, 2011; revised September 8, 2011; accepted October 24, 2011. Date of publication December 1, 2011; date of current version May 1, 2012. This work was supported in part by the Research Grants Council of the Hong Kong Special Administrative Region, China, under Grant GRF 610109. This paper was recommended by Associate Editor L. Onural.

L. Fang, O. C. Au, K. Tang, and X. Wen are with the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong (e-mail: eefanglu@ust.hk; eeau@ust.hk; tkt@ust.hk; wxxab@ust.hk).

H. Wang is with the Department of Computer Science and Technology, Tongji University, Shanghai 200092, China (e-mail: hanliwang@tongji.edu.cn).

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

Digital Object Identifier 10.1109/TCSVT.2011.2179458

Subpixel rendering techniques originate from the problem of monochromatic font rendering on LCDs. Previously, simple pixel-based font display was used and the smallest level of detail that a computer could display on an LCD was a single pixel. On the other hand, subpixel rendering takes advantage of the fact that each pixel on a color LCD is actually composed of individual subpixel stripes as shown in Fig. 1 to give greater details of text [1]–[6]. Some major works were done by Apple and Microsoft. About 20 years ago, the Apple II personal computer introduced a proprietary high resolution LCD graphics display in which each pixel has two vertical stripe subpixels with green and magenta colors, respectively. The system controlled the individual subpixels in such a way that it could display fonts with better details than pixel-based rendering displays. Depending on the relative intensity of components in a pixel, the apparent position or orientation of a line (such as the edge of a font) can be microshifted (i.e., shifted by one or two subpixel width). Later, in 1998, Microsoft announced a subpixel-based font display technology called "ClearType," which is a software technology capable of improving the readability of small text on regular LCD with three vertical stripe subpixels (red, green, and blue). With ClearType running on an LCD monitor, features of text as small as a fraction of a pixel in width can be displayed.

Fig. 2 shows a common problem when a sloping edge is displayed by pixel rendering, and how it can be suppressed by subpixel rendering. Simple pixel-based rendering causes sawtooth in sloping edge in Fig. 2(a). Thanks to the fact that a pixel is composed of three separable subpixels, we can "borrow" subpixels from adjacent whole pixels. Fig. 2(b) depicts that using subpixel rendering, the apparent position of the sloping edge is microshifted by one or two subpixel width, giving a much smoother result compared to Fig. 2(a). However, subpixel rendering may cause local color imbalance called "color fringing artifact" [6], [9], [10], because, for some pixels, only one or two subpixels are turned on/off, as shown in Fig. 2(c).

In this paper, we are concerned about image down-sampling using subpixel techniques to achieve superior sharpness when a high resolution image or video is to be displayed on low resolution LCD display terminals. For example, while digital pictures are usually captured at very high resolution (e.g., 10 mega-pixels), many of them would be displayed on LCD computer monitors, photo-frames, or small LCD screens on mobile phones or PDAs, which have considerably lower resolutions (e.g., 0.8 mega-pixel on SVGA, or 0.2

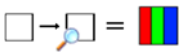


Fig. 1. Single white pixel appears as an inseparable region with solid color (left), and magnified showing separate R, G, and B subpixels (right).

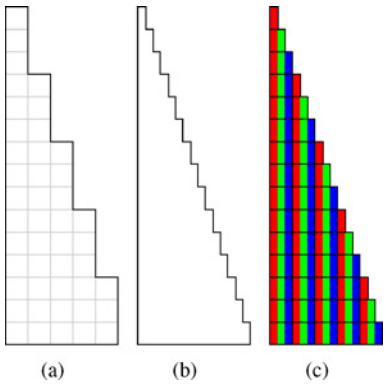


Fig. 2. Rendering of a sloping edge. (a) Pixel-based rendering. (b) Subpixel rendering (conceptual) result. (c) Subpixel rendering (actual color pattern).

mega-pixel on some smart phones). A similar situation exists for video, where full-HD (1920×1080) TV or full-HD movies in Blu-ray may be viewed on HD-ready (1366×768) or SD (720×576 or 720×480) TV or monitors. To view high resolution image/video on low resolution display, a down-sampling procedure is required.

A simple way, called direct pixel-based down-sampling (DPD) in this paper, is to perform simple down-sampling by selecting one out of every N pixels. (In this paper, the term *Direct* means no antialiasing filter is applied.) It can incur severe aliasing artifacts in regions with high spatial frequency [such as staircase artifacts and broken lines as shown in Fig. 3(b)]. An improved scheme is called pixel-based down-sampling with antialiasing filter (PDAF) in which an antialiasing filter is applied before direct pixel-based down-sampling. It suppresses aliasing artifacts at the price of blurring the image, as only the low frequency information can be retained in the process. Note that both DPD and PDAF are pixel-based methods and do not incur color artifacts.

Since the number of individual reconstruction points in LCD can be increased by three times by considering subpixels, application of subpixel rendering in down-sampling schemes may lead to improvement in apparent resolution. Higher apparent resolution is always attractive to consumers because, for a given physical size of display, higher resolution can give much more details making the image more realistic. For example, for two displays with the same physical size of 4 in $\times$ 3 in, a 704×576 display would look much better with much more details than a 352×288 resolution display. This is also the reason why HDTV looks nicer than SDTV, on displays with the same physical size.

Unfortunately, direct application of a subpixel approach to down-sampling may cause a color fringing problem and result in annoying artifacts [6], [9], [10]. Thus, some filtering is needed to suppress the color fringing artifacts without significant damage to the improved apparent resolution. This paper is about an optimal design of such a filter. In [6], Gibson used a five-tap low-pass filter to smooth the result of subpixel-based down-sampling. However, the low-pass filter relieves color fringing at the price of image blurring, and can be only adopted as an enhancement technique for achromatic image. Based on psychophysical experiments, Platt [7] defined an error metric in frequency domain, and derived the filter coefficients by minimizing this metric. Betrisey *et al*. [8] applied the results of Platt in [7] to achromatic (gray-scale) font rendering which is the basis of Microsoft's ClearType system. In [10] and [11], an algorithm based on human visual system is proposed to suppress visible chrominance aliasing. However, all these methods process subpixel-based down-sampling horizontally. Researchers typically do not attempt to apply subpixel-based down-sampling to vertical down-sampling as there is a common conception that little can be gained in the vertical direction due to the horizontal arrangement of the subpixels. In [12], we formulated subpixel-based down-sampling as a directional min-max problem and showed that the min-max solution can give superior performance over other subpixel-based down-sampling methods in terms of apparent sharpness. However, there are still considerable remaining color fringing artifacts.

In this paper, we propose an optimal 2-D MMSE subpixel-based down-sampling scheme which we call "MMSE-SD." In Section II, we start by proposing a novel diagonal direct subpixel-based down-sampling (DDSD) pattern. Then we propose a 2-D "virtual model" for a high resolution image reconstructed from the corresponding low-resolution image. Based on the virtual model, we formulate subpixel-based down-sampling as a MMSE problem and solve for the optimal solution which turns out to be very complicated. We then prove that the optimal solution is equivalent to a 2-D linear filter followed by DDSD, and such an implementation has significantly lower computational complexity. To further reduce the computational complexity, we find a simplified 2-D linear filter with even lower complexity and yet similar visual quality. To evaluate the proposed MMSE-SD, we also propose novel objective measures for apparent luminance sharpness and chrominance distortion in Section III. We then give the experiment results in Section IV. Finally, Section V concludes this paper.

II. MMSE Subpixel-Based Down-Sampling (MMSE-SD)

A. Existing Down-Sampling Schemes

For simplicity, we assume that an input high resolution image L (meaning *large*) of size $3M \times 3N$ is to be down-sampled to a low resolution image S (meaning *small*) of size $M \times N$, and to be displayed on a $M \times N$ device. (Note that if L is not of size $3M \times 3N$, i.e., the down-sampling ratio is not 3, we can use regular interpolation or decimation methods to resize L to be $3M \times 3N$.) Daly *et al*. proposed a simple subpixel-based down-sampling pattern which we call direct subpixel-based down-sampling (DSD). DSD decimates the red, green, and blue components alternately in the horizontal direction [9]–[11]. Let $(r_{i,j}, g_{i,j}, b_{i,j})$ be the (i, j)th pixel of S. DSD copies red, green and blue components (i.e., the three

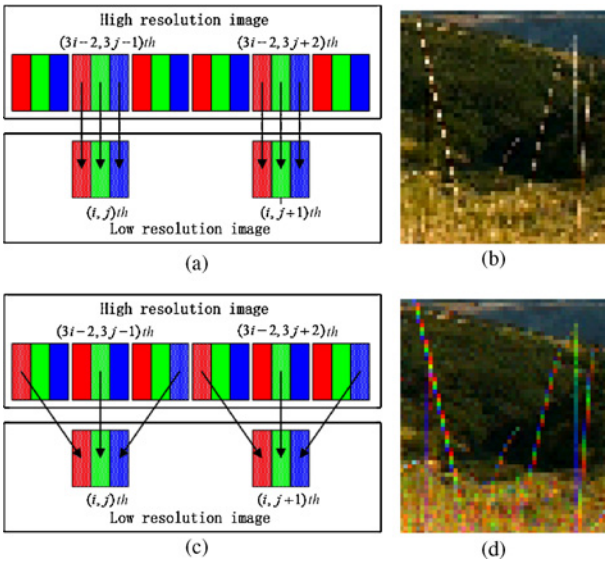


Fig. 3. (a) DPD. (b) Magnified result of DPD, where "grass" is broken due to aliasing artifacts. (c) DSD. (d) Magnified result of DSD, where "grass" is smooth but has color fringing artifacts.

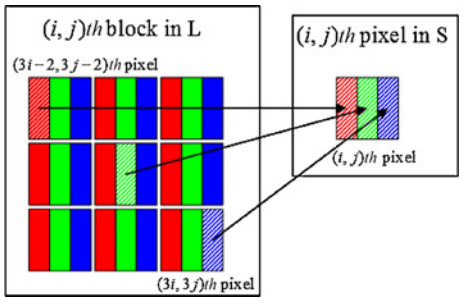


Fig. 4. Proposed DDSD pattern.

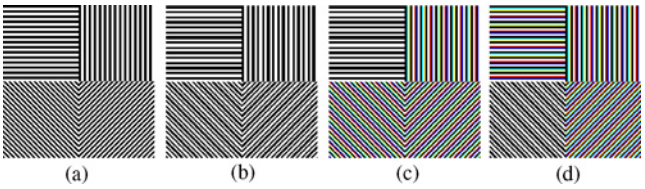


Fig. 5. Artificial image with four subimages. (a) Original L image. (b) Result of DPD. (c) Result of DSD. (d) Result of DDSD.

subpixels) of the pixel from three different pixels in L, such that $r_{i,j} = R_{3i-2,3j-2}$, $g_{i,j} = G_{3i-2,3j-1}$, $b_{i,j} = B_{3i-2,3j}$ as shown in Fig. 3(c), where $R_{3i-2,3j-2}$ is the red component of the $(3i-2, 3j-2)$th pixel of L and so on. It is clear that DSD considers only the horizontal direction, but not the vertical.

Fig. 3 depicts the resultant images of two down-sampling patterns: DPD and DSD. It is interesting to see that DSD can potentially preserve more details than DPD thanks to the increase in the number of individual reconstruction points. A close examination of Fig. 3(d) reveals that DSD fills in the gaps of the grass, making the grass continuous and sharp at the expense of annoying color artifacts.

Therefore, exploiting subpixels in down-sampling brings both opportunity and problem. The opportunity is that we can potentially increase the apparent resolution of a patterned display up to the subpixel resolution. The problem is the associated color distortion. The goal of our paper is to achieve subpixel resolution (i.e., apparent luminance resolution) while suppressing color artifacts (i.e., chrominance distortion).

B. Proposed DDSD Pattern

In our experiments, we observe that the improvement of apparent resolution in DSD tends to happen at regions with vertical edges or edges with vertical component. There is typically no improvement at smooth regions or regions with horizontal edges, due to the fact that in DSD the sampling pattern is merely in a horizontal way, which is parallel to horizontal edges. To achieve improved resolution in both horizontal and vertical directions, we propose a DDSD pattern, changing the sampling direction from horizontal to diagonal. We divide original image L into 3×3 blocks so that there are $M \times N$ blocks, one for each pixel in the down-sampled low resolution image S, such that the (i, j)th block in L corresponds to the (i, j)th pixel in S. For the (i, j)th pixel in S, DDSD copies the red, green and blue components from three different pixels in the (i, j)th block of L along the diagonal direction. (Our method works also for antidiagonal direction but we use diagonal in this paper.) Fig. 4 shows an example of DDSD

$$r_{i,j} = R_{3i-2,3j-2} \quad g_{i,j} = G_{3i-1,3j-1} \quad b_{i,j} = B_{3i,3j}. \tag{1}$$

To further understand the potential and limitation of various down-sampling schemes, we generate an artificial large image (L) of size 420×420, containing four subimages as shown in Fig. 5. The four subimages, which we call Subimage-H, Subimage-V, Subimage-D, and Subimage-AD, contain 15 pairs of black and white lines in horizontal, vertical, diagonal, and antidiagonal directions, respectively. The width of each black or white line is 7 pixels (with a total of 21 subpixels). In our experiment, L is down-sampled by a factor of 3 with DPD, DSD and DDSD to produce three 140×140 images, as shown in Fig. 5(b), (c), and (d), respectively.

We then define a subpixel-based regularity measure for each subimage

$$\mu = \frac{\sum_{k=1}^{m} w_k}{m} - \frac{w_0}{3} \qquad \sigma^2 = \frac{\sum_{k=1}^{m} (w_k - w_0/3)^2}{m} \tag{2}$$

where m is the number of black lines, w_0 is the width of black lines in the L image and w_k $(k = 1, \ldots, m)$ is the width of the kth black line in DPD, DSD or DDSD image, and the unit of w_0 and w_k is *subpixel*. In our experiment, $m = 15$ and $w_0 = 21$. The mean μ and variance σ^2 of the line width of DPD, DSD, and DDSD are shown in Table I. To account for color fringing artifacts caused by subpixel-based down-sampling, we introduce a simple color distortion measure for each subimage as follows:

$$\Delta_{\mathrm{RGB}} = \sum_{i=1,j=1}^{M,N} \min\left\{|\mathbf{c}_{i,j} - \mathbf{0}|, |\mathbf{c}_{i,j} - \mathbf{255}|\right\}/255 \tag{3}$$

TABLE I
LINE WIDTH AND COLOR DISTORTION OF DPD, DSD, AND DDSD

	Subimage-H			Subimage-V		
	DPD	DSD	DDSD	DPD	DSD	DDSD
μ	0	0	0	0	0	0
σ^2	2	2	0	2	0	0
Δ_{RGB}	0	0	2	0	2	2
	Subimage-D			Subimage-AD		
	DPD	DSD	DDSD	DPD	DSD	DDSD
μ	0	0	0	0	0	0
σ^2	2	0	2	2	0	0
Δ_{RGB}	0	2	0	0	2	2

where $\mathbf{c}_{i,j} = \begin{bmatrix} r_{i,j} & g_{i,j} & b_{i,j} \end{bmatrix}^T$, $\mathbf{0} = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}^T$, and $\mathbf{255} = \begin{bmatrix} 255 & 255 & 255 \end{bmatrix}^T$. Examining (3), the value of $\min\left\{|\mathbf{c}_{i,j} - \mathbf{0}|, |\mathbf{c}_{i,j} - \mathbf{255}|\right\}$ would be either 0 or 255. And Δ_{RGB} indicates the frequency (how often) of color artifacts happen for $i = 1, 2, \ldots, M$ and $i = 1, 2, \ldots, N$. Due to the 3:1 down-sampling ratio, the behavior (frequency) of color artifacts is periodic every three black/while lines (21 pixel width) in L. In other words, the frequency of color artifacts is given by computing Δ_{RGB} for a 7×7 ($M = 7, N = 7$) block in S, as shown in Table I. According to definition, $\Delta_{\text{RGB}} = 0$ indicates the result is free of color distortion, while $\Delta_{\text{RGB}} \neq 0$ suggests the result has color artifacts.

As expected, $\mu = 0$ for all methods, suggesting that the average line width is correct for all methods. For the DPD image, σ^2 is nonzero in all the four directions, indicating that the line spacing of DPD is irregular, as verified in Fig. 5(b). For the DSD image, σ^2 is nonzero in Subimage-H, indicating that DSD may not manage to keep the horizontal line spacing regular, as verified in Fig. 5(c), due to the horizontal decimation of DSD. On the contrary, DDSD manages to keep the line spacing regular for both horizontal and vertical lines at the expense of color fringing artifacts in Fig. 5(d). Of course, DDSD has its own limitation too, which cannot keep the line spacing regular for Subimage-D. Fortunately, diagonal edges tend to occur less frequently than horizontal and vertical edges in real situations.

Examining Table I, we find $\Delta_{\text{RGB}}(DPD) = 0$ for all subimages, suggesting that DPD is free of color artifacts. Both DSD and DDSD have nonzero Δ_{RGB} for three of the four subimages. For DSD, the three are Subimage-V, Subimage-D, Subimage-AD. For DDSD, the three are Subimage-H, Subimage-V, and Subimage-AD. So DSD and DDSD achieve higher apparent resolution at the expense of color artifacts. As the line width of the original image is 7 pixels which is not a multiple of 3, both DSD and DDSD would sample across the boundary of the black and white lines in four possible ways: (black, white, white), (white, white, black), (white, black, black), or (black, black, white), leading to four corresponding colors: cyan (0, 255, 255), yellow (255, 255, 0), red (255, 0, 0), and blue (0, 0, 255), as shown in Fig. 5(c).

C. 2-D Virtual Reconstruction Model

Here, we need to develop a tool to enable us to compare the similarity between the original high-resolution image L (of size $3M \times 3N$) and a down-sampled low-resolution image S (of size $M \times N$) generated using one of a variety of down-sampling methods, with the goal of finding a down-sampling method that gives maximum similarity. Our approach is to develop in this subsection a method to reconstruct a *virtual image* L′ of size $3M \times 3N$ from S. Then, in the next subsection, we can measure the mean square error between L and L′.

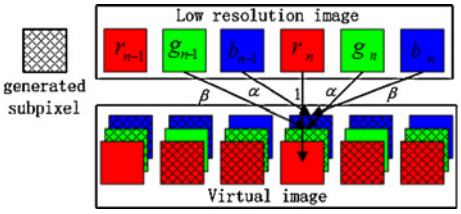


Fig. 6. 1-D reconstruction model for virtual image generation.

To generate the virtual image L′, we extend the existing 1-D reconstruction method by Kim [13] as illustrated in Fig. 6 to our 2-D problem. Assuming a matrix display (i.e., stripe LCD display), Kim constructs 3 full pixels of the virtual image out of every pixel (r_n, g_n, b_n) of the down-sampled image, one full pixel at $3n-2$ for r_n, one at $3n-1$ for g_n, and one at $3n$ for b_n. As a subpixel such as r_n can already contribute one color component to the full pixel, there are two missing color components to be generated. The missing green component is estimated by $\beta g_{n-1} + \alpha g_n$, where $\alpha + \beta = 1$. A similar expression is used to estimate the blue component.

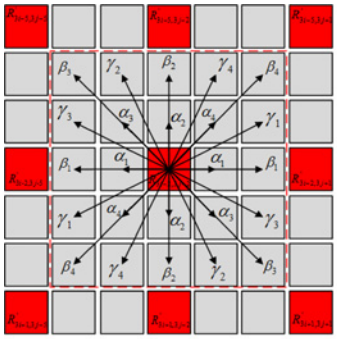
In a similar way, we generate L′ from S, except that we use the 2-D neighborhood instead of the 1-D neighborhood, as illustrated in Fig. 7 where each square represents a pixel in L′. Note that the (i, j)th pixel $(r_{i,j}, g_{i,j}, b_{i,j})$ in S corresponds to a 3×3 block of pixels in L′ at locations (k, l), with $k = 3i-2$, $3i-1$, or $3i$, and $l = 3j-2$, $3j-1$, or $3j$. We copy $r_{i,j}$ to location $(3i-2, 3j-2)$, $g_{i,j}$ to $(3i-1, 3j-1)$ and $b_{i,j}$ to $(3i, 3j)$. For the missing components, we generate them using weighted average of neighboring available components. For example, in Fig. 7, the missing red components at locations $(3i-3, 3j-3)$ and $(3i-4, 3j-4)$ are generated as $\alpha_3 r_{i,j} + \beta_3 r_{i-1,j-1} + \gamma_1 r_{i,j-1} + \gamma_4 r_{i-1,j}$ and $\beta_3 r_{i,j} + \alpha_3 r_{i-1,j-1} + \gamma_4 r_{i,j-1} + \gamma_1 r_{i-1,j}$, respectively, and those at locations $(3i-2, 3j-1)$ and $(3i-2, 3j)$ are $\alpha_1 r_{i,j} + \beta_1 r_{i,j+1}$ and $\beta_1 r_{i,j} + \alpha_1 r_{i,j+1}$, respectively. The missing values of the other two colors are generated in a similar way.

D. MMSE Subpixel-Based Down-Sampling (MMSE-SD)

After generating R′, G′, B′ using the reconstruction virtual model in Fig. 7, we have a virtual image L′ with the same resolution as the original high resolution image L. To find a good S, we formulate the problem as a MMSE problem by minimizing the mean square error (MSE) between L and L′

$$\begin{aligned} \min_{\mathbf{r},\mathbf{g},\mathbf{b}} \quad & ||\mathbf{R} - \mathbf{R}'||_2^2 + ||\mathbf{G} - \mathbf{G}'||_2^2 + ||\mathbf{B} - \mathbf{B}'||_2^2 \\ \text{s.t.} \quad & \mathbf{R}' = H_{\text{interp}} * (\mathbf{r} \otimes I_r) \\ & \mathbf{G}' = H_{\text{interp}} * (\mathbf{g} \otimes I_g) \\ & \mathbf{B}' = H_{\text{interp}} * (\mathbf{b} \otimes I_b) \end{aligned} \tag{4}$$

where $(\mathbf{r}, \mathbf{g}, \mathbf{b})$ are the $M \times N$ color components of S, $(\mathbf{R}, \mathbf{G}, \mathbf{B})$ and $(\mathbf{R}', \mathbf{G}', \mathbf{B}')$ are the $3M \times 3N$ color components of L and



(a)

$$H_{\text{interp}} = \begin{bmatrix} \beta_3 & \gamma_2 & \beta_2 & \gamma_4 & \beta_4 \\ \gamma_3 & \alpha_3 & \alpha_2 & \alpha_4 & \gamma_1 \\ \beta_1 & \alpha_1 & 1 & \alpha_1 & \beta_1 \\ \gamma_1 & \alpha_4 & \alpha_2 & \alpha_3 & \gamma_3 \\ \beta_4 & \gamma_4 & \beta_2 & \gamma_2 & \beta_3 \end{bmatrix}$$

(b)

Fig. 7. (a) 2-D reconstruction model for red component generation, where $\alpha_k, \beta_k, k = 1, 2, 3, 4$, are weighting coefficients in horizontal, vertical, diagonal, and antidiagonal directions and γ_k is for the 22.5° directions. (b) Equivalent interpolation filter.

L′, respectively. I_r, I_g, I_b are 3×3 matrices

$$I_r = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad I_g = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad I_b = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{5}$$

and $\otimes$ is the Kronecker product that expands $\mathbf{r}, \mathbf{g}, \mathbf{b}$ to be $3M \times 3N$ matrices. H_{interp} is the interpolation filter defined in Fig. 7 and $*$ is convolution operation. Without loss of generality, we use regular bilinear interpolation ($H_{\text{interp}} = H_{\text{BI}}$) in the following derivation. The derivation for the directional interpolation (H_{DI}) is similar.

Note that the optimization problem in (4) can be decomposed into three individual optimization problems. Here, we consider the red component optimization

$$\begin{aligned} \min_{\mathbf{r}} \quad & ||\mathbf{R} - \mathbf{R}'||_2^2 \\ \text{s.t.} \quad & \mathbf{R}' = H_{\text{BI}} * (\mathbf{r} \otimes I_r). \end{aligned} \tag{6}$$

Let $E \triangleq ||\mathbf{R} - \mathbf{R}'||_2^2$ with $\mathbf{R}' = H_{\text{BI}} * (\mathbf{r} \otimes I_3^r)$. Then E would contain $3M \times 3N$ terms. Consider a particular (i, j), we are interested in those terms containing $r_{i,j}$. While there are altogether 25 such terms (inside the red square of Fig. 7), we have limited space and will thus only show the nine terms at location (k, l), with $k = 3i-3, 3i-2, 3i-1$ and $l = 3j-3, 3j-2, 3j-1$ in (7), at the bottom of the page.

Differentiating E with respect to $r_{i,j}$ and setting it to zero, we obtain (8), where $k_0 = 1+2\alpha_1^2+2\beta_1^2+2\alpha_2^2+2\beta_2^2+2\alpha_3^2+2\beta_3^2+2\alpha_4^2+2\beta_4^2+2\gamma_1^2+2\gamma_2^2+2\gamma_2^2+2\gamma_4^2$, $k_1 = 2\alpha_1\beta_1+2\alpha_3\gamma_1+2\beta_3\gamma_4+2\alpha_4\gamma_3+2\beta_4\gamma_2$, $k_2 = 2\alpha_2\beta_2+2\alpha_3\gamma_4+2\beta_3\gamma_1+2\alpha_4\gamma_2+2\beta_4\gamma_3$, $k_3 = 2\alpha_3\beta_3+2\gamma_2\gamma_3$, and $k_4 = 2\alpha_4\beta_4+2\gamma_1\gamma_4$. Considering all $i = 1, 2, \ldots, M$, $j = 1, 2, \ldots, N$, we have a total of $M \times N$ equations which can be represented in a matrix form as

$$\mathbf{H}_r \; r = \mathbf{H}_R \; R \tag{9}$$

where $\mathbf{H}_r$ and $\mathbf{H}_R$ are matrices of sizes $MN \times MN$ and $MN \times 9MN$, respectively. The expressions of $\mathbf{H}_r$ and $\mathbf{H}_R$ are given in Appendix A. R is a row-ordered vector of size $9MN \times 1$ from the red component of L, and r is a row-ordered vector of size $MN \times 1$ from the red component of S, respectively. Solving (9) by taking inverse of $\mathbf{H}_r$, we get our MMSE-SD solution

$$r = (\mathbf{H}_r^{-1}\mathbf{H}_R) \; R = \mathbf{H} \; R \tag{10}$$

where $\mathbf{H} = \mathbf{H}_r^{-1}\mathbf{H}_R$ is a matrix of size $MN \times 9MN$.

E. Analysis and Simplification of MMSE-SD

The optimal output r is obtained by multiplying $\mathbf{H}$ with the input R as shown in (10), where $\mathbf{H} = \mathbf{H}_r^{-1}\mathbf{H}_R$. The total computational complexity of the optimal solution of MMSE-SD is $\mathcal{O}(M^3N^3)$. Nowadays, digital images can easily contain 10 mega pixels. For simplicity, suppose $M = 1000$ and $N = 1000$, $\mathbf{H}$ would have 9×10^{12} elements. If each element is a floating point number requiring 4 bytes to store, $\mathbf{H}$ would require 36 TB of memory. The multiplication of $\mathbf{H}_r^{-1}$ and $\mathbf{H}_R$ would require $\mathcal{O}(10^{22})$ operations. It is obvious the computation and memory requirements of MMSE-SD are excessive. Therefore, it is not sensible to compute r directly from (10). This motivates us to: 1) express MMSE-SD as a 2-D spatial-invariant linear filter followed by a 3:1 down-sampling which reduces computational complexity to $\mathcal{O}(M^2N^2)$; and 2) approximate the 2-D filter by restricting its filter size to $k \times k$, which further reduces the complexity to $\mathcal{O}(MN)$ with minor performance degradation.

1) *Express MMSE-SD as a 2-D Filter Followed by 3:1 Down-Sampling:* Note that $\mathbf{H}_r$ and $\mathbf{H}_R$ are block-circulant with some special structures, as detailed in Appendix A. We give an analytical proof in Appendix B that $\mathbf{H} = \mathbf{H}_r^{-1}\;\mathbf{H}_R$ is a block-circulant matrix as well, with blocks of size $N \times 9N$.

$$\begin{aligned} E = \cdots \quad & + \big[(R_{3i-3,3j-3} - (\alpha_3 r_{i,j} + \beta_3 r_{i-1,j-1} + \gamma_1 r_{i,j-1} + \gamma_4 r_{i-1,j}))^2 + (R_{3i-3,3j-2} - (\alpha_2 r_{i,j} + \beta_2 r_{i-1,j}))^2 \\ & + (R_{3i-3,3j-1} - (\alpha_4 r_{i,j} + \beta_4 r_{i-1,j+1} + \gamma_2 r_{i-1,j} + \gamma_3 r_{i,j+1}))^2 \\ & + (R_{3i-2,3j-3} - (\alpha_1 r_{i,j} + \beta_1 r_{i,j-1}))^2 + (R_{3i-2,3j-2} - r_{i,j})^2 + (R_{3i-2,3j-1} - (\alpha_1 r_{i,j} + \beta_1 r_{i,j+1}))^2 \\ & + (R_{3i-1,3j-3} - (\alpha_4 r_{i,j} + \beta_4 r_{i+1,j-1} + \gamma_3 r_{i,j-1} + \gamma_2 r_{i+1,j}))^2 + (R_{3i-1,3j-2} - (\alpha_2 r_{i,j} + \beta_2 r_{i+1,j}))^2 \\ & + (R_{3i-1,3j-1} - (\alpha_3 r_{i,j} + \beta_3 r_{i+1,j+1} + \gamma_4 r_{i+1,j} + \gamma_1 r_{i+1,j+1}))^2\big] + \cdots \end{aligned} \tag{7}$$

$$\begin{aligned} & k_3 r_{i-1,j-1} + k_2 r_{i-1,j} + k_4 r_{i-1,j+1} + k_1 r_{i,j-1} + k_0 r_{i,j} + k_1 r_{i,j+1} + k_4 r_{i+1,j-1} + k_2 r_{i+1,j} + k_3 r_{i+1,j+1} \\ & = \beta_3 R_{3i-4,3j-4} + \gamma_2 R_{3i-4,3j-3} + \beta_2 R_{3i-4,3j-2} + \gamma_4 R_{3i-4,3j-1} + \beta_4 R_{3i-4,3j} + \gamma_3 R_{3i-3,3j-4} + \alpha_3 R_{3i-3,3j-3} \\ & + \alpha_2 R_{3i-3,3j-2} + \alpha_4 R_{3i-3,3j-1} + \gamma_1 R_{3i-3,3j} + \beta_1 R_{3i-2,3j-4} + \alpha_1 R_{3i-2,3j-3} + R_{3i-2,3j-2} + \alpha_1 R_{3i-2,3j-1} \\ & + \beta_1 R_{3i-2,3j} + \gamma_1 R_{3i-1,3j-4} + \alpha_4 R_{3i-1,3j-3} + \alpha_2 R_{3i-1,3j-2} + \alpha_3 R_{3i-1,3j-1} + \gamma_3 R_{3i-1,3j} \\ & + \beta_4 R_{3i,3j-4} + \gamma_4 R_{3i,3j-3} + \beta_2 R_{3i,3j-2} + \gamma_2 R_{3i,3j-1} + \beta_3 R_{3i,3j}. \end{aligned} \tag{8}$$

Each $N \times 9N$ block contains three sub-blocks of size $N \times 3N$, and each of which is block-circulant with a block size of 1×3. In this paper, we describe such $N \times 9N$ matrix as "block tri-circulant." More details of block tri-circulant matrices can be found in Appendix A.

With the structure of $\mathbf{H}$ described above, there are a total of $M \times 3M$ sub-blocks each of size $N \times 3N$. Each sub-block of $\mathbf{H}$ is block-circulant with a block-size of 1×3. In other words, each row of the (k, l)th sub-block, $k = 1, ..., M$ and $l = 1, ..., 3M$, has $3N$ coefficients and is equal to the previous row rotated to the right by three positions. Note that the mth row of the (k, l)th sub-block performs an inner-product with the lth row of L and adds a term to the mth element of the kth row of S, i.e., the (k, m)th pixel of S.

Consider the first row of the (k, l)th sub-block, we can generate a $3N \times 3N$ circulant matrix with the same first row. Then the sub-block is equal to the product of a matrix $\mathbf{I}_a$ of size $N \times 3N$ and this circulant matrix, where

$$\mathbf{I}_a(i, j) = \begin{cases} 1, & i = 1, \ldots, N \quad j = 3i - 2 \\ 0, & \text{otherwise.} \end{cases} \tag{11}$$

Note that $\mathbf{I}_a$ performs effectively a 3:1 down-sampling on the circulant matrix in the vertical direction. Thus, the operation of the (k, l)th sub-block on L is effectively a 1-D convolution of its first row (with $3N$ coefficients) with the periodic extension of the lth row of L to give a row of size $3N$, followed by a 3:1 down-sampling in the horizontal direction before being added to the kth row of S.

Analyzing in the same way, the operation of the whole first row of sub-blocks (with $k = 1$, $l = 1, ..., 3M$) on L is effectively a 2-D spatial-invariant linear filter applied on the periodic extension of the entire L to give a row of size $3N$, followed by a 3:1 down-sampling in the horizontal direction. This gives the first row of S. The coefficients of the 2-D spatial-invariant filter come from the first row of $\mathbf{H}$.

Recall that $\mathbf{H}$ is a block-circulant matrix, with blocks of size $N \times 9N$. Each $N \times 9N$ block is block-tri-circulant. In other words, each $N \times 9N$ block contains three sub-blocks of size $N \times 3N$, each of which is block-circulant. Each row of the $N \times 9N$ blocks can also be considered simply as a row of sub-blocks, each of size $N \times 3N$. Then each row of sub-blocks is simply the previous row rotated to the right by three sub-block positions.

Consider the first row of sub-blocks (with $k = 1$, $l = 1, ..., 3M$), with each sub-block being of size $N \times 3N$. We can generate a block-circulant matrix of size $3MN \times 9MN$, with a block size of $N \times 3N$. Then $\mathbf{H}$ is equal to the product of a matrix $\mathbf{I}_b$ of size $MN \times 3MN$ and this block-circulant matrix. The matrix $\mathbf{I}_b$ contains $M \times 3M$ sub-blocks, each of size $N \times N$, and the (i, j)th sub-block is an $N \times N$ identity matrix ($\mathbf{I}_N$) if $j = 3i - 2$. Otherwise, the (i, j)th sub-block is an $N \times N$ zero matrix ($\mathbf{0}_N$)

$$\mathbf{I}_b(i, j) = \begin{cases} \mathbf{I}_N, & i = 1, \ldots, M \quad j = 3i - 2 \\ \mathbf{0}_N, & \text{otherwise.} \end{cases} \tag{12}$$

Note that the $\mathbf{I}_b$ performs effectively a 3:1 down-sampling on the sub-blocks of the block-circulant matrix in the vertical direction. Thus, the operation of $\mathbf{H}$ on L is effective a 2-D spatial-invariant filter applied on the periodic extension of the entire L to give an image of size $3N \times 3N$, followed by a 3:1 down-sampling in the horizontal direction, and then finally a 3:1 down-sampling in the vertical direction. Alternatively, MMSE-SD can be obtained by simply applying the 2-D filter to the circular extension of L at the down-sampling locations.

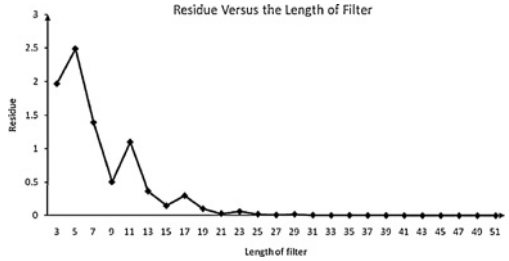


Fig. 8. Absolute approximation error between S and S(k).

This 2-D filtering can greatly reduce the computational complexity and storage requirement of MMSE-SD. By definition, the 2-D filter from $\mathbf{H}$ is of size $3M \times 3N$. It is independent of the input image and thus can be pre-computed and stored. This requires the storage of $9MN$ coefficients which is much fewer than the $9M^2N^2$ coefficients for direct implementation of (10). As for computational complexity, the 2-D filtering requires $\mathcal{O}(M^2N^2)$ operations which is much less than $\mathcal{O}(M^3N^3)$. For the case of $M = 1000$ and $N = 1000$, the storage requirement of the 2-D filter is about 36 MB, which is 10^6 times smaller than 36 TB for direct implementation of (10). The computational complexity of 2-D filtering is $\mathcal{O}(10^{12})$, which is 10^6 times smaller than $\mathcal{O}(10^{18})$ of (10).

2) *Approximate Optimal Filter by $k \times k$ Restricted Filter:* To further reduce the complexity of the 2-D filtering, we explore the possibility of reducing the filter size. We observe that many of the filter coefficients in MMSE-SD have very small magnitude and thus it may be possible to approximate them by zero without too much performance degradation. We perform an experiment to study the filter restricted to a size of $k \times k$. We will call MMSE-SD with this restricted $k \times k$ filter as MMSE-SD(k).

When the original optimal 2-D filter is used, the output of MMSE-SD is S. Let S(k) be the output of MMSE-SD(k). We compute the average absolute approximation error (difference) of S and S(k) (l_1 norm) for many different images. The typical behavior is shown in Fig. 8. We observe that, when k is large, the approximation error is approximately zero as expected. It appears that $k = 9$ is a good compromise between complexity and approximation error. We thus choose $k = 9$ in our experiments.

With $k = 9$, the size of the 2-D spatial filter in MMSE-SD(k) is 9×9. We observe that the 9×9 filter coefficients are symmetric in such a way that six coefficients appear eight times, eight coefficients appear four times, and one coefficient appears one time. Consequently, only $6 + 8 + 1 = 15$ different coefficients need to be stored and only 15 multiplications are needed for every down-sampled location.

As for computational complexity, the 2-D filtering in MMSE-SD ($k = 9$) requires $\mathcal{O}(MN)$ operations which is much less than $\mathcal{O}(M^2N^2)$ of the full 2-D filtering. These are summa-

TABLE II
STORAGE REQUIREMENT (IN BYTES) AND COMPUTATIONAL COMPLEXITY, VERSUS $M \times N$

	Gibson	Kim	MMDE-ER	MMSE	MMSE(9)
Storage	20	36	60	36MN	60
Complexity	$\mathcal{O}(MN)$	$\mathcal{O}(MN)$	$\mathcal{O}(MN)$	$\mathcal{O}(M^2N^2)$	$\mathcal{O}(MN)$

rized in Table II, assuming 4 bytes are used to store a floating point coefficient. For comparison's sake, we have also included those of Gibson [6], Kim [13], and Min-Max [12].

III. PROPOSED OBJECTIVE MEASURE

Until now, as far as we know, all researchers use subjective measures to evaluate the output images of subpixel-based down-sampling [14], [15]. There are no systematic, repeatable objective measures that can be used for evaluating such images, which motivates us to propose a few measures in this section. In subpixel-based down-sampling, our goal is to improve apparent luminance resolution (or apparent sharpness) and to suppress chrominance distortion (color fringing artifacts). We need different measures for luminance and chrominance components. The measures for the luminance component should measure its apparent sharpness, while the measures for the chrominance component should measure its chrominance distortion. Here we choose YUV color space [16], though other color spaces could have been used also.

A. Measures for Luminance Sharpness

As pointed out in Section II, in smooth regions with low frequency details, DSD barely provides improvement over DPD. The major resolution differences of various methods happen in their high frequency details. Thus, our luminance component measure computes the average of directional high-frequency "energy" (l_1 or l_2 norm). For any image x to be measured, we convolve it with four filters, H_a^k, with $k = 1, 2, 3, 4$. Basically, the four filters are simply the same 1-D high-pass filter $a = [1 - 1]$ applied in horizontal ($k = 1$), vertical ($k = 2$), diagonal ($k = 3$), and antidiagonal ($k = 4$) directions. Then, the average of directional high frequency energy (HFE) is given by $\frac{1}{4}\sum_{k=1}^{4} ||H_a^k * x||_1$ or $\frac{1}{4}\sum_{k=1}^{4} ||H_a^k * x||_2^2$.

Basically an image with larger HFE is sharper. As HFE values can be very large and difficult for the readers to comprehend, we use HFE of DDSD as a normalization factor as shown in (13). The normalized values P_{a1} and P_{b1} are typically between 0 and 2. We find DDSD as a reasonable and convenient choice as the DDSD images should be very sharp. A image with P_{a1} and P_{b1} larger than 1 is sharper than DDSD. However, HFE of other schemes (such as DPD, DSD, and PDAF) can also be used as the normalization factor

$$P_{a1} \triangleq \frac{\sum_{k=1}^{4} ||H_a^k * x||_1}{\sum_{k=1}^{4} ||H_a^k * \mathrm{DDSD}||_1}. \quad (13)$$

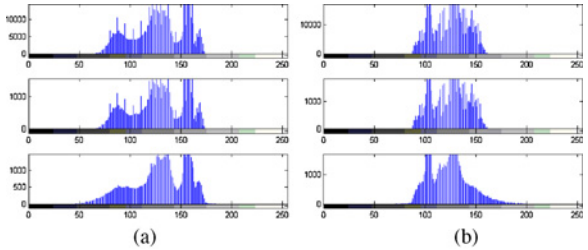


Fig. 9. Histogram distributions of chrominance components (U, V): (upper) original image, (middle) DPD, (lower) DDSD. (a) U. (b) V.

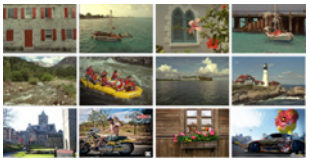


Fig. 10. Source images used in experiment, where images 1–8 are of size 768×512, 9–12 are of size 1024×768.

Similarly, we can use another 1-D high pass filter $b = [1 - 2\ 1]$ to implement the four filters H_b^k and define P_{b1} accordingly. By definition, these measures are positive numbers. A higher value indicates that there is more luminance HFE which suggests higher apparent luminance sharpness.

B. Measures for Chrominance Distortion

In our experiments, we observe that DPD typically does not cause chrominance distortion while DDSD can give significant distortion. In Fig. 9, we show the typical histograms of the chrominance (U, V) components of DPD and DDSD, and compare them with those of the original image. The U, V histograms of DPD are similar to those of the original image, while the U, V histograms of DDSD are observably different confirming the chrominance distortion.

According to Nyquist sampling theory, aliasing artifacts can occur whenever the bandwidth of signal is larger than half of the sampling frequency [17]. In DPD, aliasing can easily occur in Y component as the large image L typically contains high frequency details that exceeds the sampling frequency of S. On the other hand, aliasing in DPD is very minor for U and V components as the majority of chrominance details tends to concentrate in low frequency bands. Therefore, for any image x to be measured, we choose UV components of DPD as reference to compute $\mathrm{PSNR_U}$ and $\mathrm{PSNR_V}$

$$\mathrm{PSNR_U} \triangleq 10 \times log_{10}(255^2/\mathrm{MSE_U})$$
$$\mathrm{MSE_U} \triangleq \frac{1}{MN}(||\mathrm{U}_x - \mathrm{U_{DPD}}||_2^2) \quad (14)$$

where $\mathrm{U_{DPD}}$ and U_x are the U components of DPD and x, respectively. The $\mathrm{PSNR_V}$ is defined similarly.

IV. SIMULATION RESULTS

In this section, we simulate the proposed MMSE-SD and MMSE-SD(k) using two possible simple interpolation filters as shown in (15): the regular bilinear interpolation filter (denoted as H_{BI}), and a directional interpolation filter (denoted

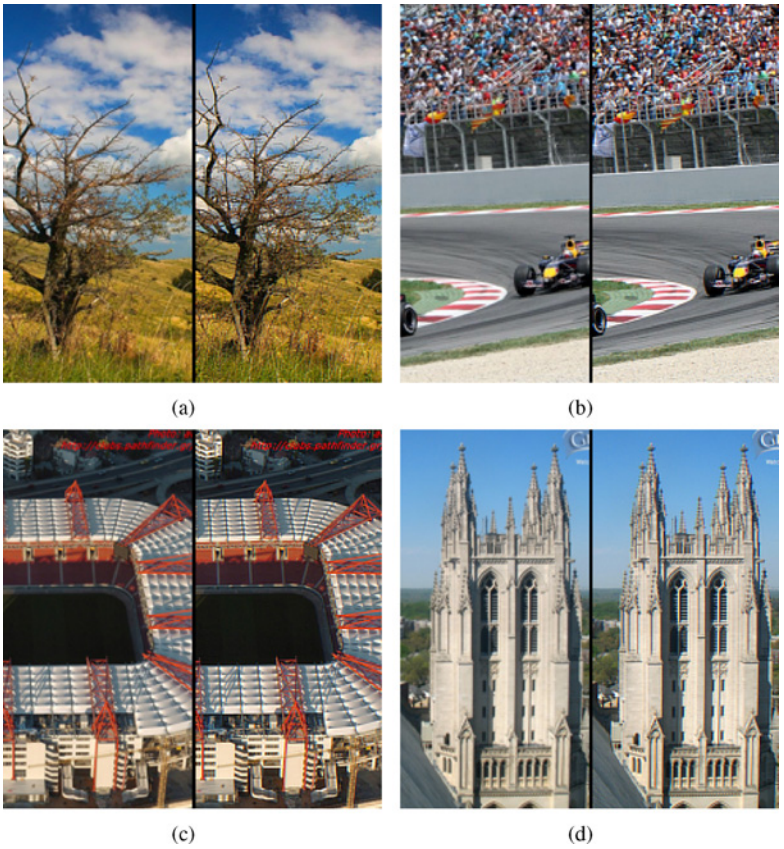


Fig. 11. (a)–(d) Several down-sampled images, where the left part uses PDAF, and the right part uses MMSE-$SD_{DI}(k = 9)$.

as H_{DI}) obtained by approximating the 22.5° contribution to the diagonal contribution, and treating the diagonal and antidiagonal interpolation in the same way as horizontal and vertical interpolation. The corresponding results are denoted as MMSE-SD_{BI}, MMSE-SD_{DI}, MMSE-$SD_{BI}(k)$ and MMSE-$SD_{DI}(k)$, respectively. As said earlier, we choose $k = 9$

$$H_{BI} = \begin{bmatrix} \frac{1}{9} & \frac{2}{9} & \frac{1}{3} & \frac{2}{9} & \frac{1}{9} \\ \frac{2}{9} & \frac{4}{9} & \frac{2}{3} & \frac{4}{9} & \frac{2}{9} \\ \frac{1}{3} & \frac{2}{3} & 1 & \frac{2}{3} & \frac{1}{3} \\ \frac{2}{9} & \frac{4}{9} & \frac{2}{3} & \frac{4}{9} & \frac{2}{9} \\ \frac{1}{9} & \frac{2}{9} & \frac{1}{3} & \frac{2}{9} & \frac{1}{9} \end{bmatrix} \quad H_{DI} = \begin{bmatrix} \frac{1}{3} & 0 & \frac{1}{3} & 0 & \frac{1}{3} \\ 0 & \frac{2}{3} & \frac{2}{3} & \frac{2}{3} & 0 \\ \frac{1}{3} & \frac{2}{3} & 1 & \frac{2}{3} & \frac{1}{3} \\ 0 & \frac{2}{3} & \frac{2}{3} & \frac{2}{3} & 0 \\ \frac{1}{3} & 0 & \frac{1}{3} & 0 & \frac{1}{3} \end{bmatrix}. \tag{15}$$

In particular, we examine their luminance apparent sharpness and chrominance distortion using the proposed measures in Section III. Twelve test images shown in Fig. 10 are used in experiment. MMSE-SD and MMSE-SD(k) are applied to these images to down-sample them by a factor of 3 both horizontally and vertically.

We compare MMSE-SD_{BI}, MMSE-SD_{DI}, MMSE-$SD_{BI}(k)$, and MMSE-$SD_{DI}(k)$ with PDAF which is the standard pixel-based down-sampling method. We also compare proposed methods with three subpixel-based down-sampling methods: Gibson [6], Kim [13], and Min-Max [12]. Unfortunately, we do not have enough details from [8] and [11] and, thus, cannot include them in our experiments.

Gibson [6] treats R, G, B subpixels as individual signal samples and applies a five-tap low-pass filter directly to smooth the result of subpixel rendering. As a result, the R, G, B values within any pixel tend to be very similar, if not identical, making the resulting image appear monochrome. Therefore, we expect the chrominance distortion of Gibson to be large for color images. Kim [13] only performed horizontal down-sampling without any vertical down-sampling, resulting in images with incorrect aspect ratios. Thus, we apply an additional pixel-based down-sampling in the vertical direction to Kim so that its results have the correct aspect ratio and can be compared with other methods in our experiments. In terms of Min-Max paper [12], there are two simplified methods: MMDE-DR and MMDE-ER. Basically MMDE-ER is much simpler than MMDE-DR with comparable performance. Thus we only compare with MMDE-ER.

To compare luminance sharpness of various methods, we show their P_{a1} and P_{b1} values in Tables III and IV. Since the special color fringing artifacts and blurring are best seen on LCD display, we only show some typical down-sampled images in Fig. 11, and we recommend the readers to view our results on our homepage directly: http://ihome.ust.hk/~fanglu/MMSE-SD2.htm.

While we are not sure whether filter a or filter b is better for measuring luminance sharpness, both filters give consistent results for our performance comparison. Tables III and IV show that the values P_{a1} and P_{b1} of Kim are significantly larger than PDAF and Gibson, while those of MMSE-SD and

TABLE III
OBJECTIVE MEASURE FOR LUMINANCE SHARPNESS USING FILTER a: P_{a1}

Image	PDAF	Gibson	Kim	MMDE-ER	MMSE-SD_{BI}	MMSE-$SD_{BI}(k)$	MMSE-SD_{DI}	MMSE-$SD_{DI}(k)$
1	0.777	0.701	0.903	1.065	0.980	1.045	1.067	1.113
2	0.779	0.729	0.916	1.031	0.934	0.992	1.008	1.046
3	0.834	0.749	0.933	1.009	0.938	1.000	1.002	1.043
4	0.848	0.791	0.975	1.041	0.950	1.019	1.019	1.062
5	0.841	0.743	0.990	1.048	0.968	1.052	1.054	1.107
6	0.791	0.723	0.883	1.072	0.973	1.042	1.062	1.108
7	0.926	0.865	1.023	1.154	1.065	1.157	1.153	1.211
8	0.857	0.762	1.019	1.092	1.009	1.082	1.091	1.138
9	0.892	0.812	1.020	1.103	1.034	1.106	1.112	1.157
10	0.884	0.823	0.989	1.090	1.024	1.102	1.106	1.154
11	0.851	0.764	0.991	1.136	1.064	1.129	1.152	1.197
12	0.932	0.900	1.007	1.122	1.061	1.129	1.130	1.174
Ave	0.851	0.780	0.971	1.080	1.000	1.071	1.080	1.126

TABLE IV
OBJECTIVE MEASURE FOR LUMINANCE SHARPNESS USING FILTER b: P_{b1}

Image	PDAF	Gibson	Kim	MMDE-ER	MMSE-SD_{BI}	MMSE-$SD_{BI}(k)$	MMSE-SD_{DI}	MMSE-$SD_{DI}(k)$
1	0.699	0.616	0.854	1.059	0.974	1.031	1.075	1.118
2	0.672	0.605	0.861	1.024	0.908	0.967	1.004	1.045
3	0.702	0.597	0.863	0.980	0.882	0.957	0.976	1.024
4	0.732	0.646	0.933	1.036	0.910	0.992	1.010	1.061
5	0.750	0.642	0.936	1.011	0.927	1.010	1.027	1.077
6	0.699	0.626	0.822	1.078	0.961	1.032	1.071	1.120
7	0.855	0.778	1.000	1.173	1.067	1.169	1.181	1.244
8	0.772	0.658	0.984	1.087	0.995	1.064	1.095	1.139
9	0.806	0.706	0.986	1.103	1.027	1.097	1.125	1.169
10	0.812	0.738	0.948	1.088	1.018	1.093	1.115	1.161
11	0.783	0.685	0.959	1.148	1.080	1.132	1.181	1.220
12	0.847	0.791	0.967	1.154	1.078	1.152	1.177	1.226
Ave	0.761	0.674	0.926	1.078	0.986	1.058	1.086	1.134

MMSE-SD(k) are in turn considerably, if not significantly, larger than Kim. Such higher values suggest that MMSE-SD and MMSE-SD(k) retain more high frequency details than PDAF, Gibson and Kim, as verified in http://ihome.ust.hk/~fanglu/MMSE-SD2.htm. In other words, our proposed MMSE-SD and MMSE-SD(k) tend to give higher apparent luminance resolution, leading to clearer and sharper down-sampled images. Comparing the two interpolation filters, the P_{a1} and P_{b1} of MMSE-SD_{BI} are lower than MMSE-SD_{DI} implying that the directional interpolation can give sharper images than bilinear interpolation.

By definition, the P_{a1} and P_{b1} of DDSD are always 1, and are similar to MMSE-SD and MMSE-SD(k). Recall that DDSD is well known to give very sharp images. Thus, Tables III and IV suggest that MMSE-SD and MMSE-SD(k) can give comparable sharpness as DDSD.

From Table V, we observe that the $PSNR_U$ and $PSNR_V$ of pixel-based PDAF is much higher than all other subpixel-based methods, including the proposed MMSE-SD and MMSE-SD(k), because pixel-based methods tend not to introduce any chrominance distortion. The large $PSNR_U$ and $PSNR_V$ confirm that the chrominance distortion of PDAF is comparable with DPD, our pixel-based reference method. The $PSNR_U$ and $PSNR_V$ of Kim are similar to those of MMSE-SD and MMSE-SD(k) suggesting similar chrominance distortion level for all three.

As expected, the $PSNR_U$ and $PSNR_V$ of Gibson are much worse than other methods, because it tends to remove all colors leading to huge chrominance distortion. We also observe that the $PSNR_U$ and $PSNR_V$ of DDSD are about $4-9$ dB lower than MMSE-SD and MMSE-SD(k) which suggests that MMSE-SD and MMSE-SD(k) can suppress chrominance distortion effectively while achieving similar level of apparent luminance sharpness. Basically, the proposed MMSE-SD achieves comparable if not higher P_a and P_b values, but 1–2 dB higher $PSNR_{U(V)}$ values than MMDE-ER. In terms of computational complexity, the proposed MMSE-SD(k = 9) is comparable to MMDE-ER.

Comparing the two interpolation filters, the $PSNR_U$ and $PSNR_V$ of MMSE-SD_{BI} are slightly larger than MMSE-SD_{DI} implying that the directional interpolation tends to introduce slightly more color fringing artifacts than bilinear interpolation. Considering the P_{a1}, P_{b1}, $PSNR_U$ and $PSNR_V$ of MMSE-SD and MMSE-SD(k), the 9×9 simplified filter appears to achieve good trade-off between apparent luminance sharpness and chrominance distortion, while achieving significant complexity reduction.

Fig. 12 shows cropped edge regions of various down-sampled images. When zoomed in, some minor color artifacts are visible in edge regions for subpixel-based down-sampling methods (i.e., DDSD and MMSE-SD), as shown in Fig. 12(b) and (d). Nevertheless, the color artifacts are less objectionable

TABLE V
OBJECTIVE MEASURE FOR CHROMINANCE DISTORTION: $PSNR_U$ AND $PSNR_V$

$PSNR_U$ (dB)									
Image	PDAF	DDSD	Gibson	Kim	MMDE-ER	MMSE-SD_{BI}	MMSE-$SD_{BI}(k)$	MMSE-SD_{DI}	MMSE-$SD_{DI}(k)$
1	51.82	27.04	26.55	32.92	31.64	34.07	33.34	33.17	32.70
2	50.17	34.50	20.25	39.53	38.20	40.59	39.94	39.80	39.36
3	49.33	35.86	20.88	39.53	38.67	40.53	39.71	39.72	39.18
4	53.07	33.88	26.72	37.19	36.30	38.27	37.53	37.46	37.02
5	46.54	27.09	25.76	30.82	30.14	32.04	31.21	31.17	30.66
6	48.70	28.90	24.76	37.01	34.66	37.94	37.19	37.03	36.55
7	50.90	31.34	24.74	35.24	34.98	36.36	35.57	35.54	35.04
8	47.16	23.58	27.60	27.95	27.76	29.20	28.50	28.28	27.85
9	51.53	27.86	25.29	31.66	31.38	32.78	32.10	31.94	31.51
10	42.63	25.45	24.41	29.58	29.57	30.64	29.95	29.85	29.43
11	44.18	28.11	24.41	32.21	32.26	33.46	32.84	32.56	32.14
12	52.70	30.90	24.94	36.84	36.41	37.83	37.01	37.03	36.51
Ave	49.06	29.54	24.69	34.21	33.50	35.31	34.57	34.46	34.00
$PSNR_V$ (dB)									
Image	PDAF	DDSD	Gibson	Kim	MMDE-ER	MMSE-SD_{BI}	MMSE-$SD_{BI}(k)$	MMSE-SD_{DI}	MMSE-$SD_{DI}(k)$
1	27.38	50.46	26.32	34.27	32.39	35.54	34.80	34.52	34.03
2	33.54	47.24	15.72	38.30	37.19	39.34	38.69	38.61	38.13
3	35.55	50.01	23.86	39.12	38.29	40.05	39.30	39.31	38.84
4	34.21	47.85	19.57	38.24	36.92	39.35	38.53	38.51	37.97
5	27.04	48.28	26.44	31.42	30.40	32.69	31.84	31.79	31.28
6	28.29	50.77	30.95	37.62	34.79	38.41	37.71	37.53	37.10
7	31.76	50.71	26.84	36.17	35.77	37.30	36.47	36.43	35.92
8	23.66	47.73	28.10	28.64	28.30	29.85	29.17	28.96	28.51
9	27.95	53.23	31.00	32.61	32.09	33.65	33.00	32.86	32.45
10	26.01	40.99	25.63	30.51	30.26	31.54	30.88	30.81	30.41
11	28.37	41.75	23.94	32.96	32.92	34.10	33.51	33.35	32.92
12	31.54	52.81	23.40	37.98	37.33	39.00	38.14	38.18	37.62
Ave	29.61	48.49	25.15	34.82	33.89	35.90	35.17	35.07	34.60

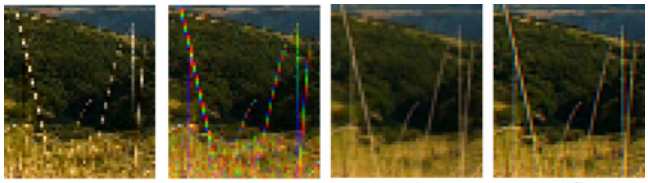


Fig. 12. Cropped edge region of various down-sampled images. (a) DPD. (b) DDSD. (c) PDAF. (d) MMSE-SD.

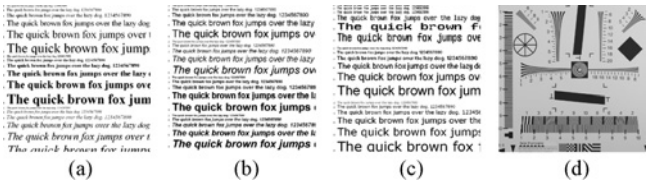


Fig. 13. Font images (13–15) and monochromatic image (16) used in experiment. (a) 13. (b) 14. (c) 15. (d) 16.

than the loss of details caused by not taking advantage of subpixel resolution (i.e., DPD and PDAF), as shown in Fig. 12(a) and (c). This is mainly because our human eyes are much more sensitive to the high frequency of luminance than to that of chrominance [18]. Fig. 12(d) depicts that MMSE-SD achieves higher apparent sharpness than those of DPD (i.e., line is broken) and PDAF (i.e., line is blurring), while it can effectively suppress annoying color artifacts in DDSD as well.

For the k:1 down-sampling ratio with the large image L of size $kM \times kN$ ($k > 3$), we use a common pixel-based decimation method (with bi-cubic filter) to resize L to be $3M \times 3N$ and then apply MMSE-SD. We simulate this method for down-sampling ratios of 4:1 and 5:1. Due to limited space, we only show P_{a1} and $PSNR_U$ values of several source images for the case 4:1 in Table VI, from which, similar conclusions can be obtained as 3:1 down-sampling.

Since Gibson's method is designed for monochromatic display types (fonts), we also include some font images and monochromatic images in our experiments, as shown in Fig. 13. The corresponding performance is given in Table VII. As expected, the $PSNR_U$ and $PSNR_V$ of Gibson's method are about 3–4 dB higher than Kim's method as well as the proposed MMSE-SD and MMSE-SD(k), which indicates that Gibson's method is more effective in suppressing color fringing artifacts for font images and monochromatic images. On the contrary, the values P_{a1} and P_{b1} of the proposed methods are larger than other methods, indicating that the proposed methods manage to provide sharper and clearer images with sacrifice of chrominance distortion. Since the special color fringing artifacts and blurring are best seen on LCD display, we recommend the readers to view our results on our homepage directly: http://ihome.ust.hk/~fanglu/MMSE-SD2.htm.

TABLE VI

LUMINANCE SHARPNESS AND CHROMINANCE DISTORTION FOR DOWN-SAMPLING FACTOR 4:1

P_{a1}								
Image	PDAF	Gibson	Kim	MMDE-ER	MMSE-SD$_{BI}$	MMSE-SD$_{BI}(k)$	MMSE-SD$_{DI}$	MMSE-SD$_{DI}(k)$
9	0.896	0.815	1.024	1.105	1.042	1.109	1.117	1.161
10	0.873	0.811	0.984	1.078	1.010	1.089	1.092	1.141
11	0.866	0.776	1.002	1.136	1.071	1.136	1.156	1.201
12	0.939	0.905	1.016	1.122	1.069	1.135	1.137	1.180

PSNR$_U$ (dB)									
Image	PDAF	DDSD	Gibson	Kim	MMDE-ER	MMSE-SD$_{BI}$	MMSE-SD$_{BI}(k)$	MMSE-SD$_{DI}$	MMSE-SD$_{DI}(k)$
9	49.92	27.92	25.29	31.24	31.06	32.36	31.65	31.54	31.10
10	41.42	24.67	24.41	29.27	29.02	30.37	29.63	29.57	29.12
11	43.38	27.84	24.43	32.50	32.44	33.55	33.04	32.80	32.43
12	51.02	30.37	24.90	35.90	35.50	36.89	36.08	36.09	35.58

TABLE VII

LUMINANCE SHARPNESS AND CHROMINANCE DISTORTION MEASURES FOR FONT AND MONOCHROMATIC IMAGES 13--16

P_{a1}							
Image	PDAF	Gibson	Kim	MMSE-SD$_{BI}$	MMSE-SD$_{BI}(k)$	MMSE-SD$_{DI}$	MMSE-SD$_{DI}(k)$
13	0.920	0.863	0.985	0.997	1.000	1.016	1.016
14	0.922	0.869	0.974	0.991	0.986	1.005	1.001
15	0.915	0.855	0.977	0.992	0.996	1.015	1.015
16	0.857	0.787	1.045	1.066	1.140	1.161	1.215
P_{b1}							
Image	PDAF	Gibson	Kim	MMSE-SD$_{BI}$	MMSE-SD$_{BI}(k)$	MMSE-SD$_{DI}$	MMSE-SD$_{DI}(k)$
13	0.893	0.817	0.959	0.978	0.975	0.996	0.992
14	0.876	0.813	0.944	0.968	0.958	0.982	0.975
15	0.872	0.795	0.947	0.974	0.968	0.995	0.989
16	0.784	0.699	1.003	1.042	1.108	1.140	1.186

PSNR$_V$ (dB)								
Image	PDAF	DDSD	Gibson	Kim	MMSE-SD$_{BI}$	MMSE-SD$_{BI}(k)$	MMSE-SD$_{DI}$	MMSE-SD$_{DI}(k)$
13	9999.00	18.18	25.43	22.34	23.07	22.70	22.56	22.34
14	9999.00	18.91	25.82	23.28	23.83	23.54	23.41	23.21
15	9999.00	19.18	26.25	23.29	23.97	23.62	23.46	23.26
16	55.08	26.18	32.77	28.45	29.57	29.02	28.70	28.32
PSNR$_V$ (dB)								
Image	PDAF	DDSD	Gibson	Kim	MMSE-SD$_{BI}$	MMSE-SD$_{BI}(k)$	MMSE-SD$_{DI}$	MMSE-SD$_{DI}(k)$
13	9999.00	18.69	26.56	23.34	24.12	23.72	23.59	23.35
14	9999.00	19.28	26.94	24.23	24.77	24.51	24.36	24.19
15	9999.00	19.53	27.38	24.28	24.97	24.62	24.45	24.25
16	55.52	27.07	34.16	29.55	30.69	30.13	29.81	29.42

V. CONCLUSION

In this paper, we proposed a new MMSE-SD method for displaying a high resolution image/video on a low-resolution LCD device. We showed analytically that MMSE-SD is equivalent to a 2-D spatial-invariant linear filter followed by DDSD, which is much faster than the straightforward implementation. The 2-D optimal filter of MMSE-SD can be approximated by a 9×9 filter to achieve lower complexity. Experimental results show that MMSE-SD and its approximated form can provide sharper and clearer images than the conventional pixel-based down-sampling method (PDAF) and is more effective in suppressing chrominance distortion than some subpixel-based down-sampling schemes.

APPENDIX A
MMSE-SD EQUATION AND ITS CHARACTERISTICS

Recall that the relationship between input R and output r is given by $\mathbf{H}_r\ r = \mathbf{H}_R\ R$, where $\mathbf{H}_r$ and $\mathbf{H}_R$ are block-circulant matrices of sizes $MN \times MN$ and $MN \times 9MN$, respectively

$$\mathbf{H}_r = \begin{bmatrix} \mathbf{A} & \mathbf{B}_2 & \mathbf{0} & \cdots & \mathbf{B}_1 \\ \mathbf{B}_1 & \mathbf{A} & \mathbf{B}_2 & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_1 & \mathbf{A} & \mathbf{B}_2 & \cdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \mathbf{B}_2 & \mathbf{0} & \cdots & \mathbf{B}_1 & \mathbf{A} \end{bmatrix}$$

$$\mathbf{H}_R = \begin{bmatrix} \mathbf{C} & \mathbf{D}_2 & \mathbf{E}_2 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{E}_1 & \mathbf{D}_1 \\ \mathbf{0} & \mathbf{E}_1 & \mathbf{D}_1 & \mathbf{C} & \mathbf{D}_2 & \mathbf{E}_2 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{E}_1 & \mathbf{D}_1 & \mathbf{C} & \mathbf{D}_2 & \mathbf{E}_2 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{E}_1 & \mathbf{D}_1 & \mathbf{C} & \mathbf{D}_2 & \mathbf{E}_2 \end{bmatrix}.$$

$\mathbf{H}_r$ contains $M \times M$ blocks (which are either $\mathbf{A}$, $\mathbf{B}_1$, $\mathbf{B}_2$ or $\mathbf{0}$), each being $N \times N$ circulant

$$\mathbf{A} = \begin{bmatrix} k_0 & k_1 & 0 & \cdots & k_1 \\ k_1 & k_0 & k_1 & 0 & \cdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & k_1 & k_0 & k_1 \\ k_1 & 0 & \cdots & k_1 & k_0 \end{bmatrix}$$

$$\mathbf{B}_1 = \begin{bmatrix} k_2 & k_4 & 0 & \cdots & k_3 \\ k_3 & k_2 & k_4 & 0 & \cdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & k_3 & k_2 & k_4 \\ k_4 & 0 & \cdots & k_3 & k_2 \end{bmatrix}$$

$$\mathbf{B}_2 = \begin{bmatrix} k_2 & k_3 & 0 & \cdots & k_4 \\ k_4 & k_2 & k_3 & 0 & \cdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & k_4 & k_2 & k_3 \\ k_3 & 0 & \cdots & k_4 & k_2 \end{bmatrix}.$$

$\mathbf{H}_R$ contains $M \times 3M$ blocks (which are either $\mathbf{C}$, $\mathbf{D}_1$, $\mathbf{D}_2$, $\mathbf{E}_1$, $\mathbf{E}_2$ or $\mathbf{0}$), each being a matrix of size $N \times 3N$

$$\mathbf{C} = \begin{bmatrix} 1 & \alpha_1 & \beta_1 & 0 & 0 & \cdots & \beta_1 & \alpha_1 \\ 0 & \beta_1 & \alpha_1 & 1 & \alpha_1 & \beta_1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \beta_1 & \alpha_1 & 1 & \alpha_1 & \beta_1 \end{bmatrix}$$

$$\mathbf{D}_1 = \begin{bmatrix} \alpha_2 & \alpha_4 & \gamma_1 & 0 & 0 & \cdots & \gamma_3 & \alpha_3 \\ 0 & \gamma_3 & \alpha_3 & \alpha_2 & \alpha_4 & \gamma_1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \gamma_3 & \alpha_3 & \alpha_2 & \alpha_4 & \gamma_1 \end{bmatrix}$$

$$\mathbf{D}_2 = \begin{bmatrix} \alpha_2 & \alpha_3 & \gamma_3 & 0 & 0 & \cdots & \gamma_1 & \alpha_4 \\ 0 & \gamma_1 & \alpha_4 & \alpha_2 & \alpha_3 & \gamma_3 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \gamma_1 & \alpha_4 & \alpha_2 & \alpha_3 & \gamma_3 \end{bmatrix}$$

$$\mathbf{E}_1 = \begin{bmatrix} \beta_2 & \gamma_4 & \beta_4 & 0 & \cdots & 0 & \beta_3 & \gamma_2 \\ 0 & \beta_3 & \gamma_2 & \beta_2 & \gamma_4 & \beta_4 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \beta_3 & \gamma_2 & \beta_2 & \gamma_4 & \beta_4 \end{bmatrix}$$

$$\mathbf{E}_2 = \begin{bmatrix} \beta_2 & \gamma_2 & \beta_3 & 0 & \cdots & 0 & \beta_4 & \gamma_4 \\ 0 & \beta_4 & \gamma_4 & \beta_2 & \gamma_2 & \beta_3 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \beta_4 & \gamma_4 & \beta_2 & \gamma_2 & \beta_3 \end{bmatrix}.$$

Consider three elements in the first row of $\mathbf{C}$ as a block (e.g., $[1\ \alpha_1\ \beta_1]$), such a block is repeatedly shifted to the right by three positions in the subsequent rows. Thus, each of $\mathbf{C}$, $\mathbf{D}_1$, $\mathbf{D}_2$, $\mathbf{E}_1$, $\mathbf{E}_2$ or $\mathbf{0}$ is block-circulant. Similarly, considering three sub-blocks of $\mathbf{H}_R$ in horizontal direction (e.g., $[\mathbf{C}\ \mathbf{D}_2\ \mathbf{E}_2]$) as a block, which appears in the first row of $\mathbf{H}_R$ and is repeatedly shifted to the right by three block positions in the subsequent rows, $\mathbf{H}_R$ is a block-circulant matrix.

Here, we call a matrix of size $N \times 9N$ "block-tri-circulant," if it contains three sub-blocks of size $N \times 3N$, each of which is block-circulant (e.g., $[\mathbf{C}\ \mathbf{D}_2\ \mathbf{E}_2]$ is block-tri-circulant). Therefore, $\mathbf{H}_R$ is block circulant, with blocks of size $N \times 9N$ that are themselves block-tri-circulant.

APPENDIX B

Recall that the MMSE-SD solution is $r = (\mathbf{H}_r^{-1}\mathbf{H}_R)\ R$. Let $\mathbf{H} \triangleq \mathbf{H}_r^{-1}\mathbf{H}_R$, in this section, we will prove that $\mathbf{H}$ is a block-circulant matrix, with blocks of size $N \times 9N$ that are themselves block-tri-circulant. We will give the proofs for a class of matrices $\mathbf{H'}_r$ and $\mathbf{H'}_R$, of which $\mathbf{H}_r$ and $\mathbf{H}_R$ are special cases.

Lemma 1: Let $\mathbf{H'}_R \in C^{MN \times 9MN}$ be the block-circulant matrix

$$\mathbf{H'}_R \triangleq \begin{bmatrix} \hat{\mathbf{H}}_1 & \hat{\mathbf{H}}_2 & \cdots & \hat{\mathbf{H}}_M \\ \hat{\mathbf{H}}_M & \hat{\mathbf{H}}_1 & \cdots & \hat{\mathbf{H}}_{M-1} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\mathbf{H}}_2 & \hat{\mathbf{H}}_3 & \cdots & \hat{\mathbf{H}}_1 \end{bmatrix} \tag{16}$$

with $M \times M$ blocks, each of which is of size $N \times 9N$ and is block-tri-circulant. Then, there exist two Hermitian matrices that factorize $\mathbf{H'}_R$ into a block-diagonal matrix of which all blocks are block-tri-circulant.

Proof: As $\mathbf{H}'_R$ is block-circulant, following Lemma 1 in [19], $\mathbf{H}'_R$ satisfies the following factorization equality:

$$\Phi = \mathbf{U}_1^H \mathbf{H'}_R \mathbf{U}_2 = diag(\Phi_1, \Phi_2, \ldots, \Phi_M) \tag{17}$$

where Φ is block-diagonal and $\Phi_i \in C^{N \times 9N}$. $\mathbf{U}_1 \in C^{MN \times MN}$ and $\mathbf{U}_2 \in C^{9MN \times 9MN}$ are Hermitian and

$$\begin{aligned} \mathbf{U}_1 &\triangleq \begin{bmatrix} u_{11}\mathbf{I}_N & u_{12}\mathbf{I}_N & \cdots & u_{1M}\mathbf{I}_N \\ u_{21}\mathbf{I}_N & u_{22}\mathbf{I}_N & \cdots & u_{2M}\mathbf{I}_N \\ \vdots & \vdots & \ddots & \vdots \\ u_{M1}\mathbf{I}_N & u_{M2}\mathbf{I}_N & \cdots & u_{MM}\mathbf{I}_N \end{bmatrix} \\ \mathbf{U}_2 &\triangleq \begin{bmatrix} u_{11}\mathbf{I}_{9N} & u_{12}\mathbf{I}_{9N} & \cdots & u_{1M}\mathbf{I}_{9N} \\ u_{21}\mathbf{I}_{9N} & u_{22}\mathbf{I}_{9N} & \cdots & u_{2M}\mathbf{I}_{9N} \\ \vdots & \vdots & \ddots & \vdots \\ u_{M1}\mathbf{I}_{9N} & u_{M2}\mathbf{I}_{9N} & \cdots & u_{MM}\mathbf{I}_{9N} \end{bmatrix} \end{aligned} \tag{18}$$

where $\mathbf{I}_N$ and $\mathbf{I}_{9N}$ are identity matrices of size $N \times N$ and $9N \times 9N$, respectively ■

$$u_{pq} = \frac{1}{\sqrt{M}} w^{(p-1)(q-1)} \qquad p = 1, \ldots, M \quad q = 1, \ldots, M \tag{19}$$

where $w = e^{-i2\pi/M}$ is the Mth root of unity.

Meanwhile [20]

$$\Phi_k = \hat{\mathbf{H}}_1 + \hat{\mathbf{H}}_2 w^{(k-1)} + \hat{\mathbf{H}}_3 (w^{(k-1)})^2 + \cdots + \hat{\mathbf{H}}_M (w^{(k-1)})^{M-1}. \tag{20}$$

Since each $\hat{\mathbf{H}}_k$ is block-tri-circulant, Φ_k is also block-tri-circulant because scalar multiplication and summation do not change the block-tri-circulant structure. Therefore, Φ is block-diagonal where each block is block-tri-circulant. This proves Lemma 1.

Let $\mathbf{H}'_r \in C^{MN \times MN}$ be a block-circulant matrix

$$\mathbf{H'}_r \triangleq \begin{bmatrix} \mathbf{H}_1 & \mathbf{H}_2 & \cdots & \mathbf{H}_M \\ \mathbf{H}_M & \mathbf{H}_1 & \cdots & \mathbf{H}_{M-1} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{H}_2 & \mathbf{H}_3 & \cdots & \mathbf{H}_1 \end{bmatrix} \tag{21}$$

with $M \times M$ blocks, each of which is of size $N \times N$ and is circulant. Following Lemma B1, we can easily get $\Lambda = \mathbf{U}_1^H \mathbf{H}'_r \mathbf{U}_1 = diag(\Lambda_1, \Lambda_2, \ldots, \Lambda_M)$, where $\Lambda_k \in C^{N\times N}$ is circulant.

Lemma 2: Let $\Lambda = diag(\Lambda_1, \Lambda_2, \ldots, \Lambda_M)$, where $\Lambda_k \in C^{N\times N}$ is circulant. Then Λ^{-1} is also block-diagonal containing $M \times M$ blocks that themselves are circulant.

Proof: As $\Lambda = diag(\Lambda_1, \Lambda_2, \ldots, \Lambda_M)$, $\Lambda^{-1} = diag(\Lambda_1^{-1}, \Lambda_2^{-1}, \ldots, \Lambda_M^{-1})$. It is shown in [21] for any $N \times N$ circulant matrix $\Lambda_k, k = 1, \ldots, M$, the product $\mathbf{P}\Lambda_k\mathbf{P}^H$ is diagonal, where $\mathbf{P}$ is the $N \times N$ discrete Fourier transform matrix. Let $\Delta = \mathbf{P}^H\Lambda_k\mathbf{P}$. Then Λ_k is given by

$$\Lambda_k = \mathbf{P}\Delta\mathbf{P}^H. \tag{22}$$

As $\mathbf{P}$ is Hermitian ($\mathbf{P}^H = \mathbf{P}^{-1}$), the inverse of Λ_k is

$$\Lambda_k^{-1} = (\mathbf{P}\Delta\mathbf{P}^H)^{-1} = (\mathbf{P}^H)^{-1}\Delta^{-1}\mathbf{P}^{-1} = \mathbf{P}\Delta^{-1}\mathbf{P}^H. \tag{23}$$

Since Δ is diagonal, Δ^{-1} is also diagonal. Following from the results in [21] that for any diagonal matrix Δ^{-1}, the product $\mathbf{P}\Delta^{-1}\mathbf{P}^H$ is circulant, we conclude that Λ_k^{-1} is circulant. Therefore, Λ^{-1} is block-diagonal where each block is circulant. This proves Lemma 2. ■

Let $\Psi \triangleq \Lambda^{-1}\Phi \in C^{MN\times 9MN}$. Note that both $\Lambda^{-1} \in C^{MN\times MN}$ and $\Phi \in C^{MN\times 9MN}$ are block-diagonal. Thus Ψ is also block-diagonal that $\Psi = diag(\Psi_1, \Psi_2, \ldots, \Psi_M)$ where $\Psi_k = \Lambda_k^{-1}\Phi_k \in C^{N\times 9N}$. As Λ_k^{-1} is circulant and Φ_k is block-tri-circulant, after some algebra operations, it can be shown that $\Psi_k = \Lambda_k^{-1}\Phi_k \in C^{N\times 9N}$ is also block-tri-circulant.

Lemma 3: Let $\Psi = diag(\Psi_1, \Psi_2, \ldots, \Psi_{\mathbf{M}})$, where $\Psi_k \in C^{N\times 9N}$ is block-tri-circulant. Then $\mathbf{U}_1\Psi\mathbf{U}_2^H$ is block-circulant containing blocks of size $N \times 9N$ that are block-tri-circulant.

Proof: Let $\mathbf{J} \triangleq \mathbf{U}_1\Psi\mathbf{U}_2^H$. For $\mathbf{U}_1$ and $\mathbf{U}_2$ being the Hermitian matrices as in (18), and for any block-diagonal matrix Ψ, it is shown in [22] that $\mathbf{U}_1\Psi\mathbf{U}_2^H$ is block-circulant. ■

Let $\mathbf{T} \triangleq \mathbf{U}_1\Psi \in C^{MN\times 9MN}$ such that $\mathbf{J} = \mathbf{T}\mathbf{U}_2^H$ and

$$\mathbf{T} \triangleq \begin{bmatrix} \mathbf{T}_{11} & \mathbf{T}_{12} & \ldots & \mathbf{T}_{1M} \\ \mathbf{T}_{21} & \mathbf{T}_{22} & \ldots & \mathbf{T}_{2M} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{T}_{M1} & \mathbf{T}_{M2} & \ldots & \mathbf{T}_{MM} \end{bmatrix}$$

$$= \begin{bmatrix} u_{11}\mathbf{I}_N & u_{12}\mathbf{I}_N & \cdots & u_{1M}\mathbf{I}_N \\ u_{21}\mathbf{I}_N & u_{22}\mathbf{I}_N & \cdots & u_{2M}\mathbf{I}_N \\ \vdots & \vdots & \ddots & \vdots \\ u_{M1}\mathbf{I}_N & u_{M2}\mathbf{I}_N & \cdots & u_{MM}\mathbf{I}_N \end{bmatrix} \tag{24}$$

$$\times \begin{bmatrix} \Psi_1 & 0 & \cdots & 0 \\ 0 & \Psi_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Psi_M \end{bmatrix}$$

where $\mathbf{T}_{pq} = u_{pq}\Psi_q$, $p = 1, \ldots, M$, $q = 1, \ldots, M$. Recall

$$\mathbf{J} = \mathbf{T}\mathbf{U}_2^H$$
$$= \begin{bmatrix} u_{11}\Psi_1 & u_{12}\Psi_2 & \cdots & u_{1M}\Psi_M \\ u_{21}\Psi_1 & u_{22}\Psi_2 & \cdots & u_{2M}\Psi_M \\ \vdots & \vdots & \ddots & \vdots \\ u_{M1}\Psi_1 & u_{M2}\Psi_2 & \cdots & u_{MM}\Psi_M \end{bmatrix} \tag{25}$$
$$\times \begin{bmatrix} u_{11}^*\mathbf{I}_{9N} & u_{21}^*\mathbf{I}_{9N} & \cdots & u_{M1}^*\mathbf{I}_{9N} \\ u_{12}^*\mathbf{I}_{9N} & u_{22}^*\mathbf{I}_{9N} & \cdots & u_{M2}^*\mathbf{I}_{9N} \\ \vdots & \vdots & \ddots & \vdots \\ u_{1M}^*\mathbf{I}_{9N} & u_{2M}^*\mathbf{I}_{9N} & \cdots & u_{MM}^*\mathbf{I}_{9N} \end{bmatrix}.$$

It is easy to obtain that

$$\begin{aligned} \mathbf{J}_{pq} &= [u_{p1}\Psi_1 \ \cdots \ u_{pq}\Psi_q \ \cdots \ u_{pM}\Psi_M] \\ &\times [u_{q1}^*\mathbf{I}_{9N} \ \cdots \ u_{qq}^*\mathbf{I}_{9N} \ \cdots \ u_{qM}^*\mathbf{I}_{9N}]^T \\ &= u_{p1}u_{q1}^*\Psi_1 + \cdots + u_{pq}u_{qq}^*\Psi_q + \cdots + u_{pM}u_{qM}^*\Psi_M. \end{aligned} \tag{26}$$

Since each Ψ_k is block-tri-circulant, $\mathbf{J}_{pq}$ is also block-tri-circulant because scalar multiplication and addition do not change the block-tri-circulant structure. Therefore, $\mathbf{J} = \mathbf{U}_1\Psi\mathbf{U}_2^H$ is block-circulant where $\mathbf{J}_{pq}$ is block-tri-circulant. This proves Lemma 3.

As $\mathbf{H}_r$ and $\mathbf{H}_R$ are special cases of $\mathbf{H}'_r$ and $\mathbf{H}'_R$, all of the above lemmas and conclusions apply to $\mathbf{H}_r$ and $\mathbf{H}_R$. As such, we give a summary of how the MMSE-SD equation from (9) leads to the special structure of $\mathbf{H}$

$$\begin{aligned} \mathbf{H}_r r = \mathbf{H}_R R &\Rightarrow \mathbf{H}_r(\mathbf{U}_1\mathbf{U}_1^H)r = \mathbf{H}_R(\mathbf{U}_2\mathbf{U}_2^H)R \\ &\Rightarrow (\mathbf{U}_1^H\mathbf{H}_r\mathbf{U}_1)\mathbf{U}_1^H r = (\mathbf{U}_1^H\ \mathbf{H}_R\mathbf{U}_2)U_2^H R \\ &\Rightarrow \Lambda\mathbf{U}_1^H r = \Phi\mathbf{U}_2^H R \\ &\Rightarrow \mathbf{U}_1^H r = (\Lambda^{-1}\Phi)\mathbf{U}_2^H R \\ &\Rightarrow \mathbf{U}_1^H r = \Psi\mathbf{U}_2^H R \\ &\Rightarrow r = \mathbf{U}_1\Psi\mathbf{U}_2^H R. \end{aligned}$$

Recall that $r = \mathbf{H}R$. Then Lemma 3 shows that $\mathbf{H} = \mathbf{U}_1\Psi\mathbf{U}_2^H$ is block-circulant containing blocks of size $N \times 9N$ that are block-tri-circulant.

REFERENCES

[1] Y. Amano, "A flat-panel TV display system in monochrome and color," *IEEE Trans. Electron Devices*, vol. 22, no. 1, pp. 1–7, Jan. 1975.

[2] T. Benzschawel and W. Howard, "Method of and apparatus for displaying a multicolor image," U.S. Patent 5 341 153, Aug. 23, 1994.

[3] C. H. B. Elliott, S. Han, M. H. Im, M. Higgins, and P. Higgins, "Co-optimization of color AMLCD subpixel architecture and rendering algorithms," *Soc. Inform. Display Dig.*, vol. 33, no. 1, pp. 172–175, May 2002.

[4] L. M. Chen and S. Hasegawa, "Influence of pixel-structure noise on image resolution and color for matrix display devices," *J. Soc. Inform. Display*, vol. 1, no. 1, pp. 103–110, Jan. 1993.

[5] M. A. Klompenhouwer and G. de Haan, "Subpixel image scaling for color matrix displays," *SID Dig.*, vol. 33, no. 1, pp. 176–179, May 2002.

[6] S. Gibson. (2003). *Sub-Pixel Font Rendering Technology* [Online]. Available: http://www.grc.com/cleartype.htm

[7] J. C. Platt, B. Keely, B. Hill, B. Dresevic, C. Betrisey, D. P. Mitchell, G. Hitchcock, J. F. Blinn, and T. Whitted, "Displaced filtering for patterned displays," in *Proc. SID Symp. Dig. Tech. Papers*, vol. 31. 2000, pp. 296–299.

[8] J. C. Platt, "Optimal filtering for patterned displays," *IEEE Signal Process. Lett.*, vol. 7, no. 7, pp. 179–181, Jul. 2000.

[9] S. J. Daly, "Analysis of subtriad addressing algorithms by visual system models," *SID Dig.*, vol. 32, no. 1, pp. 1200–1203, 2001.

[10] S. J. Daly, "Methods and systems for improving display resolution in images using sub-pixel sampling and visual error filtering," U.S. Patent 6 608 632 B2, Aug. 2003.
[11] D. S. Messing and S. J. Daly, "Improved display resolution of subsampled color images using subpixel addressing," in *Proc. IEEE Int. Conf. Image Process.*, vol. 1. Sep. 2002, pp. 22–25.
[12] L. Fang and O. C. Au, "Subpixel-based image down-sampling with min-max directional error for stripe display," *IEEE J. Sel. Top. Signal Process.*, vol. 5, no. 2, pp. 240–251, Apr. 2011.
[13] J. S. Kim and C. S. Kim, "A filter design algorithm for subpixel rendering on matrix displays," in *Proc. Eur. Signal Process. Conf.*, Sep. 2007, pp. 1487–1491.
[14] J. H. Kranz and L. D. Silverstein, "Color matrix display image quality: The effects of luminance and spatial sampling," *SID Dig.*, pp. 29–32, May 1990.
[15] L. D. Silverstein, R. W. Monty, F. E. Gomer, and Y. Y. Yeh, "A psychophysical evaluation of pixel mosaics and gray-scale requirements for color matrix displays," *SID Dig.*, vol. 20, pp. 128–131, May 1989.
[16] ITU, *Studio Encoding Parameters of Digital Television for Standard 4:3 and Wide-Screen 16:9 Aspect Ratios*, ITU-R Rec. BT.601-5, 1995.
[17] P. S. R. Diniz, E. A. B. D. Silva, and S. L. Netto, *Digital Signal Processing: System Analysis and Design*. Cambridge, MA: Cambridge Univ. Press, 2002.
[18] B. A. Wandell, *Foundations of Vision*. Sunderland, MA: Sinauer Associates, 1995.
[19] Y. Tian, "Some equalities for generalized inverses of matrix sums and block circulant matrices," *Arch. Math.*, vol. 37, no. 4, pp. 301–306, 2001.
[20] G. J. Tee, "Eigenvectors of block circulant and alternating circulant matrices," *Res. Lett. Inf. Math. Sci.*, vol. 8, no. 6, pp. 123–142, 2005.
[21] R. M. Gray, "Toeplitz and circulant matrices: A review," *Found. Trends Commun. Inform. Theory*, vol. 2, no. 3, pp. 155–239, 2006.
[22] T. De Mazancourt and D. Gerlic, "The inverse of a block-circulant matrix," *IEEE Trans. Antennas Propag.*, vol. 31, no. 5, pp. 808–810, Sep. 1983.

Lu Fang received the B.S. degree from the Department of Electronic Engineering and Information Science, University of Science and Technology of China, Hefei, China, in June 2007. She used to visit Northwestern University, Evanston, IL, under the support of Prof. A. K. Katsaggelos. She received the Ph.D. degree from the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong, in July 2011.

She is currently with the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology.

Oscar C. Au (SM'XX) received the B.A.Sc. degree from the University of Toronto, Toronto, ON, Canada, in 1986, and the M.A. and Ph.D. degrees from Princeton University, Princeton, NJ, in 1988 and 1991, respectively.

After being a Post-Doctoral Researcher with Princeton University for one year, he joined the Hong Kong University of Science and Technology (HKUST), Clear Water Bay, Hong Kong, as an Assistant Professor in 1992. He is/has been a Professor with the Department of Electronic and Computer Engineering, the Director of the Multimedia Technology Research Center, and the Director of the Computer Engineering Program in HKUST. He has published about 320 technical journals and conference papers. His fast motion estimation algorithms were accepted into the ISO/IEC 14496-7 MPEG-4 International Video Coding Standard and the China AVS-M Standard. His light-weight encryption and error resilience algorithms are accepted into the China AVS Standard. He has eight U.S. patents and is applying for more than 60 on his signal processing techniques. He has performed forensic investigation and stood as an Expert Witness in the Hong Kong courts many times. His main research interests include video and image coding and processing, watermarking and light weight encryption, speech, and audio processing. Research topics include fast motion estimation for MPEG-1/2/4, H.261/3/4 and AVS, optimal and fast suboptimal rate control, mode decision, transcoding, denoising, deinterlacing, postprocessing, multiview coding, scalable video coding, distributed video coding, subpixel rendering, JPEG/JPEG2000, HDR imaging, compressive sensing, halftone image data hiding, GPU-processing, software–hardware codesign, and so on.

Dr. Au is a Board of Governor Member of the Asia Pacific Signal and Information Processing Association. He is/was an Associate Editor of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, the IEEE TRANSACTIONS ON IMAGE PROCESSING, and the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS, PART 1. He is on the editorial boards of the *Journal of Signal Processing Systems*, the *Journal of Multimedia*, and the *Journal of Franklin Institute*. He is/was the Chair of the CAS Technical Committee on Multimedia Systems and Applications, the Vice Chair of the SP TC on Multimedia Signal Processing, and a member of the CAS TC on Video Signal Processing and Communications, the CAS TC on DSP, and the SP TC on Image, Video and Multidimensional Signal Processing. He has served on the steering committees of the IEEE TRANSACTIONS ON MULTIMEDIA and the IEEE International Conference of Multimedia and Expo. He also served on the organizing committees of the IEEE International Symposium on Circuits and Systems in 1997, the IEEE International Conference On Acoustics, Speech and Signal Processing in 2003, the ISO/IEC MPEG 71st Meeting in 2005, the International Conference on Image Processing in 2010, and other conferences. He was the General Chair of the Pacific-Rim Conference on Multimedia in 2007, the IEEE ICME in 2010, and the Packet Video Workshop in 2010. He received Best Paper Awards in SiPS 2007 and PCM 2007. He was an IEEE Distinguished Lecturer in 2009 and 2010, and has been a Keynote Speaker for a few times.

Ketan Tang received the B.S. degree from the Department of Electronic Engineering and Information Science, University of Science and Technology of China, Hefei, China, in 2009. He is currently pursuing the Ph.D. degree with the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong.

Xing Wen received the B.E. degree in electronic engineering from Xidian University, Xi'an, China, in 2007. He is currently pursuing the Ph.D. degree with the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong, under the supervision of Prof. O. C. Au and Prof. J. Xu.

His current research interests include video coding standard, motion estimation, HW/SW codesign, hardware implementation of multimedia algorithms, high throughput very large scale integration design, multiple description video coding, and visual quality assessment.

Hanli Wang received the B.S. and M.S. degrees in electrical engineering from Zhejiang University, Hangzhou, China, in 2001 and 2004, respectively, and the Ph.D. degree in computer science from the City University of Hong Kong, Kowloon, Hong Kong, in 2007.

From 2007 to 2008, he was a Research Fellow with the Department of Computer Science, City University of Hong Kong. From 2007 to 2008, he was a Visiting Scholar with Stanford University, Palo Alto, CA, invited by Prof. C. K. Chui. From 2008 to 2009, he was a Research Engineer with Precoad, Inc., Menlo Park, CA. From 2009 to 2010, he was an Alexander von Humboldt Research Fellow with the University of Hagen, Hagen, Germany. In 2010, he joined the Department of Computer Science and Technology, Tongji University, Shanghai, China, as a Professor. His current research interests include digital video coding, image processing, pattern recognition, and video content analysis.