Paper

Preserved Color Pixel: high-resolution and high-color-fidelity image acquisition using single image sensor with sub-half-micron pixels

Yuichiro Yamashita† (member), Rihito Kuroda†† (member) and Shigetoshi Sugawa†† (fellow)

Abstract A preserved-color-pixel (PCP) concept is proposed. The PCP color filter array (CFA) is arranged to construct "PCP pixels". A PCP pixel is surrounded by "buffer pixels" having color filters of the same color spectrum as that of the PCP pixel, so that most of color cross-talk from pixels of different colors are absorbed by the buffer pixels. The color cross-talk components of the buffer-pixel signals are computationally canceled by a proposed non-parametric method called "similarity-based blind cross-talk correction (SBC)," where signals of PCP pixels are used as the ground truth to estimate the signals of buffer-pixels without influence of the cross-talk. The demosaicing of each color planes' images sampled with a PCP-CFA arrangement is implemented by the adaptive normalized convolution (ANC) in conjunction with the proposed "post-convolutional-variation-minimization (PCVM)" algorithm for its cost function. Both SBC and PCVM-ANC are especially useful for image acquisition with a pixel array in a sub-half-micron generation, where its pixel pitch is approximately, or smaller than, 0.5 μm. The concept is verified with image simulation, and its effectiveness is quantified with the slanted-edge based spatial frequency response (SFR) modular transfer function (MTF) method by using the parametric color cross-talk analysis based on proposed "scalable-single-parameter (SSP)" color cross-talk model. The image simulation confirms the color reproducibility, together with the effectiveness of image resolution improvement under the influence of the complication of color cross-talk between pixels and lateral chromatic aberration (LCA) of the taking-lens. The benefit is also verified by peak-signal-to-noise-ratio (PSNR) analysis with simulated images based on a real-world picture, indicating that the proposed concept can maintain PSNR when color cross-talk increases.

Keywords: preserved-color-pixel, color filter array, color cross-talk correction, normalized convolution, image resolution, color fidelity, sub-half-micron pixel.

1. Introduction

The small camera module has been widely implemented in mobile devices such as the smartphone, and its image quality has been consistently improving owing to the progress of the technology and design of small camera module and image sensor device, combined with an advanced image post-processing algorithm supported by low-power and high-performance computing capabilities. Image resolution and color fidelity are part of crucial indices describing the image quality of camera modules employing a single image sensor with a color filter array (CFA), and the sampling frequency of pixel array has been increased by shrinking the pixel pitch while improving the intrinsic pixel performance. The significance of the conventional approach remains unchanged; whereas there will be emerging challenges as the pixel pitch shrinks, given the conditions that both the size of the camera module for the mobile device and the wavelength range of visible light are kept constant. It will be more difficult to confine electromagnetic energy of light by the micro-lens and wave-guide, and its leakage to the adjacent pixels, i.e., a cross-talk, comes to be more evident. The improvement of camera image quality with a single sensor has a possibility to hit a plateau; it is therefore expected to support the continuing improvement trend with an additional, such as computational, approach. With regard to the cross-talk correction, so far, the existing algorithms either assumes a cross-talk...
function\(^1\) or estimates it using a known target\(^2\) and apply an inverse filter using the function. The former approach may have a large estimation error if there is a gap between the assumption on the cross-talk and the actual one, whereas the latter requires time-consuming calibration step in a factory before the shipment of a camera. It is practically difficult to apply them for mainstream consumer use.

This article proposes a concept named "preserved color pixel (PCP)\(^3\)", enabling a non-parametric cross-talk correction algorithm free from calibration in a factory. The concept consists of a special CFA arrangement: PCP-CFA, as well as a set of supporting image processing algorithms to correct the color cross-talk in non-parametric approach and to demosaic the data captured by the PCP-CFA arrangement so that the image resolution and color fidelity are effectively enhanced for a camera with a single image sensor. Here, the color cross-talk in this article corresponds specifically to the transmission of the electromagnetic energy from another color planes induced by "spectral crosstalk" and "optical spatial crosstalk\(^3\)" unless otherwise stated. In Sections 2 and 3, as a background to this study, the key spatial property of the sub-half-micron pixel, together with effect from the complication of the color cross-talk and a lateral chromatic aberration (LCA) of a taking lens to image quality will be briefly described. In Section 4, the PCP concept, alongside the proposed cross-talk correction algorithm: "similarity-based blind cross-talk correction (SBC)" and the interpolation algorithm: "adaptive normalized convolution (ANC)" with "post-convolutional-variation-minimization (PCVM)," will also be explained. In Section 5, a proposed "scalable single parameter (SSP)" cross-talk model will be explained in order to enable parametric analysis of the image resolution and color fidelity with various levels of cross-talks. In Sections 6 and 7, the testing setup and simulation flow will be described, followed by the results. Section 8 will conclude the study with discussions about future directions.

2. Property of Images Captured with a Sub-Half-Micron Pixel Array

A small camera module for the mobile phone has been composed of an image sensor with small pixel pitch, e.g., 1.5–0.8 µm as of the year 2019, to enhance the image resolution under a limitation on possible optical format. To compensate for the small sensitivity of pixel, a fast taking-lens with small F-number of, e.g., 1.5–2.4 has been employed. When considering a combination of the pixel pitch of 1.0 µm and F-number of 1.5 as an example, the MTF of the pixel and the lens show comparable first-zero-frequencies of 1000 lp/mm and 1234 lp/mm in their frequency response characteristics curve, respectively, where the wavelength of 540nm is used to calculate the values based on the diffraction MTF theory\(^4\). Assuming there is no aliasing, both the lens and pixel determine the system frequency by almost equal weights. If the sub-half-micron pixel of, e.g., 0.5 µm is used, the first zero of the pixel MTF increases to 2000 lp/mm. Consequently, in the sub-half-micron pixel generation, the taking-lens determines the frequency of the system, and the intensity difference of two adjacent pixels turns out to be small. Figures 1, showing the correlation plot of intensities of adjacent pixels (Center to top left, and to right).

3. The Combined Effect of Color Cross-Talk and LCA in a Single Image Sensor Camera

The types of cross-talk were categorized into spectral cross-talk, optical spatial cross-talk, and electrical cross-talk\(^3\). The spectral cross-talk is an intra-pixel color cross-talk due to the color filter property. The electrical cross-talk occurs between photodiodes or subsequent signal readout blocks, which can be minimized by introducing Deep Trench Isolation\(^5\) in the photodiode and by eliminating signal-coupling paths in the circuit. The optical spatial cross-talk of sub-half-micron pixels in the same color plane does not noticeably reduce the
system frequency response since the taking-lens limits the system frequency, as explained in Section 2. However, the spectral and optical spatial cross-talk occurring between different color planes, i.e., the color cross-talk, causes severe degradation of the frequency response when it is combined with LCA. The LCA causes a misalignment in color channels, followed by spatial averaging by the cross-talk. LCA correction by geometric translation cannot recover the effect of averaging, and the spatial frequency response is reduced. The degradation is expected to be severer for the case of smaller camera module with sub-half-micron pixel, since the color cross-talk will be more evident as the pixel pitch shrinks, and the LCA of the camera module lens, originating from material dispersion and lens design, is difficult to be suppressed under the limitation set by the cost and the form-factor. Although the use of sub-half micron pixel for the small camera module originally aims at the enhancement of image resolution, the inverse consequence under the LCA and color cross-talk needs to be resolved to gain the benefit from smaller pixels.

4. PCP: Preserved Color Pixel Concept

4.1 Color filter array arrangement of preserved color pixel concept

Figure 2 shows the top view of the Bayer PCP-CFA arrangement with one line of buffer pixels for R, G, and B color planes. The PCP concept is based on the color filter arrangement to reserve at least one "preserved" color pixel (PCP pixel) surrounded by buffer pixels having filters of the same color as that of the PCP pixel in order to construct a small array of pixels, i.e., the patch. In this example, each patch consists of a 3-by-3 array of pixels, and the patches are arranged in the same manner as the Bayer CFA, which is called Bayer PCP-CFA in the following. The PCP pixel and buffer pixels of a patch possess individual micro-lenses while sharing the color filter of the same color. The cross-talk from the red pixels, for example, is absorbed by the green buffer pixels on its nearest neighbor, and the PCP pixel only receives the cross-talk from the buffer pixels of the same color; thus, the local color information is preserved. If the color cross-talk is severe and may penetrate a buffer pixel, more than two lines of buffer pixels can be placed around the PCP pixel.

The PCP-CFA needs both a dedicated cross-talk correction algorithm to compensate the cross-talk affecting the buffer pixels and an interpolation method to fill-in missing patches of pixels, which will be explained in following subsections.

4.2 Similarity-based blind cross-talk correction

Similarity-based blind cross-talk correction (SBC) method corrects the color cross-talk based on the assumption that buffer pixels of the same colors in periodic similarity are affected by the same coupling ratio of cross-talk, i.e., cross-talk coefficient. In the following, the proposed algorithm is first explained using the one-dimensional representation, as shown in Figure 3 and later extended to the two-dimensional system.

The following variables are defined to develop the system equation, where the index \( i \) indicates the location of the pixel. 1) Ground truth: \( G(i) \) and \( R(i) \) for green and red channels; 2) Observation, i.e., sensor output: \( G'(i) \) and \( R'(i) \); 3) Guide: \( G^#(i) \) and \( R^#(i) \), which are generated by signals from PCP pixels; 4) Estimate: \( G^*(i) \) and \( R^*(i) \), corresponding to the cross-talk corrected signals; 5) Cross-talk coefficient: \( k_{RG,+}(i) \) and \( k_{RG,-}(i) \), where the index \( i \) and subscript RG+/- corresponds to the cross-talk coefficient affecting to the pixel of \( i \)-th address, and the directional information indicates that the cross-talk is transferred from R pixel to G pixel in either positive or negative direction, respectively. The cross-talk cancellation in the

![Fig. 2. PCP color filter arrangement top view.](image)

![Fig. 3. One-dimensional representation of similarity-based blind cross-talk correction.](image)
following is intended for identifying the Estimate by deriving the cross-talk coefficient based on simultaneous equations composed of the Observation and the Guide. In this study, as an example of implementation, bilinear interpolation was applied to the signals of PCP pixels to create the Guide.

The derivation starts from a detailed description of the cross-talk model as shown in Figure 4, where coefficients \( s_0, s_1, \) and \( s_2 \) are the ratios of signals detected by the center pixel, and transferred to the neighboring pixels on the left and right, respectively. The other coefficients \( t_1, \) and \( t_2 \) are the gain factors for the transfer of cross-talk. When a unit energy enters to the \( i \)th pixel, it is distributed to the own pixel and neighbors with the specified coefficients and gain factors. The relationship between the Ground truth and the Observation is given as \( s_0G(i) + s_2t_2G(i – 1) + s_1t_1R(i + 1) = G'(i) \) using these coefficients, where the address indices for \( s \) and \( t \) are omitted. For the first order derivation of the correction algorithm, the following assumptions are given for simplicity: 1) Energy conservation \( (s_0 + s_1 + s_2 = 1) \); 2) Unity gain at transfer \( (t_1 + t_2 = 1) \); 3) The coefficients \( s_0, s_1, \) and \( s_2 \) are identical for the local pixels; 4) The spectral cross-talk is negligible compared with the optical spatial cross-talk. The assumptions transform the cross-talk equation as

\[
G'(i) = G(i) – k_{RG-}(i)[G(i) – R(i + 1)] + n_1(i),
\]

where \( k_{RG-}(i) = s_1 \) (cross-talk coefficient to negative direction), and \( n_1(i) = s_2(G(i – 1) – G(i)) \) respectively. The Ground truth \( G(i) \) can be represented by the superposition of the low-frequency and high-frequency parts of images, and it is now given as \( G(i) = G_L(i) + G_{HP}(i) = G(0) + G_{HP}(i) \) based on a proposed expression\(^6\), where the superscripts \( LP \) and \( HP \) mean the low-frequency and high-frequency parts, respectively. The derivation also assumes the Guide image created from the PCP pixels as its low-frequency part. Substituting the equation above to Equation 1 now yields to its final form:

\[
G'(i) = G_L(i) – k_{RG-}(i)(G_L(i) – R(i + 1)) + n_1(i) + n_2(i).
\]

\( n_2(i) = G_{HP}(i) – k_{RG-}(i)(G_{HP}(i) – R_{HP}(i + 1)) \).

Deriving the cross-talk coefficient \( k_{RG-}(i) \) by directly solving Equation 2 has two issues: 1) The noises: \( n_1 \) and \( n_2 \) are unknown; 2) The Guide and Observation signals contain random and photon shot noises. Assuming the coefficient is identical for the pixels in similarity and noises are uncorrelated to the Guide and Observation, solving a matrix equation built in a local window can provide more reliable estimate of the cross-talk coefficient. When using the window size of 3 as an example, both the pixels at the address \((i-6, i-5) \) in Figure 3, and \((i+6, i+7) \) (not shown in the figure) are the nearest pixels in similarity, and the matrix equation to derive the cross-talk coefficient \( k_{RG-}(i) \) is given as

\[
\begin{pmatrix}
G'(i-6) - n_1(i-6) - n_2(i-6) \\
G'(i) - n_1(i) - n_2(i) \\
G'(i+6) - n_1(i+6) - n_2(i+6)
\end{pmatrix}
\begin{pmatrix}
k_{RG-}(i-6) \\
k_{RG-}(i) \\
k_{RG-}(i+6)
\end{pmatrix}
\end{pmatrix}
\]

\( \begin{pmatrix}
G_L(i-6) - k_{RG-}(i)G_L(i-6) - R_L(i-5) \\
G_L(i) - k_{RG-}(i)G_L(i) - R_L(i+1) \\
G_L(i+6) - k_{RG-}(i)G_L(i+6) - R_L(i+7)
\end{pmatrix} \) \( G_L(i) \) (3)

The matrix equation means that the Observations \( G' \) with extra components \( n_1 \) and \( n_2 \) is expressed by the Guide image \( G^2 \) and \( R^2 \) and an unknown cross-talk coefficient. Here, the first extra component \( n_1 \) is considered to be negligible in the sub-half-micron pixel generation, as explained in Section 2. And the second component \( n_2 \) is composed of the high-frequency part which is uncorrelated to the equations of the right-hand side. Therefore, the use of regularization on Equation 3 without including \( n_1 \) and \( n_2 \) cancels the interferences that are uncorrelated to the low-frequency part and eventually gives the cross-talk coefficient for the low-frequency part. The Estimate is now given as

\[
G'(i) = G_L(i) – k_{RG-}(i)(G_L(i) – R_L(i + 1))
\]

\( G(i) = G(i) + n_1(i) + n_2(i) \)

\( G(i) – k_{RG-}(i)(G_{HP}(i) – R_{HP}(i + 1)) \) (4)

Here the equation means that the Estimate is equivalent to the Ground truth with the cross-talk on the high-frequency part uncorrected.

Figure 5 describes the extension of SBC to the two-dimensional system. With the upper-right green pixel at the address \((i, j) \) as an example, the equation with cross-
talk coefficients from "blue to green" and "red to green" is given as

\[ G'_{BG}(i,j) = G^a(i,j) - k^{(0)}_{BG}(i,j)(G^a(i,j) - R^a(i + 1,j)) \]
\[ - k^{(-1)}_{BG}(i,j)(G^a(i,j) - R^a(i + 1,j - 1)) \]
\[ - k^{(0,1)}_{BG}(i,j)(G^a(i,j) - B^a(i,j + 1)) \]
\[ - k^{(1,-1)}_{BG}(i,j)(G^a(i,j) - B^a(i - 1,j + 1)). \]  (5)

In order to simplify the calculation, the average signals of two adjacent red pixels and blue pixels are used with unified cross-talk coefficients for the red channel and blue channels, respectively, yielding to the reduced cross-talk correction equation as

\[ G'_{BG}(i,j) = G^a(i,j) \]
\[ - k_{BG}(i,j) \left[ G^a(i,j) - \frac{R^a(i + 1,j) + R^a(i + 1,j - 1)}{2} \right] \]
\[ - k_{BG}(i,j) \left[ G^a(i,j) - \frac{B^a(i,j + 1) + B^a(i - 1,j + 1)}{2} \right]. \]  (6)

Based on Equation 6, a set of total 9 equations from the window-size of 3-by-3, for example, can be generated to compose a matrix equation, one for the pixel located in the center of the window and the other 8 for the surrounding pixels in local similarity, and the Estimate is obtained in the same manner as it was applied to Equation 3.

4.3 Adaptive normalized convolution with post-convolutional-variation-minimization

Normalized convolution (NC)\(^7\), which enables the resampling of the non-uniformly sampled data, can also be applied to interpolate a large patch of missing pixels seen in the PCP-CFA arrangement. The convolution result is given as:

\[ \hat{f}(x,y) = \frac{f(x,y) \cdot c(x,y) + g(x,y)}{c(x,y) \cdot g(x,y)}, \]  (7)

where \( f, c, \) and \( g \) correspond to the convolution result, image data, certainty map, and applicability function, respectively. The optimal applicability function depends on the edge direction in the local area of the two-dimensional image, and it is often desirable to choose a variety of functions adaptively to interpolate the patches of missing pixels by employing the adaptive normalized convolution (ANC)\(^8\). In this article, an optimal applicability function is chosen among many candidates based on the proposed cost function: "post-convolutional-variation-minimization (PCVM)," given as

\[ i = \arg \min_{i \in \mathbb{N}} \sum_{y=m_y-r_y/2}^{m_y+r_y/2} \sum_{x=m_x-r_x/2}^{m_x+r_x/2} \left\| \frac{f(x,y) \cdot c(x,y) + g(x,y)}{c(x,y) \cdot g(x,y)} - f(x,y) \cdot c(x,y) \right\|. \]  (8)

Here, \( i, m_{x/y}, \) and \( r_{x/y} \) are the index of applicability function, \( x/y \) addresses of the center of the patch, and the window size in \( x/y \) directions, respectively. It assumes that normalized convolution operation with the optimal applicability function to interpolate a patch of missing pixels does not substantially change the signals around the missing patch.

5. Scalable Single-Parameter Cross-Talk Model

Unified rule to describe the cross-talk for a different pixel pitch with a single cross-talk parameter is expected in order to enable benchmarking its influence on the different imaging systems. In the following, a scalable single-parameter (SSP) cross-talk model is proposed that can control the magnitude of cross-talk with a single cross-talk parameter. The model is also accompanied by a pixel pitch for a scaling factor. Although the model may not accurately reflect the physical cross-talk phenomena, it is efficient tool to incorporate its effect in the feasibility study phase with an analytical approach.

The development of a single parameter cross-talk model begins with the following equation that shows the distribution of energy \( c(x, y, \alpha) \) when the unit energy originally arrives at a unit pixel with a size and pitch of \( p \), where \( x, y, \) and \( \alpha \) are the axes in the continuous domain and the cross-talk parameter, respectively.
Here Equation 9 is based on the following assumptions: a) cross-talk energy is conveyed to the pixels of orthogonal and diagonal nearest neighbors; b) the total signal energy is conserved after the cross-talk event; c) the ratio of diagonal cross-talk component is given as the product of those of the horizontal and vertical components. Its integral for a period of pixel pitch provides the cross-talk in the discrete domain: \( C(i, j, \alpha) \) as
\[
c(x, y, \alpha) = \frac{1}{p^2} \left[ (1 - \alpha) \text{rect} \left( \frac{x}{p} \right) + \frac{\alpha}{3} \text{rect} \left( \frac{x}{3p} \right) \right],
\]
\[
(1 - \alpha) \text{rect} \left( \frac{y}{p} \right) + \frac{\alpha}{3} \text{rect} \left( \frac{y}{3p} \right), \tag{9}
\]

Based on Equation 10, the ratio of energy entering to the center pixel \( (C_c) \), the horizontal and vertical pixels \( (C_h) \) and \( (C_v) \), and diagonal pixels \( (C_d) \) are expressed with the cross-talk parameter as follows:
\[
C_c(\alpha) = C(0, 0, \alpha) = 1 - \frac{4}{3} \alpha - \frac{4}{9} \alpha^2,
\]
\[
(11)
\]
\[
C_h(\alpha) = C(\alpha, \alpha) = 1 - \frac{4}{3} \alpha - \frac{4}{9} \alpha^2,
\]
\[
(12)
\]
\[
C_v(\alpha) = C(\alpha, -\alpha) = 1 - \frac{4}{3} \alpha - \frac{4}{9} \alpha^2,
\]
\[
(13)
\]

The model is scalable, i.e., it can generate the cross-talk of a large pixel with the superposition of the cross-talk of unit pixels, assuming the amount of energy transferred from the unit pixel is consistent for pixels of different pitches. The cross-talk between the unit pixels of the same color are complementary hence they are contained in the large pixel, and the cross-talk going outward from the large pixel are integrated with each direction and calculated as follows:
\[
C_c(\alpha, n) = C_c \left( \frac{\alpha}{n} \right),
\]
\[
(14)
\]
\[
C_h(\alpha, n) = \frac{1}{n^2} \left( \frac{\alpha}{9} \alpha^2 \right) = C_h \left( \frac{\alpha}{n} \right),
\]
\[
(15)
\]
\[
C_v(\alpha, n) = \frac{1}{n^2} \left( C_h + C_v \right) + \left( n - 2 \right) (C_h + 2C_d) = C_h \left( \frac{\alpha}{n} \right).
\]
\[
(16)
\]

Here, \( n \) is the number of unit pixels extended in the horizontal and vertical directions in the case of a square pixel. The derivation indicates that the cross-talk parameter of the large pixel enlarged by the factor of \( n \)-by-\( n \) reduces to \( 1/n \) times compared with the original. Table I shows the relationships between the cross-talk parameter and the cross-talk to all the directions for the pixel pitch ratio of \( n=1 \).

### 6. Experimental Setups

#### 6.1 Experimental setup

The effectiveness of the PCP concept was evaluated by applying a slanted-edge based spatial frequency response (SFR) modular transfer function (MTF) method on a black and white simulated image with an edge angle of 5 degrees. The MTF was calculated using a "sfrmat3" MATLAB program. The testing setup is summarized in Table II. The cut-off modifier \( k_{MTF} \) is introduced to adjust the lens point-spread-function by attenuating the frequency response to incorporate the effect of lens MTF degradation compared to the diffraction limit MTF, which often occurs when the image height increases. The effect of LCA is implemented as parallel shifts of red and blue planes assuming the size of the simulated area is small compared to the total image area. An interpolation algorithm for Bayer CFA employs Adaptive Residual Interpolation.

### Table I. Cross-talk Parameter and Cross-talk Ratio for Unit Pixel.

| cross-talk parameter(s) | pixel size: p (unit size, n=1) |
|-------------------------|--------------------------------|
| **Center (C)**          | **VH (Cv, Ch)** |
| 0                       | 100.0%  0.0%        |
| 0.1                     | 87.1%   3.1%  0.1% |
| 0.2                     | 75.1%   5.8%  0.4% |
| 0.3                     | 64.0%   8.0%  1.0% |
| 0.4                     | 53.9%   9.6%  1.9% |
| 0.5                     | 44.4%   11.1% 2.8% |
| 0.6                     | 36.0%   12.0% 4.0% |
| 0.7                     | 28.4%   12.4% 5.4% |
| 0.8                     | 21.8%   12.4% 7.1% |
| 0.9                     | 16.0%   12.0% 9.0% |
| 1.0                     | 11.1%   11.1% 11.1% |

### Table II. Experimental Setup.

| Camera simulator parameters | Settings |
|-----------------------------|----------|
| **Lens F#**                 | 1.5      |
| **Cutoff Modifier**         | 0.7 (equivalent F# 2.14) |
| **CFA**                     | 0.5µm-Bayer, 0.5µm-Bayer, 0.5µm-Bayer-PCP |
| **Cross-talk Parameter**    | 0 - 0.7 |
| **LCA**                     | SRG for Bayer-PCP, 1.02µm |
| **X-talk cancellation**      | no cancellation for Bayer |
| **Interpolation algorithm**  | Adaptive residual interpolation for Bayer, PVC-ANC for Bayer-PCP |
6.2 Simulator detail

A custom-made camera simulator was implemented based on MATLAB, which includes the diffraction by the taking-lens, detector footprint, and interpolation, as well as pixel random noise, photon shot noise, LCA, and cross-talk as noise-options. The object image was generated by using 0.1 \( \mu \text{m} \) virtual sampling grid and projected onto the simulated image sensor with a magnification of 1.0×. The aperture of the pixel was set to the square-shape with the height and width of the pixel pitch. The simulator was validated by confirming the matching on the MTF curves between the ones from the simulated image without noise option, and the diffraction theory. Photon shot noise was generated firstly by defining the normalized number of electrons per 1 \( \mu \text{m}^2 \)-unit-pixel which corresponds to a signal level, and then by using the number of electrons per actual pixel as the average value parameter for the Poisson probability distribution function. Regarding the slanted-edge image, the signal levels of the R, G, and B signals on the brighter side were set to 200DN (Digital Number) for a bit depth of 8-bit per channel.

7. Results and Discussions

Figure 6 shows the comparison of the MTF curves for three types of pixel arrays obtained from the result of simulation based on ideal situation, i.e., without cross-talk, LCA nor noise options. All the MTF curves are almost indentical since the frequency response of the system is determined by the taking-lens.

Figure 7 compares the MTF curves simulated with a case with the cross-talk of \( \alpha = 0.6 \) and LCA = 2 \( \mu \text{m} \). The parameter \( \alpha = 0.6 \) used in the simulation corresponds to the situation that 12% of the incoming energy to a 0.5 \( \mu \text{m} \)-pitch pixel is transferred from the pixel of origin to an orthogonal neighbor, which causes significant MTF reduction when combined with LCA of 2 \( \mu \text{m} \). The PCP-CFA showed the best frequency response from 0 to 300 lp/mm, and comparable results to the other two for higher frequency than 300 lp/mm. PCP, by overall, provides better frequency response than others.

Figures 8 shows the MTF at the frequency of 200 lp/mm for LCA = 2 \( \mu \text{m} \) with the cross-talk parameter as a variable. Bayer PCP-CFA pixel of 0.4 \( \mu \text{m} \)-pitch was also tested to see the effect of further pixel-pitch shrink. Bayer CFA with 0.5 \( \mu \text{m} \)-pitch shows a significant decrease of frequency response, even with a minor cross-talk of \( \alpha = 0.05 \), whereas Bayer PCP-CFA renders the reduction relatively insensitive to the increase of cross-talk. The use of 0.4 \( \mu \text{m} \)-pitch PCP-CFA pixel increases the frequency response owing to the increased spatial frequency of the Guide image. The gain is, however, not
so obvious, since the frequency response is closer to the limit of the Lens MTF of 65%~70% @ 200 lp/mm as observed in Figure 6. Bayer CFA with 1.0 µm-pitch does not show drastic frequency reduction unlike Bayer CFA with 0.5 µm-pitch since the larger pixel size makes it insensitive to the effects from the cross-talk and LCA. The PCP-CFA, however, still performs better frequency response than that of 1.0 µm-pitch Bayer-CFA when the cross-talk comes to be obvious.

Figure 9 is the MTF plot of 0.5 µm-pitch PCP-CFA simulated with photon shot-noise option. For example, the average number of signal electrons with GL(gray level) = 1000 e-/µm² corresponds to 250 e-/(0.5 µm)², which yields to SNR of 15.8 (24 dB). Their frequency responses still show comparable trends to that of the noiseless condition, indicating the PCP concept with SBC and ANC-PCVM also works under the influence of noise.

Figure 10 shows the result of color recovery by SBC combined with PCP-CFA. Saturation in the HSV color space is used as an index of color purity, where the value 1.0 corresponds to the maximum purity and zero indicates a colorless gray. In the slanted edge image, the signal of green plane is set to zero, making the original color of the brighter part of the slanted edge image purple color with its saturation of 1.0. As the cross-talk parameter \( \alpha \) increases, Bayer CFA shows a drastic decrease in saturation, whereas the degradation is slower in PCP-CFA since preserved-color pixels help to maintain the chroma information. The SBC combined with PCP-CFA almost entirely recovers saturation for all the simulated range of \( \alpha \). This result can be recognized as a clear advantage of PCP-CFA.

Figure 11 shows a peak-signal-to-noise-ratio (PSNR) comparison for the simulated images captured with RGB image sensor and 3-by-3 PCP-RGB image sensor under different cross-talk parameters. The pixel pitch is 0.5 µm. The source image for simulation consists of a hairy brown teddy bear with pink flowers in green-leaf background. The output RGB image size after simulation process is 1022-by-686 pixels. The images taken by RGB sensor with \( \alpha = 0 \) is used for the reference images to derive the numbers of PSNR for images with cross-talks. For the cases of both LCA=0 µm and 2 µm, the pictures of RGB sensor show clear degradations as \( \alpha \) increase due to the reduction of color saturation as also suggested in the result in Figure 10, whereas those of PCP sensor can retain PSNR of more than 30dB.


8. Conclusion

The preserved-color-pixel concept was proposed with supporting algorithms: SBC and PCVM-ANC. They were effective in the sub-half-micron generation, where the signals of adjacent pixels showed similar values. The concept was verified with image simulation, and its effectiveness was quantified by using the slanted-edge based SFR-MTF method under the influence of cross-talk, represented by the SSP cross-talk model. The simulation with 0.5 µm-Bayer PCP, 0.5 µm-Bayer, and 1.0 µm-Bayer CFAs with LCAs of 2.0 µm showed that the PCP concept was capable of providing overall a superior frequency response when the system had been affected by the complication of color cross-talk and LCA. The result didn’t change under the influence of photon shot-noise. Furthermore, the PCP concept performed excellent color correction at a low spatial frequency pattern, which was also verified with PSNR analysis using the simulation with a sample real-world image. It showed more than 30dB PSNR for the cross-talk parameter (α) of less than 0.7.

Three challenges remain for future study. Firstly, the derivation of SBC was done with introducing several steps of approximations. Although the first feasibility study of SBC worked effectively, eliminating those approximations could provide better cross-talk cancellation. Secondly, the simple bilinear interpolation, which was used to generate the Guide image, determined the upper limit of frequency when cancelling the cross-talk. Consequently, more intelligent interpolation is preferred. Finally, the reduction of image resolution in a specific type of image needs to be resolved. When Bayer CFA arrangement is employed, it is often challenging to recover spatial frequency of the image without information of green channel, because the green channel has twice high sampling frequency than others. It may occur when the object has a color of red, blue, magenta, or purple with saturation of 1.0 in HSV color space and also there is only negligible spectral cross-talk. The Bayer-PCP CFA arrangement is no exception, which suffers even more serious spatial frequency reduction due to the existence of large voids, i.e. missing patches of green pixels. Optimization of the color filter spectrum to balance the color fidelity and spatial resolution after demosaicing is necessary.

References

1) K. Hirakawa, "Cross-talk explained," in Proc. of International Conference on Image Processing, ICIP, pp.677-680, San Diego, CA, USA (2008)
2) A. Getman, T. Uvarov, Y.I. Han, B. Kim, J.C. Ahn and Y.H. Lee, "Cross-talk, color tint and shading correction for small pixel size image sensor," in Proc. 2007 International Image Sensor Workshop, pp.166-169, Ogunquit, Maine, USA (2007)
3) G. Agranov, V. Berezin and R. Tsai, "Cross-talk and Sub-Pixel Distribution of Sensitivity in Color CMOS Image Sensor," in Proc. 2001 IEEE Workshop on Charge-Coupled Devices and Advanced Image Sensors, Lake Tahoe, NV, USA, 133-136 (2001)
4) G.D. Boreman, "Modulation Transfer Function in Optical and Electro-Optical Systems,"SPIE Press (2001)
5) A. Tournier, F. Leverd, L. Favernec, C. Perrot, L. Pinzelli, M. Gafetain, N. Cherault, D. Jeanjean, J. Carrere, F. Hirigoyen, L. Grant and F. Roy, "Pixel-to-Pixel isolation by Deep Trench technology: Application to CMOS Image Sensor," in Proc. 2011 International Image Sensor Workshop, Hakkaido, Japan (2011)
6) J.T. Korneliussen and K. Hirakawa, "Camera processing with chromatic aberraiton," IEEE Transactions on Image Processing, 23, 10, 4539-4552 (2014)
7) H. Knutsson and C. Westin, "Normalized and Differential Convolution: Methods for Interpolation and Filtering of Incomplete and Uncertain data," in Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. June , no. 16-19, 515-523 (1993)
8) V. Argyriou, T. Vlachos and P. Piroddi, "Gradient-Adaptive Normalized Convolution," IEEE Signal Processing Letters, Vol.15, 489-492 (2008)
9) P. Burns, "sfrmat3: SFR analysis for digital cameras and scanners," 12-May -2015, [Online] Available: http://losburns.com/imaging/software/SFRedge/ sfrmat3_post/index.html
10) Y. Monno, D. Kiku, M. Tanaka and M. Okutomi, "Adaptive residual interpolation for color and multispectral image demosaicking," Sensors, 17, 12, 1-21 (2017)