Image Quality Impact and Comparison of Selected State-of-the-Art CFA Interpolation Techniques

Petr Dostal¹, Milos Klima²

Faculty of Electrical Engineering, Czech Technical University in Prague
Technika 2, 166 27, Praha 6, Czech Republic
¹dostape2@fel.cvut.cz, ²klima@fel.cvut.cz

Abstract
Among recent multimedia imaging systems single chip cameras are dominating over three chip cameras. The single chip cameras with a CFA as a color splitting system provide the R G B images with a usually different an d in complete sampling. The calculation of full resolution R G B images brings some problems and selected interpolation technique can affect the final image quality significantly. This paper summarizes experimental results and performance comparison for four recent CFA in terpolation techniques. Having the reference with re constructed im ages av aila ble, the performance of each technique has bee n evaluated by objective criteria - both PSNR and SSIM.

Index Terms: demosaicing, Bayer pattern, CFA, mosaic image, image quality

I. Introduction
Almost all im aging systems are designed as trichromatic. At present, all cameras are us ed only one chip to capture color information because of ex act matching of scanning rasters is very difficult to be satisfied especially in terms of very high resolution chips (CCD). Misalignment of scanning rasters leads to smearing a final image; the Foveon® solves this problem by manufacturing a special silicon chip. This chip separates red, green and blue part of visible light in different levels of silicon substrate. In most cases, a planar on-chip color splitting system is used to separate the red, blue and green channel. The color filter array known as CFA created on the chip is used for spatial sampling and color separation together. Therefore each pixel can capture information about one primary color. This image is called a mosaic image. The most spread CF A architecture is known as a Bayer pattern [1] (see Figure 1 (a)). As you can see, the pattern contains twice more green color pixels because this color contains most information about luminance and therefore about details in a scene. Th e additive colors are used in the Bayer pattern. A camera is manufactured by Bayer or overlaying filters containing subtractive color. Hence the CFA is based on the additive colors, which are less sensitive than CFA based on sub traceble colors [2] (see Figure 1 (b)). On the other hand, using CFA on additive colors is more practicable because a display unit is based on the same color model and the processing of GB is less computationally complex. Usage of CFA in pixel location is done on the manufacturer to capture more important higher sensitivity or less computational complexity. For instance DSLR Nikon D800 uses Bayer pattern unlike the DSLR Kodak DCS620x exploiting CMY CFA [2].

In order to reconstruct a full color image, th e color information must be computed by an algorithm known as demosaicing. In literature, many interpolation algorithms have been already presented. The earliest proposed techniques were based on well-known interpolation techniques as nearest neighbor, bilinear interpolation, bicubic in terpolation. These techniques were not able to provide high-level performance. Therefore, another techniques exploiting feature of image known as inter-channel (sub-band) correlation were proposed. Alleson in [3] proposed an algorithm based on the human visual system (HVS) per ception. This algorithm exploits the inter-channel correlation to compute the full im age age b y y a summation of chrominance and luminance; both chrominance and luminance are computed directly from the mosaic image by proper filtering. Better results can be achieved by using adaptive methods changing the computation according to the content of input image. Su in [4] proposed the wavelet-based algorithm exploiting the inter-channel correlation. The next improvement can be achieved by a directional interpolation; interpolation along the edges is more effective than across them, because the signal along the edges contains lower frequency components which are better reconstructed. Another approach proposed by Hirakawa mentioned in [5] exploiting a directional interpolation with sub-band correlation. This algorithm creates two full color images (horizontal and vertical d irected interpolation is used) which leads to two estimated candidates for each color in each pixel location. Then used the local homogeneity of the image in each pixel location to decide which candidate in each pixel location is better. The another algorithm utilizing the sub-band correlation and directional in terpolation was proposed by Menon [6]. This algorithm creates two full color images (horizontal and vertical full-resolution) full-resolution green channels. The decision which green information of them was used in each pixel location is chosen according to the gradient map. Having reconstructed full-resolution green channel, the reconstruction of the red and blue channel is utilizing sub-band correlation is performed. Chung in [7] proposed the novel approach combining the edge information with adaptive heterogeneity projection in order to decide how much of the information from each of four adjacent pixels will be used to reconstruct the missing information in each pixel placed in the center of adjacent pixels. The goal of this study is to compare the performance of selected algorithms; bilinear interpolation, algorithm proposed by Alleson [3], Hirakawa [5], Menon [6] and Chung [7].

Figure 1: CFA patterns,
(a) Bayer pattern,
(b) CMY pattern.
The special synthetic image was used for this purpose and two objective criteria, P SNR and S SIM were used for the performance evaluation. The content of this paper is organized as follows. In section 2, the selected interpolation algorithms are described. Section 3 is dedicated to the preparation for testing. In section 2, the synthetic image generating, the description of objective criteria and details of subjective testing is mentioned. Furthermore, the chosen objective criteria are discussed. The performance of the chosen algorithm is presented in section 4. The reconstruction of spatial frequency is discussed in chapter 5. Finally, in the last six chapters the performance of the algorithms is reported in the conclusion.

2. Interpolation algorithms to be evaluated

This chapter is dedicated to detailed description of selected algorithms. The bilinear interpolation was used as a reference to give some idea about the performance of simple interpolation technique.

2.1. Linear demosaicing inspired by human visual system

This approach is based on a reconstruction of full resolution luminance image directly from the mosaic image utilizing the inter-channel correlation. All eysson [3] claim that the spatial information about luminance is preserved with a full spatial resolution in spite of the sampling and th at the chromatic information is sub-sampled. The imaging process is described in Retina of the single chip camera with CFA. Therefore the mosaic image can be expressed as summation of full-resolution luminance and sub-sampled chrominance. Furthermore, these in formation about the full-resolution luminance can be recovered directly by filtering the mosaic image. Alleysson in [3] proved that the full-resolution luminance is in fact the region placed into the center of the mosaic image spectrum and the chrominance represents the other regions placed on the border of the spectrum frequency domain. The filter used to recover the luminance from the mosaic image with the spectrum of the mosaic image is shown in Figure 2, where f and f_y are the spatial frequencies in horizontal and vertical direction, respectively. The filtering is done in a small neighboring area which are similar in luminance and chrominance. The decision is based on a local homogeneity, which is measured by the total number of pixels located within a small neighborhood area which are similar in luminance and chrominance within a small neighboring area. The pixel in question is exploited for each horizontal or vertically interpolated pixel in the picture. The red and blue channels are reconstructed from the difference image and full-resolution green channel. The strategy for estimation of red channel is shown in Figure 3. This approach described in [5] uses an indicator to choose between horizontally or vertically interpolated intensities instead of choosing the interpolation direction based on edge indicators, which are used in the edge-directed interpolation. For choosing between horizontal or vertically interpolated intensities instead of choosing the interpolation direction based on edge indicators, which are used in the edge-directed interpolation, the chrominance information is computed from sub-sampled chrominance information. The full-resolution chrominance information is computed from the full-resolution luminance image and full-resolution chrominance. Note, the full-resolution chrominance is composed of the full-resolution luminance and full-resolution chrominance. N ote, the full-resolution color channels are composed of the full-resolution luminance and full-resolution chrominance. The performance of this approach depends on the filter design and the interpolation technique used to recover the full-resolution chrominance image. In our case, the bilinear interpolation is used.

2.2. Adaptive homogeneity directed demosaicing algorithm

This approach proposed by Hirakawa [5] exploits the sub-band correlation. Because of high correlation between color channels, the difference image (difference between red and green channel or between blue and green channel) contains mostly low frequency components. The high frequency components contained in the difference image are related to edges in the picture. The red and blue channels are reconstructed from the difference image and full-resolution green channel. The strategy for estimation of red channel is shown in Figure 3. This approach described in [5] uses an indicator to choose between horizontally or vertically interpolated intensities instead of choosing the interpolation direction based on edge indicators, which are used in the edge-directed interpolation. For choosing between horizontally or vertically interpolated intensities instead of choosing the interpolation direction based on edge indicators, which are used in the edge-directed interpolation, the chrominance information is computed from sub-sampled chrominance information. The full-resolution chrominance information is computed from sub-sampled chrominance information by the interpolation; the low-pass filtering can be used. Note that the chrominance contains only low frequency components due to the high inter-channel correlation. The full-color image is computed by adding the luminance and chrominance. Note, the full-resolution color channels are composed of the full-resolution luminance and full-resolution chrominance. The performance of this approach depends on the filter design and the interpolation technique used to recover the full-resolution chrominance image. In our case, the bilinear interpolation is used.
2.3. Demosaicing with directional filtering and a posteriori decision

This approach proposed by Menon [6] exploits the interpolation artifacts are reduced. Menon in [6] mentioned the interpolation artifacts affect mostly high frequency content of each pixel. Correction of these artifacts is carried out by exploiting an in-ter-channel correlation of the three primary colors and separating low- and high-frequency components in each pixel and removing the high frequencies of the unknown color channels. In each pixel, the color channel correlation is computed in one step. By applying the special convolution core, the remaining red and blue channels are estimated. The red and blue information placed in the RG row is computed in horizontal direction and the missing blue information at the same position is computed in vertical direction. The red or blue missing value at blue or red position are recomputed by the same directed interpolation as the green color. After reconstructing the full-resolution red and blue color channel, the full-color image is created.

The last step of this algorithm is a refining step. In this step the interpolation artifacts are reduced. This approach proposed by Chung [7] utilizes an inter-channel correlation. At first step full-resolution green color channel is exploited, therefore the calculation considers only one case of calculation missing green value, see in Figure 4 (a), 2) vertical variation, see in Figure 4 (b), 3) the other variation, see in Figure 4 (c).

Figure 4: Three possible cases of calculation missing green value, (a) horizontal variation, (b) vertical variation, (c) other variation.

The arrows in Figure 4 denote the pixels used for the reconstruction of central missing value. In order to estimate the missing green value more accurately from its four adjacent pixels, the four proper weights depending on gradient information obtained from mosaic image for the horizontal and vertical heterogeneity projection are assigned to four corresponding spectral-correlation terms to affect their impact on reconstructed green image; for more details refer to [7]. For computation of four weights, the central eight adjacent pixels are used. After green channel reconstruction, the remaining red and blue channels are estimated. The red and blue missing value at blue position is computed from four adjacent spectral-correlation term with four proper weights; in this case the information from the heterogeneity projection is not exploited, therefore the calculation consists of only one case as shown in Figure 5. The blue and green missing values at the red positions are computed by following the same strategy.

2.4. Demosaicing with gradient edge detection mask and adaptive heterogeneity projection

This approach proposed by Chung [7] utilizes an inter-channel correlation, favoring exploitation of the information about edges of the mosaic image for the decision, which chooses the pixels to be used for the interpolation of missing value. Firstly, the information about the horizontal, vertical and both diagonal (45°, 325°) edges is calculated from the luminance by special proper convolution core (the size of the core is 5 x 5) composed of combination of the luminance core (see in Figure 2) with Sobel operator used for specific direction (horizontal, vertical or diagonal); for more details about the core design see in [7]. The special convolution core is applied directly on the mosaic image so that the information about edges contained in the luminance is computed in one step. By applying the special convolution core, the four full-resolution images containing information about the edges in four directions are created. The next step is the computation of heterogeneity projection for the mosaic image. The projection is realized by running two 1D Laplacian operators, first for horizontal and second for vertical directions. Each of two projections uses proper a mask size for each pixel position; the size of the 1D mask varies according the location of the pixel in question from 5 – 11 pixels. If the pixel is located on the horizontal oriented edge, the mask is computed for horizontal projection will be larger; otherwise, the mask will be smaller. After computing horizontal and vertical heterogeneity projection, the green channel is reconstructed. According to this projection, the calculation of missing green value considers three cases: 1) horizontal variation, see in Figure 4 (a), 2) vertical variation, see in Figure 4 (b), 3) the other variation, see in Figure 4 (c).

Figure 5: The computation of missing red value on the blue position.

3. The preparation for testing

This chapter deals with the description of synthetic test image, the description of used objective criteria used for performance evaluation and the description of subjective testing.

3.1. The generating of synthetic images

The monochromatic synthetic image (226x226 pixels) is composed of 16 boxes; see in Figure 6 (a). Each box has size
50 x 50 points. The rows from the top contain horizontal, diagonal 45°, diagonal 135° and vertical direction of spatial frequency component. The columns from the right contain concrete spatial frequencies; \( f_i/2, f_i/4, f_i/6, f_i/8 \), where \( f_i \) denotes the sampling raster frequency. In other words, the boxes on the right side contain the high est fr equency component which can be captured. The Figure 6 (c) shows the modulo spectra of monochromatic synthetic image shown in Figure 6 (a). You can see the peaks in position 1/2, 1/4, 1/6 and 1/8 represents the mentioned spatial frequencies contained in the synthetic testing image. After sampling testing image by the Bayer CFA (Figure 1 (a)) the mosaic image is created (Figure 6 (b)). Due to the frequencies higher than 1/4, there is an aliasing effect in the mosaic image, the effect of aliasing is evident. (Figure 6 (d)). Due to the frequencies higher than 1/4, there is an aliasing effect in the mosaic image, the effect of aliasing is evident.

3.2. Objective criteria

Two objective criteria were used for performance evaluation, CPSNR and SSIM.

3.2.1. CPSNR

The color peak signal to noise ratio expressed in dB gives the quality of reconstructed image in terms of the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation. This metric is computed as follows

\[
CPSNR = 10 \log_{10} \left( \frac{255^2}{CMSE} \right) \tag{1}
\]

where CMSE is objective parameter Mean Square Error for color images calculated according the equation

\[
CMSE = \frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \left[ \sum_{c=1}^{Ch} \left( I_{Orig}(i,j) - I_{Rec}(i,j) \right)^2 \right], \tag{2}
\]

where \( M, N \) denotes the sizes of image in vertical, horizontal direction, respectively. \( Ch = \{r, g, b\} \); \( I_{Orig}(i,j), I_{Rec}(i,j) \) denotes the three color components of the pixel at location \((i, j)\) in the original image; \( I_{Rec}(i,j), I_{Rec}(i,j), I_{Rec}(i,j) \) denotes the three color components of the pixel at location \((i, j)\) in the reconstructed image.

3.2.2. SSIM

The Structural similarity index, SSIM, compares local patterns of pixel intensities that are normalized for luminance and contrast. Wang [8] claims that a measure of structural information change can provide a good approximation of perceived image distortion. By exploring the structural information in the image, the influence of illumination is separated. Illumination can affect the hue and the contrast in the image, but the structural information is independent on it. The change of all the characteristics of image, luminance, contrast and structural information does not affect the quality of final image. Despite this, the perceived quality of image is most dependent on the structural information.

For overall image quality, the mean SSIM index is expressed in dB gives the color peak signal to noise ratio for color images computed according the equation

\[
SSIM \left( x_p, y_p \right) = \left( \frac{2 \mu_x \mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \right) \left( \frac{2 \sigma_{xy} + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \right) \tag{3}
\]

where \( \mu_x, \mu_y \) are weighted windows (size 11x11) from original image and reconstructed image at position \( p \), respectively. The circular symmetric weighting Gaussian function is used, for more details see in [8]. The circular symmetric weighting Gaussian function is used, for more details see in [8].

For overall image quality, the mean SSIM index is expressed in dB gives the color peak signal to noise ratio for color images computed according the equation

\[
SSIM \left( I_{Orig}, I_{Rec} \right) = \frac{1}{M} \sum_{p=1}^{M} SSIM \left( x_p, y_p \right) \tag{4}
\]

where \( I_{Orig} \) and \( I_{Rec} \) are original and reconstructed color image, respectively.

3.3. Subjective testing

In order to get some initial estimate of real subjective quality evaluation the preliminary subjective testing has been performed. The subjective testing procedure has been done according to the modified ITU-R Rec. BT-500 in the version of DSCQS method. The applied approach has been simplified and the results are preliminary at the moment and we plan to perform detailed subjective test in consequent work. So far altogether five observers have been evaluating the experimental results of tested demosaicing procedures. The scale from 0 to 5 has been used for the quality evaluation. The subjective results are demonstrated in Figure 7.

![Figure 6: Testing image.](image)
4. Results

To evaluate the algorithm performance among four concerned algorithms, none of them uses any postprocessing embedded refinement scheme; the Hirakawa’s algorithm [5] and Menon’s algorithm [6] was edited for this purpose.

The reconstructed testing image is shown in the Figure. The performance computed according the PSNR and SSIM is mentioned in Table 1. Chung’s algorithm [7] gives the higher results in terms of PSNR and SSIM. In terms of subjective testing, the assessors found no difference between Hirakawa’s [5] and Chung’s [7] algorithm. The most significant difference in evaluation is in the case of Menon’s algorithm [6]; see Figure 7.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>SSIM [%]</td>
<td>56.18</td>
<td>89.17</td>
<td>90.64</td>
<td>78.35</td>
<td>91.61</td>
</tr>
</tbody>
</table>

Table 1: PSNR and SSIM results.

The Chung’s algorithm [7] was able to reconstruct the most of boxes containing the diagonal frequency components with lower level of color artifacts. The algorithms [5], [6], [7] handle the reconstruction of horizontal and vertical spatial frequency components except the finest details on the sub-sampled channels (see in Figure 6 (b)). Therefore these parts of testing images are very difficult to be reconstructed correctly. Note, that the Hirakawa’s [5] and Menon’s [6] algorithm was able to reconstruct always one of the mentioned boxes except the horizontal and vertical stripes containing false color, respectively. None of the selected algorithms was able to reconstruct the finest details in terms of horizontal and vertical direction. In the case of finest details, there are green-purple strips in all cases. The boxes containing the diagonal frequency at fs/6 was reconstructed with the green color artifacts. Alleysson’s algorithm [3] gives almost identical results in the case of reconstructed box es as the bilinear interpolation in spite of Alleysson’s algorithm overcomes the performance of the bilinear interpolation.

Finally, we can conclude that the best performance in the case of reconstruction spatial frequency components exhibits Chung’s algorithm described in [7]. On the other hand, this algorithm is not able to compute correctly the areas containing the horizontal and vertical frequency components at fs/2. In contrast to algorithms proposed by Hirakawa [5] and Menon [6], each of them was able to reconstruct the finest details on the sub-sampled channels. Menon’s [6] algorithm was able to reconstruct always one of the mentioned boxes except the horizontal and vertical stripes containing false color, respectively. None of the selected algorithms was able to reconstruct the finest details in terms of horizontal and vertical direction. In the case of finest details, there are green-purple strips in all cases. The boxes containing the diagonal frequency at fs/6 was reconstructed with the green color artifacts. Alleysson’s algorithm [3] gives almost identical results in the case of reconstructed boxes as the bilinear interpolation in spite of Alleysson’s algorithm overcomes the performance of the bilinear interpolation.

6. Acknowledgements

This work has been supported by the project of the Czech Grant Agency No. P102/10/1320 “Research and modeling of advanced methods of image quality evaluation”. The authors would like to acknowledge the support of Mr. Axel de Cambourg, Francois Cottereau and Matthieu Bleicher for performing the subjective tests.

7. References


Figure 7: PSNR, MSSIM and Subjective results.

Figure 8: Reconstructed testing images, (a) Bilinear interpolation, (b) Hirakawa’s algorithm [5], (c) Menon’s algorithm [6], (d) Alleysson’s algorithm [3], (e) Chung’s algorithm [7].