3rd International Workshop on Perceptual Quality of Systems (PQS 2010)
6-8 September 2010, Bautzen, Germany

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

Petr Dostal [1], *Milos Klima*[2]

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

Abstract

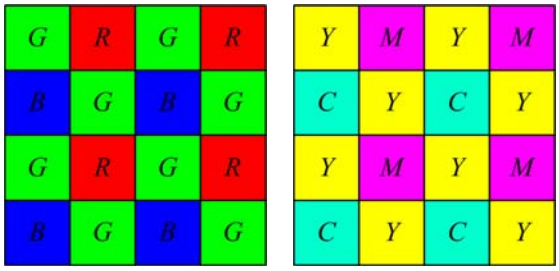
Among recent multimedia imaging systems single chip cameras are dominating over three chip cameras. The single chip cameras with a CFA as a color splitting system provide the RGB images with mutually different and incomplete sampling. The calculation of full resolution RGB images brings some problems and selected interpolation technique can affect the final image quality very significantly. This paper summarizes experimental results and performance comparison for four recent CFA interpolation techniques. Having the reference with reconstructed images available, the performance of each technique has been evaluated by objective criteria - both PSNR and SSIM.

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

1. Introduction

Almost all imaging systems are designed as trichromatic. At present, all cameras use only one chip to capture color information because the condition of exact matching of scanning rasters is very difficult to be satisfied especially in terms of very high resolution chips (CCD). Mismatching 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 as a *mosaic image*. The most spread CFA structure is known as a *Bayer pattern* [1] (see Figure 1 (a)). As you can see, this pattern contains twice more green color pixels because this color contains most information about luminance and therefore about details in a scene. These additive colors used in the Bayer CFA are manufactured by overlaying filters containing subtractive color. Hence the CFA based on the additive colors are less sensitive than CFA based on subtractive colors [2] (see Figure 1 (b)). On the other hand, using CFA based on additive colors is more practical because all display units are based on the same color model and the processing RGB is less computationally complex. Usage of some of mentioned CFA type depends only on the manufacturer – to find more important higher sensitivity or less computational complexity. For instance DSLR NikonD700 uses Bayer pattern unlike the DSLR Kodak DCS620x exploiting CMY CFA [2].

In order to reconstruct a full color image, the missing 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



(a) (b)

Figure 1: *CFA patterns,*
(a) *Bayer pattern,*
(b) *CMY pattern.*

neighbor, bilinear interpolation, bicubic interpolation. 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. Alleyson in [3] proposed an algorithm based on the *human visual system* (*HVS*) perception. This algorithm exploits the inter-channel correlation to compute the final image by 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 their computation according 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 (horizontally and vertically directed interpolation is used) which leads to two estimated candidates for each color in each pixel location. Then utilized 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 interpolation was proposed by Menon [6]. This algorithm creates two (horizontal and vertical) 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 channels utilizing sub-band correlation is perfomed. Chung in [7] proposed the novel approach combining the edge information with adaptive heterogeniety projection in order to decide how much of the information from each of four adjacent pixels will be used to reconstruction of missing information in the 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 Alleyseon [3], Hirakawa [5], Menon [6] and Chung [7].

10.21437/PQS.2010-21

The special synthetic image was used for this purpose and two objective criteria, PSNR and SSIM 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 this chapter, 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 chosen algorithm is presented in section 4. The reconstruction of spatial frequencies is discussed in chapter 5. Finally, in the last sixth chapter we report 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. Alleysson [3] claims that spatial information about luminance is preserved with a full spatial resolution in spite of retina sampling and that the chromatic information is sub-sampled. This implies that the retina samples color information as the single chip camera with CFA. Therefore the *mosaic image* can be expressed as summation of full-resolution luminance and sub-sampled chrominance (chrominance is a difference between color information and luminance). Furthermore, the information 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 *mosaic image* spectrum and the chrominance represents the other regions placed on the border of the 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_x and f_y are the spatial frequencies in horizontal and vertical direction, respectively. The 1/2 denotes the half of the sampling frequency, which is given by the resolution of the chip. A magnitude frequency spectrum of a mosaic image is used for better demonstration of suppressed parts. The filter was designed to recover the central part as precisely as possible when the computational demands are moderate; for more details about filter design see in [3]. After computing the full-resolution luminance, this part is subtracted from the *mosaic image* in order to obtain the 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 componets 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 full-resolution luminance and full-resolution assigned chrominance. The performance of this approach depends on the filter design and the interpolation technique used to recover the full-resolution chrominance signal. In our case, the bilinear interpolation is used.

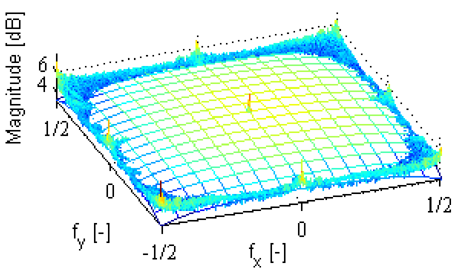


Figure 2: *Transfer function of filter used to recover luminance directly from mosaic image.*

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 method 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, the similarity of the luminance and chrominance within a small neighboring area of the pixel in question is exploited. The R, G and B channels are firstly interpolated horizontally and vertically by exploiting the inter-channel correlation, thus, the each missing color sample in R, B and G channel can be chosen from the horizontal I_H or vertical I_V full color reconstructed image, respectively. Since the decision between using horizontally or vertically interpolated pixel is done in the CIELAB space, both of the horizontally and vertically images have to be converted into this space. The decision is based on a local homogeneity, which is measured by the total number of pixels located within small neighboring area which are similar in luminance and chrominance to the pixel to be interpolated. When the value of Euclidean distance between two pixels is lower than a threshold, these pixels are considered as similar. Following this strategy, each pixel is chosen whether from the horizontally interpolated image or from vertically interpolated image and the full-color reconstructed image is obtained. All three color information in each pixel location are chosen either from horizontally I_H or vertically I_V reconstructed full color image. For more details, see in [5].

The last step of this algorithm is a refining of reconstructed image. In this step the interpolation artifacts are reduced and the signal to noise ratio is improved.

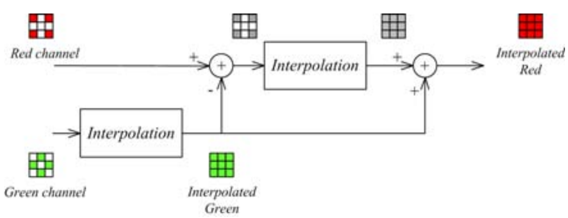


Figure 3: *Estimation of red color channel.*

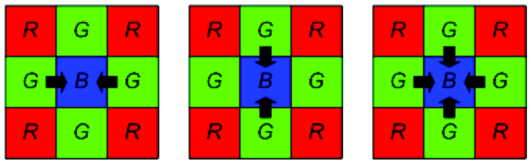
2.3. Demosaicing with directional filtering and a posteriori decision

This approach proposed by Menon [6] exploits the inter-channel correlation. At first step the horizontal and vertical full-resolution green color channels are reconstructed by exploiting the inter-channel correlation with horizontal and vertical edge-directed interpolation, respectively. After reconstructing both, horizontal and vertical, full-resolution green color channels, the decision is made to select the right direction of interpolation which gives better results. The interpolation along the edge is more effective than interpolation across them because the gradient of adjacent pixels is smaller than the gradient located across the edge. This information is exploited to detect the existence and direction of any edges in the picture. By comparing horizontal and vertical gradients of adjacent pixels the appropriate direction of interpolation in some pixel location can be found. The information about appropriate direction of interpolation is computed only for red and blue locations, in other words, only for locations where the green is missing. The computing of both gradients is mentioned in [6]. The full-resolution green color channel is reconstructed according to the result of comparing both gradients for each pixel location. Having the full-resolution green channel, the red and blue full-resolution color channels are computed by exploiting the sub-band correlation. The missing red or blue information placed at green position is computed with the help adjacent original red or blue information, respectively. Thus, the missing red 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 values placed at blue or red position are reconstructed 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. 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 inter-channel correlation of the three primary colors and separating low- and high-frequency components in each pixel and replacing the high frequencies of the unknown component with the high frequencies of the Bayer component, which is known. In each pixel, the low-frequency component is kept unchanged since these components of the color channels are less correlated then the high-frequency components. For detailed explanation see [6].

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

This approach proposed by Chung [7] utilizes an inter-channel correlation, furthermore exploits the information about edges obtained directly from mosaic image for the decision, which chooses the pixels to be used for the to interpolation of missing value. Firstly, the information about the horizontal, vertical and both diagonal (45^{o}, 325^{o}) 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 one 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 computed for horizontal projection will be large; otherwise, the mask will be small. 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).



(a) (b) (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 (information obtained from mosaic image by using special proper convolution core) are assigned to four corresponding spectral-correlation terms to affect their impact on reconstructed green value; for more details refer to [7]. For computation of four weights, one central and eight adjacent pixels are used. After green channel reconstruction, the remaining red and blue channels are estimated. The red missing value at blue position is computed from four adjacent spectral-correlation terms with four proper weights; in this case the information from the heterogeneity projection is not exploited, therefore, the calculation considers only one case shown in Figure 5. The blue missing values at the red positions are computed by following the same strategy.

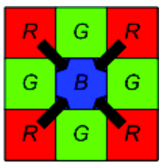


Figure 5: *The computation of missing red value on the blue position.*

3. The preparation for testing

This chapter deals with the definition of synthetic testing 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

50x50 points. The rows from the top contains horizontal, diagonal 45°, diagonal 135° and vertical direction of spatial frequency component. The columns from the right contains concrete spatial frequencies; $f_s/2$, $f_s/4$, $f_s/6$, $f_s/8$, where f_s denotes the sampling raster frequency. In other words, the boxes on the right side contain the highest frequency 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*, see in Figure 6 (d). The frequencies generated by aliasing must be eliminated to enable the correct reconstruction.

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

$$CMSE=\frac{1}{3MN}\left(\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}\sum_{ch\in Ch}\left(I_{Orig}^{ch}(i,j)-I_{Rec}^{ch}(i,j)\right)^2\right), \tag{2}$$

where M, N denotes the sizes of image in vertical, horizontal direction, respectively. $Ch=\{r,g,b\}$; $I^r_{Orig}(i,j)$, $I^g_{Orig}(i,j)$, $I^b_{Orig}(i,j)$ denotes the three color components of the pixel at location (i,j) in the original image; $I^r_{Rec}(i,j)$, $I^g_{Rec}(i,j)$, $I^b_{Rec}(i,j)$ 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 luminance and the contrast in the image, but the structural information is independent on it. The change of all three characteristics of image, luminance, contrast and structural information can affect the quality of final image. Despite this, the perceived quality of image is most dependent on the structural information. For more details, see in [8]. The *SSIM* is computed as follows

$$SSIM(x_p,y_p)=\frac{\left(2\mu_{x_p}\mu_{y_p}+C_1\right)\left(2\sigma_{x_py_p}+C_2\right)}{\left(\mu_{x_p}^2+\mu_{y_p}^2+C_1\right)\left(\sigma_{x_p}^2+\sigma_{y_p}^2+C_2\right)}, \tag{3}$$

where x_p, y_p 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]. $C_1=(K_1 \cdot L)^2$, $C_2=(K_2 \cdot L)^2$, where $K_1=0{,}01$, $K_2=0{,}03$ and L denotes the dynamic range of the pixel values; in terms of 8bit grayscale image $L=255$.

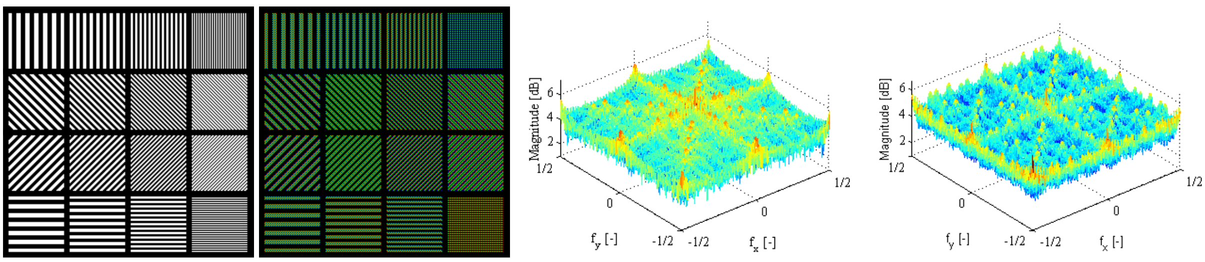
For overall image quality, the *mean SSIM* index is used

$$MSSIM(I_{Orig},I_{Rec})=\frac{1}{M}\sum_{p=1}^{M}SSIM(x_p,y_p), \tag{4}$$

where I_{Orig} and I_{Rec} is 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 tests 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.



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

Figure 6: *Testing image,*
(a) *monochromatic testing image,*
(b) *mosaic image of testing image,*
(c) *modulo spectra of testing image,*
(d) *modulo spectra of mosaic image, the effect of aliasing is evident.*

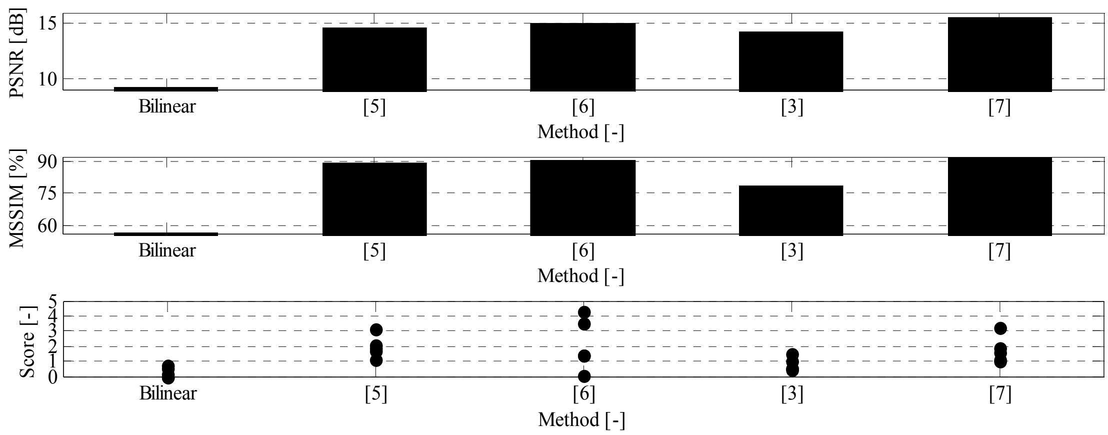


Figure 7: *PSNR, MSSIM and Subjective results.*

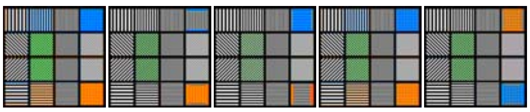
4. Results

To evaluate the algorithm performance itself 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.

Method	Bilinear	[5]	[6]	[3]	[7]
PSNR [dB]	9,16	14,59	14,99	14,26	**15,51**
SSIM [%]	56,18	89,17	90,64	78,35	**91,61**

Table 1: *PSNR and SSIM results.*



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

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].

5. Conclusion

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 ones (right top and bottom box). The frequency contained in these two boxes was higher than the sampling frequencies of red and blue Bayer sampling rasters. In the case of the horizontal or vertical directions, there is no information about stripes in red or blue sub-sampled channels, in other words, the aliasing effect is presented in both 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 $f_s/6$ was reconstructed with the green color artifacts. Alleyson's algorithm [3] gives almost identical results in the case of reconstructed boxes as the bilinear interpolation in spite of Alleyson'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 frequencies at $f_s/2$ in contrast to algorithms proposed by Hirakawa [5] and Menon [6]. Each of them was able to reconstruct the finest details only in one direction, see in Figure 8 (b), (c). The diagonal frequency at $f_s/6$ was not reconstructed even by any algorithm.

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 Cotereau and Matthieu Bleichner for performing the subjective tests.

7. References

[1] Bayer, B. E., Color imaging array, U.S. Patent 3 971 065, 1976
[2] Askey, P ., "Kodak DCS620x review", Online: http://www.dpreview.com/reviews/kodakdcs620x/
[3] Alleysson, D., Susstrunk, S., Herault, J., "Linear Demosaicing Inspired by Human Visual System", In *IEEE Transactions on Image Processing*, Vol. 14, No. 4, April 2005, pp. 439–449.
[4] Su, Ch.-Y., Kao, W.-Ch., " Effective Demosaicing Using Subband Correlation", In *IEEE Transaction on Consumer Electronics*, Vol. 14, No. 1, February 2009, pp. 199-204.
[5] Hirakawa, K., Parks, T. W., "Adaptive Homogeneity-Directed Demosaicing Algorithm", In *IEEE Transaction on Image Processing*, Vol. 14, No. 3, 2005, pp. 360–368.
[6] Menon, D., Andriani, S., Calvagno, G., "Demosaicing With Directional Filtering and a posteriori Decision", In *IEEE Transactions on Image Processing*, Vol. 16, No. 1, 2007, pp. 132–141.
[7] Chung, K. L., et al., "Demosaicing of Color Filter Array Captured Images Using Gradient Edge Detection Masks and

Adaptive Heterogeneity-P rojection", In *IEEE Transaction on Image Processing*, Vol. 17, No. 12, December 2008, pp. 2356–2367.

[8] Wang, Z., et al., "Image quality assessment: From error visibility to structural si milarity", In *IEEE Transaction on Image Processing*, Vol. 13, No. 4, April 2004, pp. 600-612.