

Zhang–Wu Directional LMMSE Image Demosaicking

Pascal Getreuer

CMLA, ENS Cachan (pascal.getreuer@cmla.ens-cachan.fr)

Communicated by Antoni Buades *Demo edited by* Pascal Getreuer

Abstract

Most digital cameras capture samples through a color filter array. At every pixel location, the camera observes one of either the red, green, or blue component. *Demosaicking* (or demosaicing) is the problem of using this incomplete information to estimate all three color components at every pixel. Zhang and Wu proposed an effective solution to this problem in “Color Demosaicking via Directional Linear Minimum Mean-Square-Error Estimation” [4].

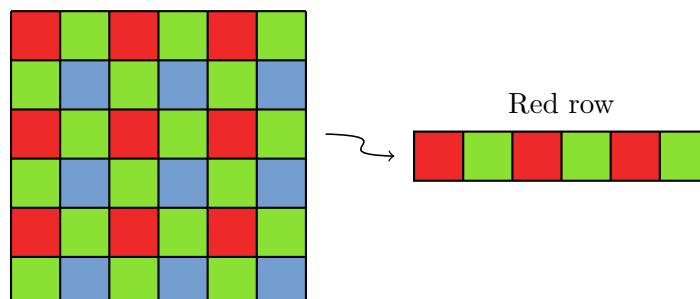
Source Code

ANSI C source code to produce the same results as the demo is accessible on the article web page https://doi.org/10.5201/ipol.2011.g_zwld. Future software releases and updates will be posted at <http://dev.ipol.im/~getreuer/code>.

Keywords:

1 The Bayer Pattern

The most commonly used color filter array is the Bayer pattern [1].



The Bayer pattern.

Green pixel locations are arranged in a quincunx lattice and cover half the array. The red and blue pixel locations are spaced uniformly every two pixels and each cover a quarter of the array. The pattern alternates between “red rows” and “blue rows.” In a red row the pattern is R, G, R, G, \dots and in a blue row it is G, B, G, B, \dots

2 LMMSE Estimation

We review linear minimum mean-square error (LMMSE) estimation for scalar variables. Suppose X and Y are scalar random variables, and we wish to use the (assumed known) value of Y to estimate the (unknown) value of X . The minimum mean-square error (MMSE) estimate is

$$\hat{X}_{\text{MMSE}} := \arg \min_F \mathbb{E}[(F(Y) - X)^2], \quad (1)$$

where $F(Y)$ is any function of Y . Under mild conditions the MMSE estimator is $E[X|Y]$, the conditional expectation of X given Y .

However, in many applications only limited information is available about X and Y . LMMSE estimation is useful in these cases since it only requires second-order statistics. The LMMSE estimator is the best estimate of the form

$$\hat{X}_{\text{LMMSE}} := a(Y - \mathbb{E}Y) + b. \quad (2)$$

The optimal coefficients a and b minimize the mean-square error

$$\min_{a,b} \mathbb{E}[\left((a(Y - \mathbb{E}Y) + b) - X\right)^2]. \quad (3)$$

Differentiating with respect to a and b yields

$$\begin{aligned} 0 &= \frac{\partial}{\partial a} \mathbb{E}[(a(Y - \mathbb{E}Y) + b - X)^2] \\ &= 2\mathbb{E}[(a(Y - \mathbb{E}Y) + b - X)(Y - \mathbb{E}Y)] && \Rightarrow a = \frac{\text{Cov}(X, Y)}{\text{Var } Y}, \\ &= 2(a \text{Var } Y - \text{Cov}(X, Y)) \end{aligned} \quad (4)$$

$$\begin{aligned} 0 &= \frac{\partial}{\partial b} \mathbb{E}[(a(Y - \mathbb{E}Y) + b - X)^2] \\ &= 2(b - \mathbb{E}X) && \Rightarrow b = \mathbb{E}X. \end{aligned} \quad (5)$$

So the LMMSE estimate is

$$\hat{X}_{\text{LMMSE}} = \frac{\text{Cov}(X, Y)}{\text{Var } Y}(Y - \mathbb{E}Y) + \mathbb{E}X, \quad (6)$$

and its mean-square error is

$$\begin{aligned} \mathbb{E}[(\hat{X}_{\text{LMMSE}} - X)^2] &= a^2 \text{Var } Y - 2a \text{Cov}(X, Y) + \text{Var } X \\ &= \text{Var } X - \frac{\text{Cov}(X, Y)^2}{\text{Var } Y}. \end{aligned} \quad (7)$$

To avoid division by zero, a small value 0.1 is added (relative to the intensity range $[0, 255]$) in the denominators to $\text{Var } Y$.

3 LMMSE Denoising

Zhang and Wu [4] describe the following simple method based on LMMSE estimation for signal denoising. Suppose that a smooth one-dimensional signal u_n is observed with independent zero-mean additive noise,

$$f_n = u_n + \epsilon_n. \quad (8)$$

Assuming the noise ϵ has limited low frequency content, an effective estimate is $s = f * h$ where h is a lowpass filter. However, lowpass filtering is a balance between noise reduction and blurring the true signal content, which is especially problematic if the noise is spatially varying.

Since u is assumed to be smooth, the lowpass filtered signal s should be closer than f to u . Statistics of s and $(f - s)$ can be used to estimate statistics of u and ϵ :

$$\mu = \frac{1}{2M+1} \sum_{m=-M}^M s_{n+m} \approx \mathbb{E}u_n \quad (9)$$

$$\sigma_u^2 = \frac{1}{2M+1} \sum_{m=-M}^M (s_{n+m} - \mu)^2 \approx \text{Var } u_n \quad (10)$$

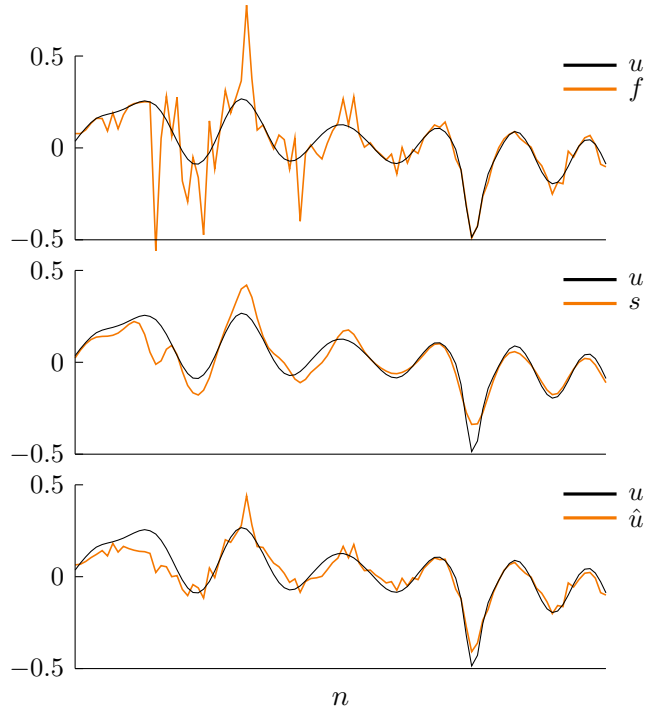
$$\sigma_\epsilon^2 = \frac{1}{2M+1} \sum_{m=-M}^M (s_{n+m} - f_{n+m})^2 \approx \text{Var } \epsilon_n, \quad (11)$$

where M is a fixed window radius. From these statistics we obtain the LMMSE denoised estimate

$$\hat{u}_n = \mu + \frac{\sigma_u^2}{\sigma_u^2 + \sigma_\epsilon^2} (f_n - \mu), \quad (12)$$

and an estimate of the error

$$\mathbb{E}[(\hat{u}_n - u_n)^2] \approx \sigma_u^2 - \frac{(\sigma_u^2)^2}{\sigma_u^2 + \sigma_\epsilon^2}. \quad (13)$$



LMMSE denoising experiment.

The figure shows an LMMSE denoising experiment on a sequence of 100 samples with spatially varying non-Gaussian noise. The top shows the exact signal u and the observed signal f . The middle shows the lowpass filtered signal $s = f * h$. The bottom shows the LMMSE denoised signal.

4 Algorithm

For typical natural images, the color channels are strongly correlated. Most demosaicking methods combine information across color channels to take advantage of this correlation. Let R_n , G_n , B_n denote the color components of the n th pixel of a row of the image and define the *primary difference signals*

$$\Phi_{g,r}(n) = G_n - R_n, \quad \Phi_{g,b}(n) = G_n - B_n. \quad (14)$$

The algorithm outline is

1. The primary difference signals are estimated in the horizontal direction and denoised using LMMSE estimation.
2. Similarly, the primary difference signals are estimated and denoised in the vertical direction.
3. The two denoised directional estimates are fused to obtain a full resolution green channel.
4. The primary difference signals are interpolated and then subtracted from green to obtain the red and blue channels.

We discuss each step in detail below.

4.1 Directional Estimates

We focus on the estimation of a red row; the process is analogous for a blue row. For each pixel in a red row, either G_n or R_n but not both are known. Zhang and Wu suggest based on empirical evidence that it is accurate to assume that the primary difference signals are smooth. Second-order Laplacian interpolation is used to obtain the missing samples,

$$\hat{G}_n = \frac{1}{2}(G_{n-1} + G_{n+1}) - \frac{1}{4}(R_{n-2} - 2R_n + R_{n+2}), \quad (15)$$

$$\hat{R}_n = \frac{1}{2}(R_{n-1} + R_{n+1}) - \frac{1}{4}(G_{n-2} - 2G_n + G_{n+2}), \quad (16)$$

which provides an estimate of the difference signal,

$$\hat{\Phi}_{g,r}^h = \begin{cases} \hat{G}_n - R_n & n \text{ is a red pixel,} \\ G_n - \hat{R}_n & n \text{ is a green pixel.} \end{cases} \quad (17)$$

The estimate is denoised as described in the section on LMMSE Denoising. The suggested parameters are a window radius $M = 4$ and an approximately Gaussian lowpass filter

$$h(z) = \frac{26}{128} + \frac{23}{128}(z + z^{-1}) + \frac{15}{128}(z^2 + z^{-2}) + \frac{9}{128}(z^3 + z^{-3}) + \frac{4}{128}(z^4 + z^{-4}). \quad (18)$$

The LMMSE denoising error estimate is also noted since it will be used in the following step.

4.2 Fusion of the Directional Interpolations

The previous steps produced horizontally and vertically directed estimates of the difference signals $\Phi_{g,r}$ and $\Phi_{g,b}$. The directional estimates are now fused to yield the final estimate of the difference signals.

Since the objective is to interpolate the green channel, fusion is only performed at red and blue pixel locations (where the green component is unknown). LMMSE estimation is applied again here to develop a method for fusion.

Let h denote the horizontally directed estimate, ϵ_h the estimation error, and σ_h^2 the LMMSE denoising error estimate, and similarly let v , ϵ_v , and σ_v^2 denote the analogous quantities for the vertically directed estimate. The fused estimate is

$$w = (1 - \lambda)h + \lambda v \quad (19)$$

where λ is selected to minimize the estimated mean-square error,

$$\begin{aligned} 0 &= \frac{\partial}{\partial \lambda} \mathbb{E} [((1 - \lambda)\epsilon_h + \lambda\epsilon_v)^2] \\ &= 2\mathbb{E} [-(1 - \lambda)\epsilon_h^2 + \lambda\epsilon_v^2 + (1 - 2\lambda)\epsilon_h\epsilon_v] \Rightarrow \lambda = \frac{\sigma_h^2}{\sigma_h^2 + \sigma_v^2}. \\ &\approx 2(1 - \lambda)\sigma_h^2 + 2\lambda\sigma_v^2, \end{aligned} \quad (20)$$

In the last line, it is assumed that ϵ_h and ϵ_v are uncorrelated. Zhang and Wu claim that this is usually true, particularly on edges and texture where accurate fusion is important. To avoid division by zero, a small value 0.2 is added (relative to the intensity range $[0, 255]$) in the denominator when computing λ .

With the fused estimates of $\Phi_{g,r}$ and $\Phi_{g,b}$, the green component G_n is now available at every pixel. Demosaicking of the green channel is now complete.

4.3 Obtaining the Red and Blue Channels

Based on original Bayer samples and the fully interpolated green channel, primary difference signal $\Phi_{g,r} = G - R$ is known at red locations and $\Phi_{g,b} = G - B$ is known at blue locations.

The primary difference signals are interpolated in two steps. First, $\Phi_{g,r}$ is interpolated at blue locations by averaging its four diagonal neighbors, all of which are red locations,

$$\Phi_{g,r}(i, j) = \frac{1}{4} (\Phi_{g,r}(i - 1, j - 1) + \Phi_{g,r}(i + 1, j - 1) + \Phi_{g,r}(i - 1, j + 1) + \Phi_{g,r}(i + 1, j + 1)). \quad (21)$$

Primary difference signal $\Phi_{g,b}$ is similarly interpolated at red locations. Second, $\Phi_{g,r}$ and $\Phi_{g,b}$ are interpolated at green locations by averaging their four axial neighbors,

$$\Phi_{g,r}(i, j) = \frac{1}{4} (\Phi_{g,r}(i - 1, j) + \Phi_{g,r}(i + 1, j) + \Phi_{g,r}(i, j - 1) + \Phi_{g,r}(i, j + 1)). \quad (22)$$

Finally, the red and blue channels are obtained as $R = G - \Phi_{g,r}$ and $B = G - \Phi_{g,b}$.

5 Examples

In the following examples, an exact image is mosaicked over the Bayer pattern and restored with Zhang–Wu demosaicking and other demosaicking methods. The difference between the exact and demosaicked images is measured with the peak signal to noise ratio (PSNR) over all three color channels. Consider two color images A and B as vectors in \mathbb{R}^{3N} , where N is the number of pixels and the components with values in $[0, 255]$ represent RGB intensities, then the PSNR is

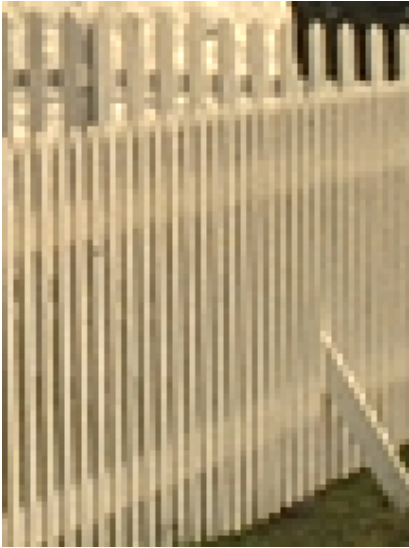
$$PSNR(A, B) := 10 \log_{10} \frac{255^2}{\frac{1}{3N} \|A - B\|_2^2}. \quad (23)$$

A typical artifact of demosaicking methods is the appearance of “zipper patterns,” a sequence of adjacent pixels of alternating brightness. It is also common to see false colors and desaturation. The test images here are selected to highlight these problems.

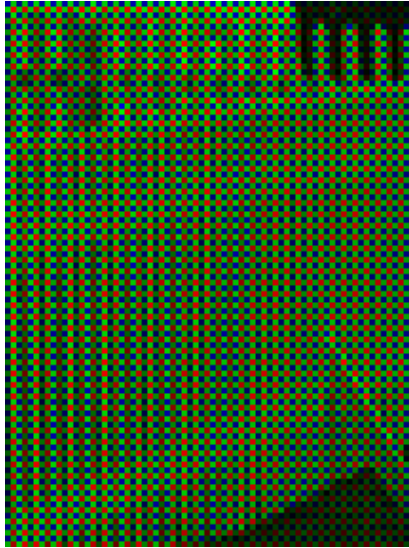
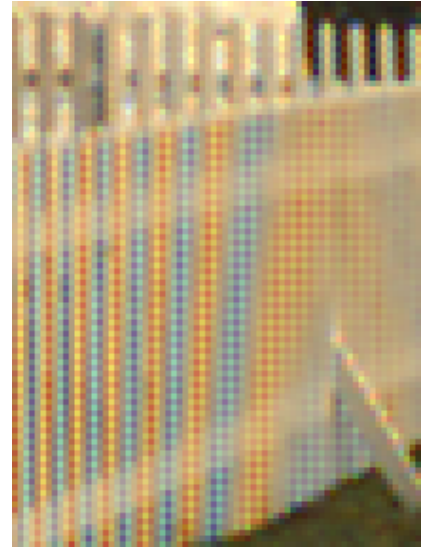
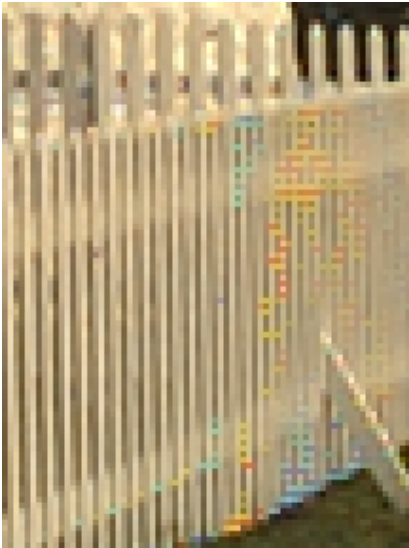
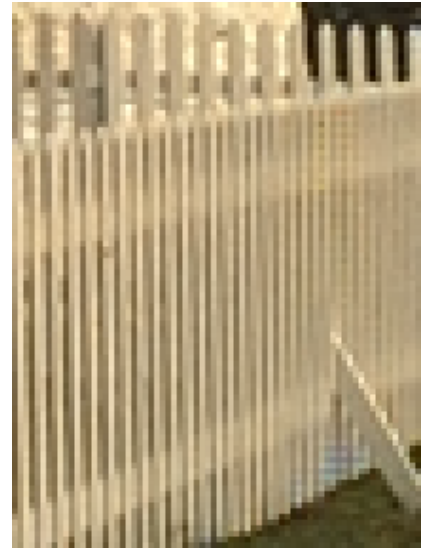
To avoid effects from the image boundaries, a margin of 5 pixels has been removed before computing PSNR and in the displayed results. All images are enlarged by nearest neighbor interpolation to show individual pixels more clearly.

Zhang–Wu is especially competitive on images with aliasing. The fence image is increasingly difficult toward the right side. The red and blue channels are sampled well below the Nyquist rate needed to sample the fence slats.

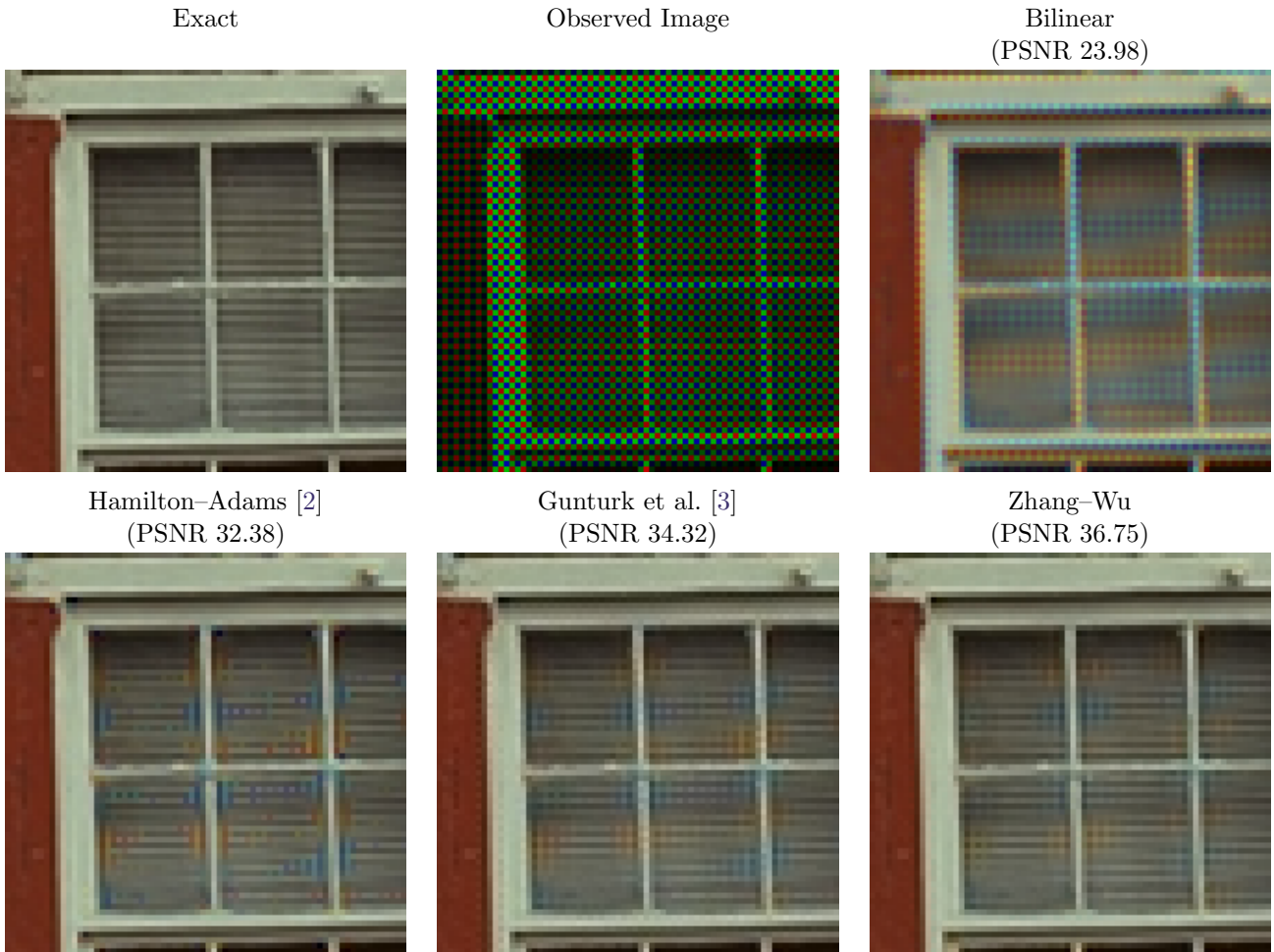
Exact



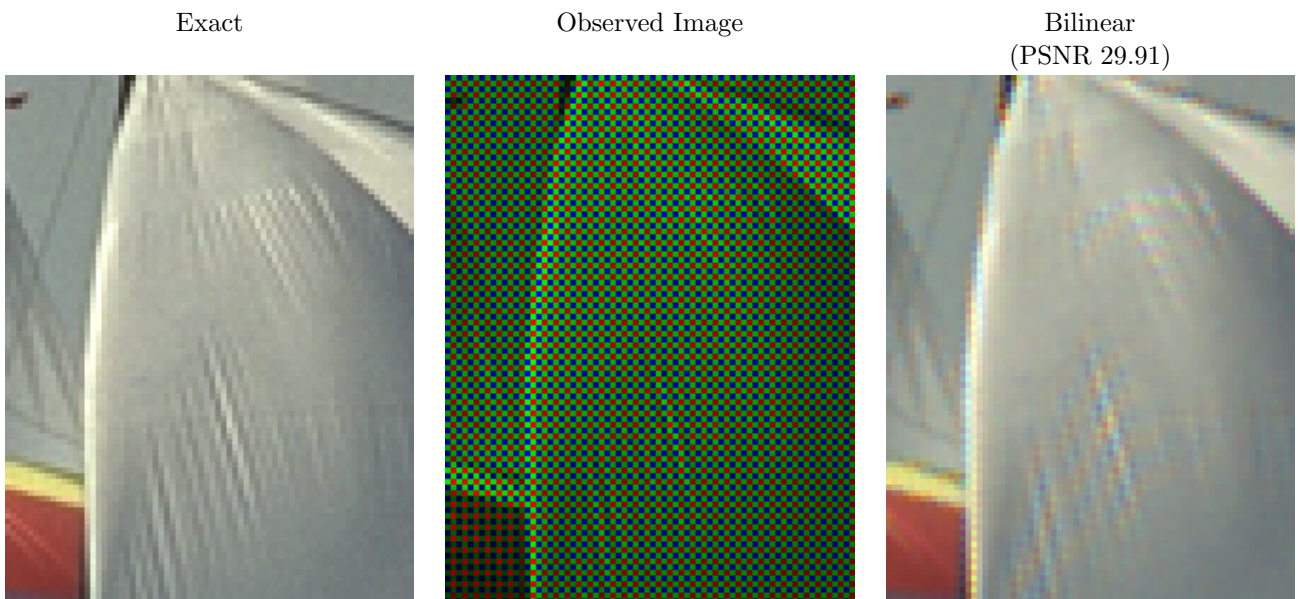
Observed Image

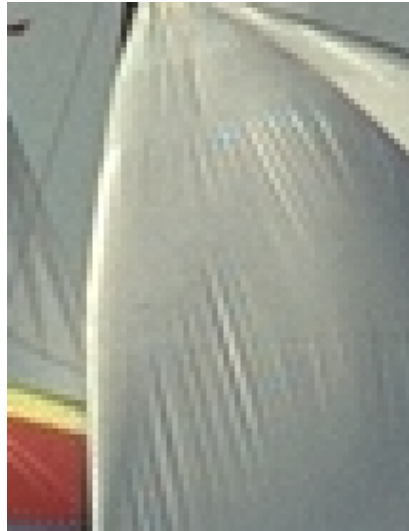
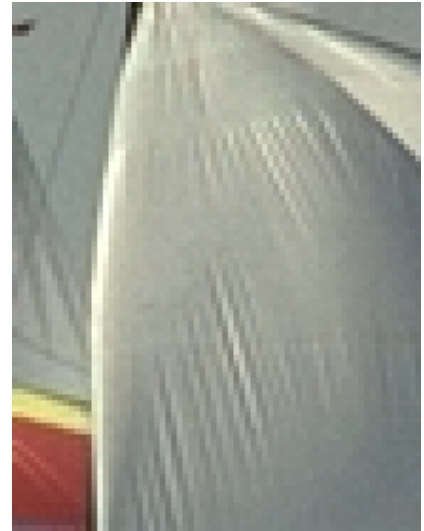
Bilinear
(PSNR 20.42)Hamilton–Adams [2]
(PSNR 30.19)Gunturk et al. [3]
(PSNR 32.35)Zhang–Wu
(PSNR 38.29)

In the window image, the window blinds violate the Nyquist rate when subsampled by the CFA. While in the previous example Zhang-Wu performs well on the aliased pattern of the fence, Zhang-Wu is less successful on the abrupt T-junctions formed between the slats and the window frame.



This example is on a mostly gray image of a sail. Demosaicking such an image is useful to compare to what degree the methods create false colors.



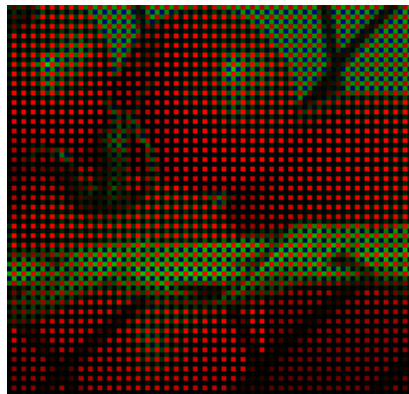
Hamilton–Adams [2]
(PSNR 37.45)Gunturk et al. [3]
(PSNR 40.95)Zhang–Wu
(PSNR 41.65)

The DLMMSSE Zhang–Wu scheme described here is known to work very well on the Kodak dataset. However, as shown by the same authors [5], the choice of dataset significantly affects the outcome when evaluating demosaicking algorithms. The Kodak dataset is somewhat gray and blurred, and Zhang–Wu was designed for these kinds of images. Zhang–Wu is less effective for example on the more colorful IMAX and McMaster datasets. For example, significant zipper artifacts appear on McMaster image 9:

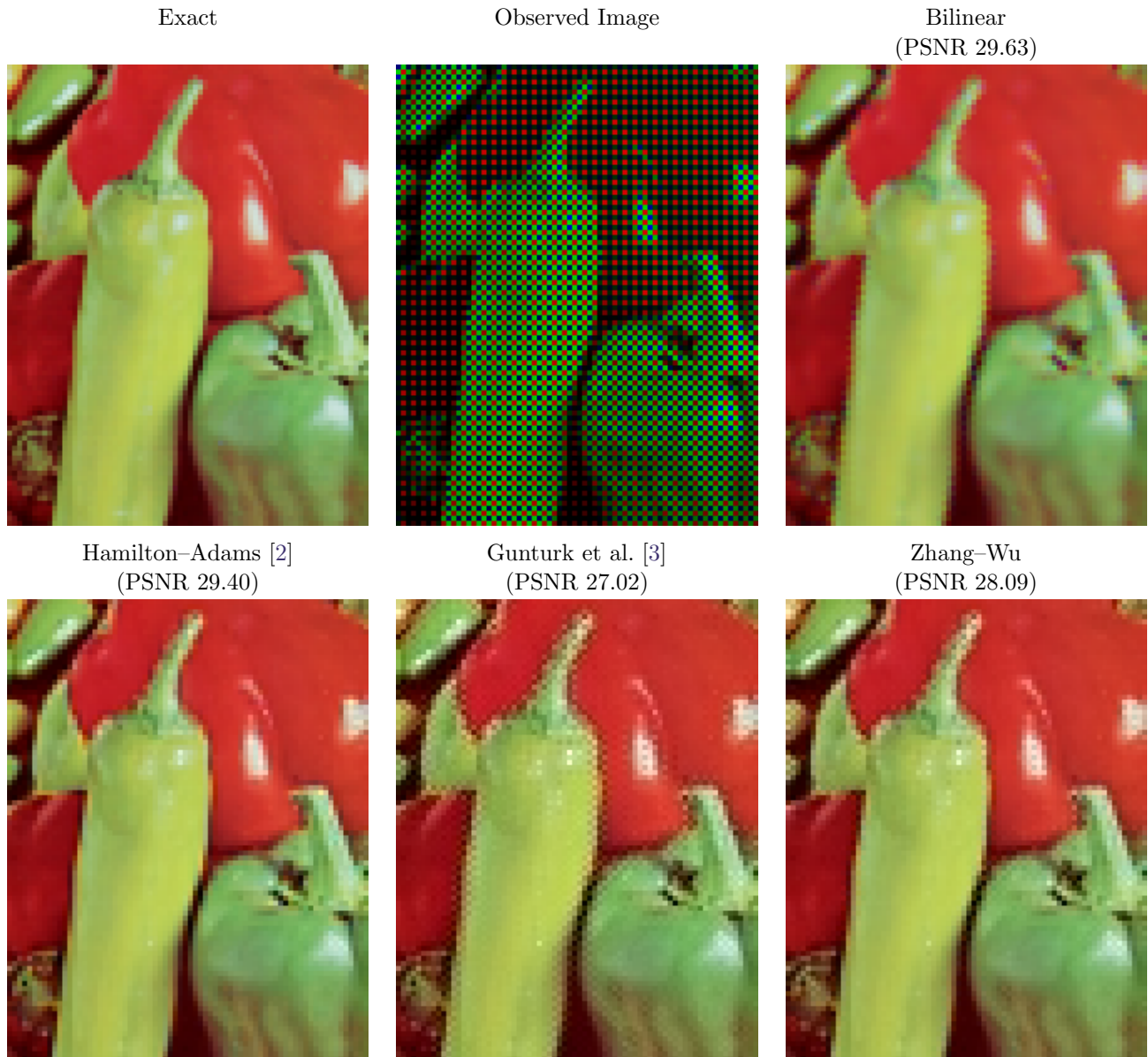
Exact



Observed Image

Bilinear
(PSNR 31.08)Hamilton–Adams [2]
(PSNR 30.27)Gunturk et al. [3]
(PSNR 27.11)Zhang–Wu
(PSNR 28.34)

As another demonstration, the last example is on the peppers image. This image has significant color discontinuities at the edges, where the channel correlation assumption is false.



Acknowledgments

This material is based upon work supported by the National Science Foundation under Award No. DMS-1004694. Work partially supported by the Office of Naval Research under grant N00014-97-1-0839 and by the European Research Council, advanced grant “Twelve labours.”

Image Credits



Kodak Image Suite, image 1 (<http://r0k.us/graphics/kodak/>)



Kodak Image Suite, image 7 (<http://r0k.us/graphics/kodak/>)



Kodak Image Suite, image 9 (<http://r0k.us/graphics/kodak/>)



Kodak Image Suite, image 19 (<http://r0k.us/graphics/kodak/>)



McMaster Dataset, image 9 (http://www4.comp.polyu.edu.hk/~cslzhang/CDM_Dataset.htm)



Peppers standard test image

References

- [1] B. E. Bayer, “Color imaging array,” U.S. Patent 3971065, 1976.
- [2] J. F. Hamilton, Jr. and J. E. Adams, Jr., “Adaptive color plan interpolation in single sensor color electronic camera,” U.S. Patent 5629734, 1997.
- [3] B. K. Gunturk, Y. Altunbasak, and R. M. Mersereau, “Color plane interpolation using alternating projections,” *IEEE Transactions on Image Processing*, vol. 11, no. 9, pp. 997–1013, 2002. <http://dx.doi.org/10.1109/TIP.2002.801121>
- [4] Lei Zhang and Xiaolin Wu, “Color demosaicking via directional linear minimum mean square-error estimation,” *IEEE Transactions on Image Processing*, vol. 14, no. 12, pp. 2167–2178, 2005. <http://dx.doi.org/10.1109/TIP.2005.857260>
- [5] Lei Zhang, Xiaolin Wu, Antoni Buades, and Xin Li, “Color Demosaicking by Local Directional Interpolation and Non-local Adaptive Thresholding,” *Journal of Electronic Imaging* 20(2), 023016, 2011. <http://dx.doi.org/10.1117/1.3600632>