

# LOW-COMPLEXITY LINEAR DEMOSAICING USING JOINT SPATIAL-CHROMATIC IMAGE STATISTICS

*Javier Portilla*

Visual Information Processing Group  
Dept. of Comp. Science and Artif. Intell.  
Universidad de Granada, Spain  
javier@decsai.ugr.es

*Deitze Otaduy, Carlos Dorrnsoro*

Centro de Invest. y Desarrollo de la Armada  
Madrid, Spain  
{dotaduyp, cdorrnsoro}@oc.mde.es

*We present an efficient Linear Minimum Mean Square Error (LMMSE) method for reconstructing full color images from single sensor Color Filter Array (CFA) data. We use a representative set of full color images to estimate the joint spatial-chromatic covariance among pixel color components. Then, we derive from it a set of joint color-space, small linear kernels which predict the missing color samples as linear combinations of their neighbor observed samples. The color arrangement of the local mosaic varies with the window's location, and this results into a different predictor for every local mosaic and color sample. As an extension, we include blur and noise in the training process, obtaining localized mosaic-constrained Wiener estimators that partially compensate for these degradations. We show that this simple method provides an excellent trade-off between performance and computational cost.*

## 1. INTRODUCTION

Single sensor color digital cameras based on Color Filter Array technology [1] capture only one color component at each spatial location. Therefore, space-color interpolation is required for reconstructing a full-color image. This process is termed *demosaicing*. In the last two decades, many different approaches have been proposed to solve this problem (see [2] for an up-to-date review). Most demosaicing methods follow a heuristic approach, with varying results in terms of performance and computational cost. Some are iterative and demand heavy computation (e.g., [3, 4]). These often provide good insight about the nature of the problem, and they may also be suitable for off-line processing of certain images. However, as pointed out in [2], recent advances in digital image capture are making those computationally demanding approaches less and less suitable for most practical situations. On the one hand, the fast increase in spa-

tial resolution of commercial CCD sensors has not had an optical counterpart, and, thus, the ratio between the optical cut-off frequencies and the sensor sampling frequencies is quickly falling. As a consequence, there is a progressive reduction in what used to be the signature of CFA technology: color aliasing. On the other hand, the increasing speed of embedded digital processing is being overrun by the larger amount of data to be processed, because of the increase in both spatial resolution and number of bits per pixel in images. Thus, keeping low the computational cost of demosaicing is still very important, and, for most practical situations low-complexity algorithms are preferable to time-consuming sophisticated algorithms.

Linear formation models are among the most comprehensive for the demosaicing problem (e.g., [5, 6, 7]). They are usually applied using LMMSE (Linear Minimum Mean Square Error) for the estimation of the full color image, typically under a Bayesian frame. As a natural extension of the image sampling model, these methods have the advantage of incorporating the effect of the Point Spread Function of camera optics, which can be partially compensated later on within the demosaicing process. Surprisingly enough, the linear image formation model has been little exploited in practical algorithms. The methods referred above suffer from a variety of problems that make them not suitable for being embedded in digital cameras. In [7] a sound mathematical approach is followed, to end up embracing assumptions that severely reduce the potential of the method, like the separability of the spectral (chromatic) and spatial correlation. In [6, 7] a global estimation approach is followed, and the main computations are done in the Fourier domain, instead of using the more efficient solution of small kernels in the spatial domain. In addition, the authors do not report any result that could be compared to other methods. Ref. [5] presents an insightful Bayesian approach to the same problem. However, that work does not provide a practical algorithm. Ref. [8] proposes an efficient method based on signal correlation, but it does not apply an LMMSE optimization.

---

This work is funded by the Ministerio de Ciencia y Tecnología (grant TIC2003-1504) and the Ministerio de Defensa (2050-559 grant contract), Spain. JP is supported by the "Ramón y Cajal" Program.

In this work we have followed an alternative LMMSE demosaicing approach, keeping the conceptual simplicity of the linear image formation model, but avoiding the drawbacks referred above. The key strategies to achieve this goal have been the following. First, instead of representing every component in the image formation model and trying to solve analytically the resulting equations, we have focused on the property that the missing/degraded color samples are estimated through linear combinations of the observations (because of applying a MSE optimization to a linear model). In addition, instead of using a prior for the image (e.g., as in [5, 6]), we demonstrate that good quality interpolation can be achieved from the joint spatial-chromatic correlation (JSCC) obtained from a small set of typical images. Second, instead of posing constraints on the image statistics, we only pose a constraint on the linear estimators themselves, to make them suitable for practical computation: a small  $N \times N$  spatial support. We demonstrate that this constraint, using  $N \geq 9$ , does not affect significantly the overall performance, whereas it allows us to perform calculations efficiently in the spatial domain. Finally, we decouple the estimation problem into a set of different prediction problems. These result from the different local mosaics framed when shifting the estimation window over the CFA.

## 2. SPATIAL-CHROMATIC IMAGE STATISTICS

Statistical modelling is a key issue conditioning the quality of image processing tasks. In this case we are only concerned with signal correlation, as we follow a LMMSE estimation approach. We can write the JSCC data as:

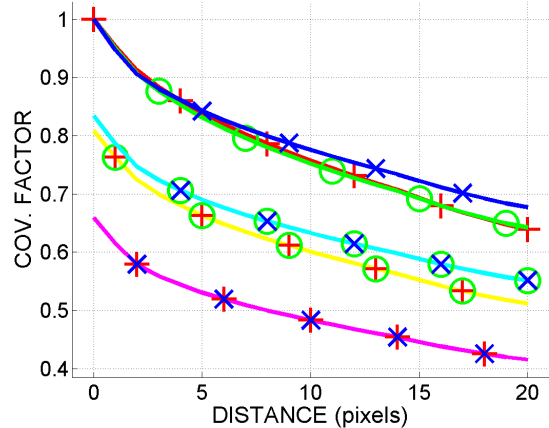
$$C_{xx}(c_1, c_2, \Delta_n, \Delta_m) = \mathbb{E}\{x(c_1, n, m)x(c_2, n + \Delta_n, m + \Delta_m)\}$$

where  $(n, m)$  are the spatial coordinates, and  $(c_1, c_2)$  are a pair of colors. Joint spatial-chromatic image statistics have been studied for a long time (see, e.g. [9]). For the examples shown in this work we have assumed a Bayer CFA mosaic [1], with red, green and blue filters, although the described procedure can be applied to any periodic CFA structure. In Fig. 1 we show an average radial profile of the estimated JSCC (normalized), for all six color pairs, obtained using the 40 images of [10]. As we demonstrate in Sections 3.2 and 3.3 the linear estimation depends exclusively on the JSCC, when there is no blur/noise, and on both the JSCC and the cross-JSCC between the clean and the degraded image, when there is blur/noise.

## 3. LINEAR MODEL AND LMMSE ESTIMATION

We use the following image formation model for the data captured in a CFA-based digital camera (see, e.g., [7])

$$\mathbf{y} = \mathbf{S}(\mathbf{H}\mathbf{x} + \mathbf{w}_0), \quad (1)$$



**Fig. 1.** Covariance factor, measured among color samples as a function of distance (up, overlapped: auto-covariance; down: cross-covariance). R, G and B are represented by "+", "O" and "x", respectively.

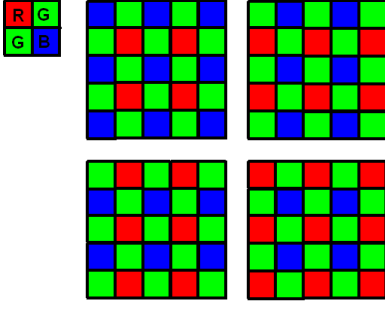
where  $\mathbf{H}$  represents the camera PSF (blurring),  $\mathbf{w}_0$  is independent zero-mean white noise, and  $\mathbf{S}$  the color sampling carried out by the CFA array. Being  $P$  the number of image pixels,  $\mathbf{x}$  has  $3P$  elements and  $\mathbf{y}$  has only  $P$  elements. It is more convenient to use, instead  $\mathbf{y} = \mathbf{S}\mathbf{H}\mathbf{x} + \mathbf{w}$ , where note that  $\mathbf{w} = \mathbf{S}\mathbf{w}_0$  is still white. The unconstrained MMSE solution of this linear system is

$$\hat{\mathbf{x}}(\mathbf{y}) = \arg \min_{\mathbf{x}} \|\mathbf{y} - \mathbf{S}\mathbf{H}\mathbf{x}\|, \quad (2)$$

which is a linear function of  $\mathbf{y}$ . However, as the involved vectors and matrices are large, the direct solution of this equation is computationally costly.

### 3.1. Tiling the CFA mosaic into local mosaics

As mentioned before, some authors have imposed constraints on the image statistics to solve Eq. 2 globally in the Fourier domain [6, 7]. Instead, we force the linear estimators to have a limited spatial support, considering only the neighbors around the missing/degraded sample. By forcing a local estimation window, we allow for a much simpler solution and ensure a modest computational cost. As a consequence of localizing the estimation window we have now different local mosaic structures for different locations on the CFA array. The CFA mosaic is made of the repetition of a basic sampling pattern, with  $J$  color components. If we take an odd-size square window around every position, then there are  $J$  different possibilities of local mosaics in that window, as shown in Fig. 2. We distinguish between these  $J$  cases, obtaining different estimators for each local mosaic. Note that by decoupling the problem this way we are not affecting the global LMMSE optimality under the spatial support constraint. Although most previous algorithms use different local interpolators for different locations on the



**Fig. 2.** Basic sampling pattern for the Bayer mosaic (up left), and the resulting  $J = 4$  local mosaics, for  $N = 5$ .

CFA, as far as we know none of them derive their estimation from the image statistics.

### 3.2. Demosaicing

Let  $\mathbf{x}_i^j$  be the vector of the unknown color samples of the  $(i, j)$  pixel, the  $i$ th pixel among the  $L_j$  pixels having a type- $j$  local mosaic around them. We build the matrices  $\mathbf{X}_j = (\mathbf{x}_1^j \dots \mathbf{x}_{L_j}^j)$ ,  $j = 1 \dots J$ . Each of these matrices has dimensions  $(N_c - 1) \times L_j$ , where  $N_c$  is the number of colors (e.g.,  $N_c = 3$ , and  $J = 4$ , for a Bayer CFA). Let  $\tilde{\mathbf{x}}_i^j$  be the vector containing all observed color samples in the window around pixel  $(i, j)$  ( $N^2$  elements). Now we build a matrix with all these neighborhoods, following the same order as before:  $\tilde{\mathbf{X}}_j = (\tilde{\mathbf{x}}_1^j \dots \tilde{\mathbf{x}}_{L_j}^j)$ . The linear estimation model tells us that every unknown color sample is predicted by a linear combination of their observed neighbors:

$$\mathbf{X}_j \simeq \widehat{\mathbf{X}}_j = \mathbf{D}_j \tilde{\mathbf{X}}_j, \quad (3)$$

and the LMMSE optimization implies that

$$\mathbf{D}_j = \arg \min_{\mathbf{H}} \mathbb{E} \left\{ \left\| \mathbf{X}_j - \mathbf{H} \tilde{\mathbf{X}}_j \right\|^2 \right\}. \quad (4)$$

Operating in the previous expression it yields:

$$\mathbf{D}_j = \mathbb{E} \{ \mathbf{X}_j \tilde{\mathbf{X}}_j^T \} (\mathbb{E} \{ \tilde{\mathbf{X}}_j \tilde{\mathbf{X}}_j^T \})^{-1} = \mathbf{C}_C^j (\mathbf{C}_M^j)^{-1}. \quad (5)$$

The elements of the matrices on the r.h.s. of Eq. 5 are joint spatial-chromatic correlation samples. Thus, we build the matrices  $\mathbf{C}_M^j$  and  $\mathbf{C}_C^j$  by (automatically) rearranging elements from  $C_{xx}$ .

### 3.3. Demosaicing with blur and noise compensation

We operate the same way as before, but now including all the  $N_c$  color samples in  $\mathbf{X}_j$ . We name  $\tilde{\mathbf{Y}}_j$  each of the  $J$  matrices made of the neighborhoods of the observed color samples, and obtain:

$$\mathbf{D}_j^b = \mathbb{E} \{ \mathbf{X}_j \tilde{\mathbf{Y}}_j^T \} (\mathbb{E} \{ \tilde{\mathbf{Y}}_j \tilde{\mathbf{Y}}_j^T \})^{-1} = \mathbf{C}_{xy}^j (\mathbf{C}_{yy}^j)^{-1}. \quad (6)$$

The estimation is then  $\widehat{\mathbf{X}}_j = \mathbf{D}_j^b \tilde{\mathbf{Y}}_j$ . For the training, we simulate the degraded "observations" through linear filtering and noise addition to the originals. Then we estimate the cross-JSCC (original-degraded)  $C_{xy}$  and the JSCC (degraded-degraded)  $C_{yy}$ . From  $C_{xy}$  and  $C_{yy}$  we rearrange terms as before to obtain  $\mathbf{C}_{xy}^j$  and  $\mathbf{C}_{yy}^j$ , respectively. Finally, using Eq. 6, we obtain the estimator  $\mathbf{D}_j^b$ . Note that when there is blur and/or noise to compensate for, a larger spatial support is typically required than for the pure interpolation case.

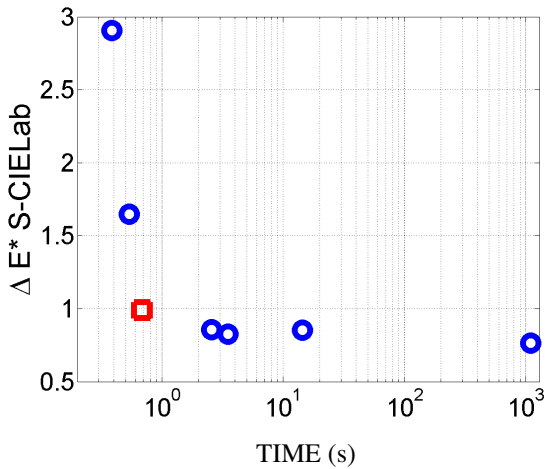
## 4. RESULTS AND DISCUSSION

From the images of [10]<sup>1</sup> we selected numbers 18, 31, 32, 33, 12, 34, 39, 15, 40, 16, 17 and 19, in  $512 \times 768$  format (the same images as in [11]). We estimated  $C_{xx}$  for each of these twelve images using the other eleven, and derived from  $C_{xx}$  the corresponding linear interpolators. Then, we simulated the Bayer mosaic and applied our demosaicing method. We have compared our method to some state-of-the-art techniques [3, 8, 4, 11]<sup>2</sup>, as well as to two reference methods (Bilinear and Nearest Neighbor). Table 1 shows average error using MSE for each color channel, and the S-CIELab metric  $\Delta E_{ab}^*$  [12] (assuming a 72 dpi display seen from 18 in.), as well as running times (using MATLAB(c) under a Pentium IV, 2.39 GHz). Because of its simplicity, our approach is very fast, but it still provides good performance. Fig. 3 displays these results on the  $\Delta E_{ab}^*$ -running time plane. Although we do not mean to equate running time to complexity, we think this graph is still significant. We see that the methods we compare to (circles) are grouped into two broad classes, namely "high speed, low quality" (nearest neighbor, bilinear: up left), and "medium-low speed, high quality" (the rest). Our method, for  $N = 9$  (square) lies on the imaginary corner where the two classes meet, sharing the speed range with the first group, but the quality with the second. Fig. 4(top) shows a visual comparison of the bilinear interpolation with our method ( $N = 9$ ). See how, despite having a comparable cost, ours preserves much better the details and it has much less color artifacts.

Finally, we obtained localized mosaic-constrained Wiener estimators that partially compensate for blur. We simulated the blur using an isotropic Gaussian filter with  $\sigma = 3$ , and then applied a 8-bit quantization. Fig. 4(bottom) shows a crop from a blurred image and its result ( $N = 15$ ). Note that usually deblurring is applied *after* the image has been demosaiced, a suboptimal procedure. A fast non-linear post-processing (see [2]) could be applied here to remove high-frequency artifacts.

<sup>1</sup>Although this set is not representative of what is captured by a modern digital camera, we have used it here for comparison purposes.

<sup>2</sup>Many thanks to Dr. Xin Li, for providing us Matlab(c) code implementing these methods.



**Fig. 3.** Average S-CIELab performance vs. running time. Circles, from left to right: Nearest Neighbor, Bilinear Interpolation, and those of references [8, 11, 3, 4], in that order. Square: our method, with  $N = 9$ .

### 5. CONCLUSIONS

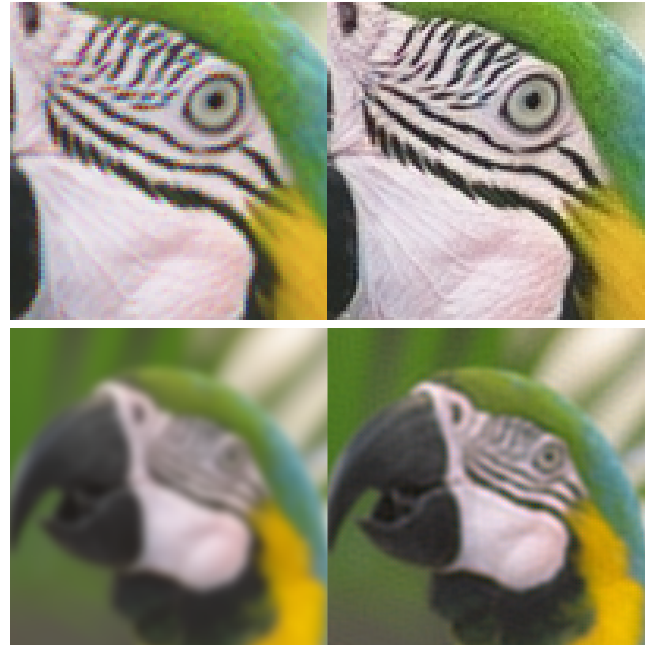
We have presented a fast demosaicing method based on joint spatial-chromatic image statistics and a linear formation model. Its performance is comparable to that of sophisticated, non-linear methods. This has been achieved by imposing a limited spatial support to the estimation kernels. We have demonstrated that a few examples of typical images suffices to derive interpolation kernels for other images. We have also shown how to include blur and noise compensation in the process. Finally, an interesting possibility to explore is the use of training sets adapted to specific types of images.

|             | MSE    |        |        | $\Delta E_{ab}^*$<br>S-CIELab | Time<br>(s) |      |
|-------------|--------|--------|--------|-------------------------------|-------------|------|
|             | Red    | Green  | Blue   |                               |             |      |
| Near. Neig. | 184.11 | 108.71 | 182.52 | 2.91                          | 0.38        |      |
| Bilinear    | 121.65 | 33.06  | 130.56 | 1.65                          | 0.48        |      |
| Pei [8]     | 13.53  | 4.41   | 14.74  | 0.86                          | 2.57        |      |
| Li [11]     | 7.87   | 3.57   | 8.48   | 0.83                          | 3.51        |      |
| Gunturk [3] | 8.08   | 3.70   | 10.57  | 0.85                          | 14.38       |      |
| Lu [4]      | 8.72   | 5.58   | 11.78  | 0.77                          | 1103.14     |      |
| Ours        | $N 7$  | 15.56  | 6.52   | 17.04                         | 1.04        | 0.59 |
|             | $N 9$  | 13.86  | 5.82   | 15.69                         | 0.99        | 0.65 |
|             | $N 11$ | 13.34  | 5.63   | 15.05                         | 0.98        | 0.74 |
|             | $N 15$ | 12.98  | 4.48   | 14.62                         | 0.97        | 0.96 |

**Table 1.** Average mean square and S-CIELab error for the 12 images, including running times. See text for details.

### 6. REFERENCES

[1] B. E. Bayer, "Color imaging array," *U.S. Patent 3971065*, 1976.



**Fig. 4.** **Up:** comparison between bilinear interpolation (left) and our method (right). **Down:** a blurred image (left), and our result of taking its Bayer mosaic, and demosaic it with blur and noise compensation (right).

[2] B. K. Gunturk, J. Glotzbach, Y. Altunbasak, R. W. Schafer, and R. M. Mersereau, "Demosaicking: Color filter array interpolation in single chip digital cameras," *IEEE Signal Processing Magazine*, September 2004.

[3] B. K. Gunturk and R. M. Mersereau, "Color plane interpolation using alternating projections," *IEEE Trans. on Image Proc.*, vol. 11, no. 9, pp. 997–1013, 2002.

[4] W. Lu and Y. P. Tan, "Color filter array demosaicing: new method and performance measures," *IEEE Trans. on Image Proc.*, vol. 12, no. 10, pp. 1194–1210, 2003.

[5] D. H. Brainard, "Bayesian method for reconstructing color images from trichromatic samples," in *Proc. IS&T 47th Annual Meeting*, Rochester NY, 1994, pp. 375–380.

[6] D. Taubman, "Generalized wiener reconstruction of images from colour sensor data using a scale invariant prior," *Proc. of Int. Conf. on Image Proc.*, vol. 3, pp. 801–804, 2000.

[7] H. J. Trussel and Robert E. Hartwig, "Mathematics for demosaicking," *IEEE Trans. on Image Proc.*, vol. 11, no. 4, pp. 485–492, 2002.

[8] S. C. Pei and I. K. Tam, "Effective color interpolation in CCD color filter arrays using signal correlation," *IEEE Tran. on Cir. and Sys. for Video Technology*, vol. 13, no. 6, pp. 503–513, 2003.

[9] G. J. Burton and I. R. Moorhead, "Color and spatial structure in natural scenes," *Applied Optics*, vol. 26, pp. 157–170, 1987.

[10] 40 scanned images, "Eastman kodak(c) photographic color image database," 1993.

[11] X. Li, "Demosaicing by successive approximation," *IEEE Trans. on Image Proc.*, In press, 2005.

[12] M. D. Fairchild, *Color Appearance Models*, Addison-Wesley, 1997.