

FILTERBANK-BASED UNIVERSAL DEMOSAICKING

Jing Gu, Patrick J. Wolfe

Harvard University
School of Engineering and Applied Sciences
Oxford Street, Cambridge, MA 02138 USA
jinggu@fas.harvard.edu
patrick@seas.harvard.edu

Keigo Hirakawa

University of Dayton
Electrical and Computer Engineering
300 College Park, Dayton, OH 45469 USA
k.hirakawa@notes.udayton.edu

ABSTRACT

As a result of renewed interests in industry and academia to provide image sensors with enhanced capabilities, both the theory and design of color filter arrays has advanced in recent years. By way of developments in spatio-spectral sampling and panchromatic pixels, advances in this area have contributed to increased resolution and enhanced noise performance. As such, the rest of the digital camera processing pipeline — including recovery methods such as demosaicking — must also evolve to maximize the benefits of improved color filter array design. In this article, we introduce a new universal demosaicking method that draws from the lessons learned in Bayer demosaicking designs, but can be applied to arbitrary array patterns. We recast the data-dependence of Bayer demosaicking as a parsimonious reconstruction of the underlying image signal that is inherently sparse in some representation. Using properties of filterbanks, we generalize this principle to yield a nonlinear recovery method that is consistent with the state-of-the-art Bayer demosaicking methods.

Index Terms— Filterbanks, color filter array, demosaicking

1. INTRODUCTION

Digital image data are typically obtained by way of a spatial subsampling procedure implemented as a color filter array (CFA), a physical construction whereby each pixel location measures only a single color. Most of these schemes involve the colors red, green, and blue (RGB). In particular, the Bayer pattern CFA [1], shown in Figure 1(a) overleaf, attempts to complement humans' spatial color sensitivity via a quincunx sampling of the green component that is twice as dense as that of red and blue.

In recent years, however, new color filter array designs have emerged that provide image sensors with enhanced capabilities. For example, panchromatic color filters — color filters whose quantum efficiencies correspond to multi-band spectral filters — are ideal for imaging when fewer photons are available, because these filter types allow the penetration of more photons. In fact, spectral selectivity stems from a combination of the transmittance of the color filter pigment and the quantum efficiency of the substrate, and thus *unfiltered* pixels¹ result in the maximal photon efficiency. Figure 1(b) shows a recently introduced color filter array pattern incorporating them, with the goal to increase robustness of the image sensor under

¹Based upon work supported in part by the Texas Instruments Incorporated Wireless University Relations Program, and the National Science Foundation under Grant No. DMS-0652743.

¹The unfiltered pixels are sometimes referred to as “white” — a term which we avoid to limit confusion with the “white point,” which typically is assumed to have equal portion of red, green, and blue.

noise [2]. Panchromatic pixels also have the ability to increase spatial resolution considerably, as RGB-based CFA sampling is unable to minimize the risk of aliasing [3]. A CFA pattern resulting from this spatio-spectral sampling strategy is shown in Figure 1(c).

Each CFA pattern considered above requires CFA-specific technique for *demosaicking* — the inverse problem of reconstructing a spatially undersampled vector field whose components correspond to RGB colors. The optimal solution to this ill-posed inverse problem, in the L^2 sense of an orthogonal projection onto the space of bandlimited functions separately for each spatially subsampled color channel, is well known to produce unacceptable visual distortions and artifacts. Key techniques used to overcome this limitation in the contemporary designs of Bayer demosaicking — by far the most studied demosaicking problem — include inter-color correlation and data-dependent processing [4].

The goal of this paper is to develop a new universal demosaicking method that yields complete RGB image given any CFA pattern. The proposed method differs from previous work in this area [5, 6], as we systematically generalize principles that enabled improved Bayer demosaicking methods. Specifically, we recast the data-dependency of demosaicking as a parsimonious spatio-spectral reconstruction of the underlying image signal that is inherently sparse in some representation, and demonstrate that the flexibility afforded by filterbank transforms will provide a rigorous framework for implementing this same reconstruction strategy regardless of the CFA sampling pattern.

The remainder of this paper is organized as follows. Section 2 provides a review of requisite mathematics for universal demosaicking. In Section 3 we leverage a filterbank-based convolution structure to model arbitrary CFA sampling, and propose reconstruction method based on robust regression. We provide simulation results in Section 4 and make concluding remarks in Section 5.

2. MATHEMATICAL BACKGROUND

2.1. Spatio-Spectral Sampling

Since three color components are typically employed to represent a full-color image, the idea to exploit correlation across color channels has been key to improving demosaicking performance. In particular, high frequency components of red, green, and blue components are highly redundant, suggesting that recovery of image features such as textures and edges are still possible if spatial and spectral contents of images were modeled jointly. Indeed, contemporary Bayer demosaicking methods have focused on this by analyzing the CFA sampled data $y[n]$ in the following manner [7]:

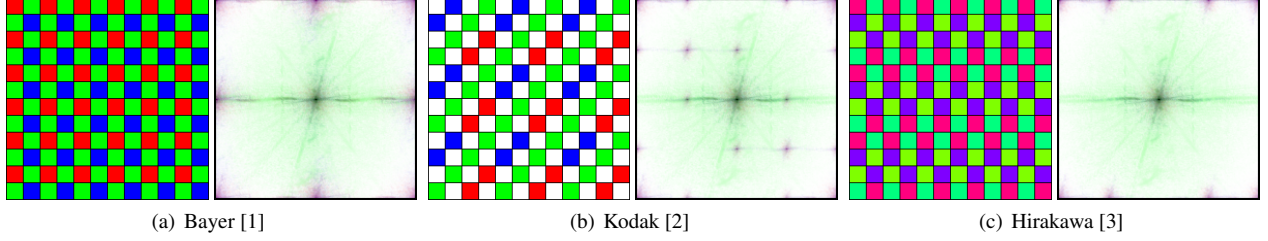


Fig. 1. Examples of color filter array patterns, and respective log-magnitude spectra of sensor data representing the “lighthouse” image.

$$\begin{aligned}
 y[n] &:= \mathbf{c}[n]^T \mathbf{x}[n] = \mathbf{c}[n]^T \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \mathbf{x}[n] \\
 &= \mathbf{c}[n]^T \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} r - g \\ g \\ b - g \end{bmatrix}
 \end{aligned}$$

where $\mathbf{x}[n] = (x_r[n], x_g[n], x_b[n])^T \in \mathbb{R}^3$ is the latent color image signal in RGB space, $\mathbf{c}[n] = (c_r[n], c_g[n], c_b[n])^T \in [0, 1]^3$ is the location-dependent color filter in RGB. With the assumption that $\sum_i c_i[n] = \gamma$ for some constant k , we have

$$y[n] = \gamma x_g[n] + c_r[n] \underbrace{(x_r[n] - x_g[n])}_{\alpha[n]} + c_b[n] \underbrace{(x_b[n] - x_g[n])}_{\beta[n]}. \quad (1)$$

Typically, $x_g[n]$ is interpreted as a proxy for luminance and the difference images $\alpha[n], \beta[n]$ represent chrominance. This is a convenient way to analyze sensor data because the high frequency components shared by $x_r[n], x_g[n], x_b[n]$ are subtracted away, yielding a subsampled lowpass difference images $c_r[n]\alpha[n]$ and $c_b[n]\beta[n]$.

Fourier analysis of y reveals that the overlapping support of $\hat{x}_g(\omega)$ and $\hat{c}_r \star \hat{\alpha}_g(\omega)$ (or $\hat{c}_b \star \hat{\beta}_g(\omega)$) results in aliasing [3, 7, 8] — examples are shown in Figure 1. Noting that Bayer CFA pattern in particular suffers from aliasing at $\omega = (0, \pi)$ or $(\pi, 0)$ when horizontal or vertical image features are present, no post-sampling linear processing will “un-do” such degradations.

This observation motivates us to consider a locally signal adaptive processing aimed at (i) identifying the type of image feature being represented by the CFA data, and (ii) smoothing the signal while preserving the detected feature. This notion is analogous to compressed sensing and wavelet-based signal denoising approaches that leverage sparse representations — conditioned on the support of significant signal coefficients (i.e., features), nonsignificant coefficients are suppressed. For example, suppose we conjecture that local image features may contribute to horizontally or vertically highpass frequency components, but not both [8]. If this is indeed the case, then the aliasing due to overlapping support of $\hat{x}_g(\omega)$ and $\hat{c}_r \star \hat{\alpha}_g(\omega)$ is limited to *either* $\omega = (0, \pi)$ or $(\pi, 0)$, allowing exact recovery of signal via amplitude demodulation conditioned on the detected feature orientation. We make this connection more precise in Section 3.

2.2. Subband Convolution Structure in Haar Filterbank

A multilevel filterbank structure is illustrated in Figure 2. If

$$\sum_n g_i[n] z^{-n} = 1 + (-1)^i z, \quad \sum_n h_i[n] z^{-n} = \frac{(-1)^i + z^{-1}}{2},$$

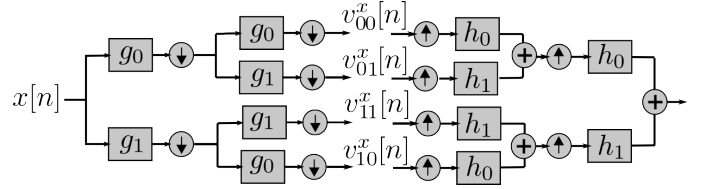


Fig. 2. A multilevel filterbank structure.

then $\{g_i, h_i\}_{i \in \mathbb{Z}_2}$ is said to comprise a Haar filterbank transform. Adopting a binary index representation, let $\mathbf{i} = (i_0, \dots, i_{I-1}) \in \mathbb{Z}_2^I$ be the subband index, where $i_k \in \mathbb{Z}_2$ reference the analysis filters used in the k th-level decomposition (i.e. g_0 or g_1). Then $v_{\mathbf{i}}^x[n]$ represents the n th filterbank transform coefficient of x in the \mathbf{i} th subband, and we have the following theorem [9].

Theorem 2.1 (Subband Convolution) *Let $v_{\mathbf{i}}^x[n], v_{\mathbf{i}}^y[n]$ be I -level Haar filterbank coefficient sequences corresponding to x and y , respectively, with $v_{\mathbf{i}}^{x \cdot y}[n]$ that of the element-wise product $x \cdot y$. Then, letting \otimes denote cyclic convolution, we have the relation:*

$$v_{\mathbf{i}}^{x \cdot y}[n] = v_{\mathbf{i}}^x[n] \otimes v_{\mathbf{i}}^y[n] := 2^{-I} \sum_{\mathbf{i}' \in \mathbb{Z}_2^I} v_{\mathbf{i}+\mathbf{i}'}^x[n] v_{\mathbf{i}'}^y[n].$$

Extending to the two-dimensional Haar filterbank transform, we immediately obtain the result for double-subband index (\mathbf{i}, \mathbf{j}) , for horizontal and vertical directions, respectively.

Corollary 2.2 (2D Subband Convolution) *Let $v_{\mathbf{i}, \mathbf{j}}^x[n], v_{\mathbf{i}, \mathbf{j}}^y[n]$ be I -level two-dimensional Haar filterbank coefficient sequences corresponding to x and y , respectively. Then we have the relation:*

$$v_{\mathbf{i}, \mathbf{j}}^{x \cdot y}[n] = v_{\mathbf{i}, \mathbf{j}}^x[n] \otimes v_{\mathbf{i}, \mathbf{j}}^y[n] := 4^{-I} \sum_{(\mathbf{i}', \mathbf{j}') \in \mathbb{Z}_2^I \times \mathbb{Z}_2^I} v_{\mathbf{i}+\mathbf{i}', \mathbf{j}+\mathbf{j}'}^x[n] v_{\mathbf{i}', \mathbf{j}'}^y[n].$$

3. UNIVERSAL DEMOSAICKING

3.1. Filterbank Analysis Of CFA Sampling

The Fourier analysis of (1) has previously yielded several useful insights into demosaicking [7, 8, 10] and CFA [3] designs. However, this analysis is inherently *global* — meaning it does not describe the influence of the *local* regularity of the image signal on recoverability. To this end, we are able to adopt the filterbank-based sampling theory described above to characterize CFA sampling in the joint time-frequency domain.

By Corollary 2.2, the filterbank transform of $y[n]$ is:

$$\begin{aligned}
 v_{\mathbf{i}, \mathbf{j}}^y[n] &= \gamma v_{\mathbf{i}, \mathbf{j}}^{x_g}[n] + v_{\mathbf{i}, \mathbf{j}}^{c_r \alpha}[n] + v_{\mathbf{i}, \mathbf{j}}^{c_b \beta}[n] \\
 &= \gamma v_{\mathbf{i}, \mathbf{j}}^{x_g}[n] + v_{\mathbf{i}, \mathbf{j}}^{c_r}[n] \otimes v_{\mathbf{i}, \mathbf{j}}^{\alpha}[n] + v_{\mathbf{i}, \mathbf{j}}^{c_b}[n] \otimes v_{\mathbf{i}, \mathbf{j}}^{\beta}[n].
 \end{aligned} \quad (2)$$

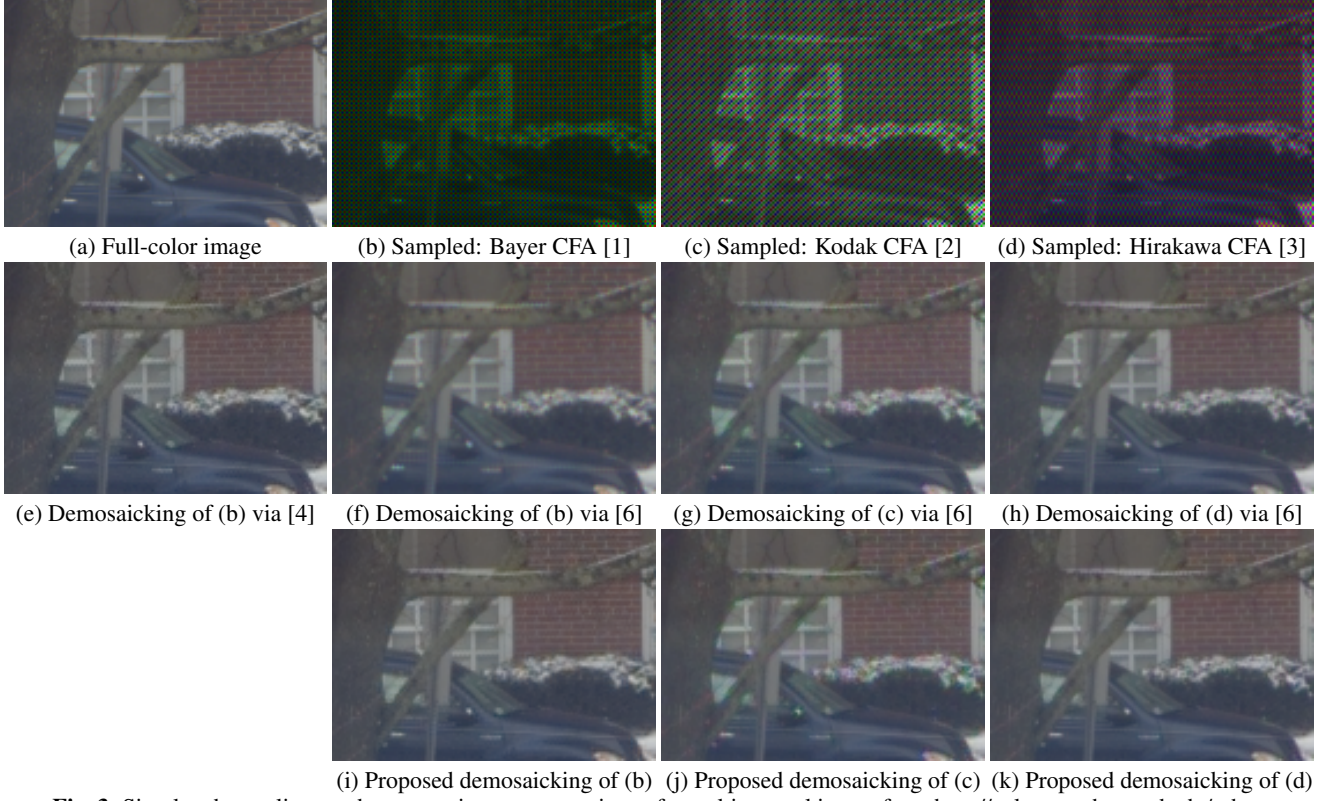


Fig. 3. Simulated samplings and comparative reconstructions of a multispectral image from <http://color.eecs.harvard.edu/rgbw>.

Owing to the lowpass nature of difference images α and β , high frequency filterbank coefficients of α and β are insignificant. In this article, we let $v_{i,j}^{\alpha}[n] = v_{i,j}^{\beta}[n] = 0$ when $(i, j) \neq (0, 0)$ to simplify notation, though it is straightforward to generalize to the cases when more coefficients are nonzero. Simplifying (2), we obtain

$$v_{i,j}^y[n] = \gamma v_{i,j}^{x_g}[n] + 4^{-I} v_{i,j}^{c_r}[n] v_{0,0}^{\alpha}[n] + 4^{-I} v_{i,j}^{c_b}[n] v_{0,0}^{\beta}[n]. \quad (3)$$

That is to say, each filterbank coefficient $v_{i,j}^y[n]$ is a linear combination of $v_{i,j}^{x_g}[n]$ and the difference images $v_{0,0}^{\alpha}[n]$, $v_{0,0}^{\beta}[n]$ (if $v_{i,j}^{c_r}[n]$, $v_{i,j}^{c_b}[n]$ are nonzero). This is similar to the analysis of [3, 7], where $v_{i,j}^{c_r}[n] \otimes v_{i,j}^{\alpha}[n] = v_{i,j}^{c_r}[n] v_{0,0}^{\alpha}[n]$ is the joint time-frequency analogue of *amplitude modulation*. However, the spatially local nature of filterbank transform makes this an attractive alternative for designing locally signal adaptive demosaicking.

3.2. Demosaicking By Parsimonious Recovery

The objective of the demosaicking is to recover an RGB image \mathbf{x} from CFA data \mathbf{y} . By linearity, this interpolation is equivalent to the “estimation” of $\{v_{i,j}^{x_g}[n], v_{i,j}^{\alpha}[n], v_{i,j}^{\beta}[n]\}$ from $\{v_{i,j}^y[n]\}$ (followed by inverse transform to recover \mathbf{x}). The filterbank analysis of CFA data in (3) implies that, in the statistical sense of conditional expectation,

$$E[v_{i,j}^{x_g}[y]] = \gamma^{-1} \left(v_{i,j}^y - 4^{-I} v_{i,j}^{c_r} E[v_{0,0}^{\alpha}|y] - 4^{-I} v_{i,j}^{c_b} E[v_{0,0}^{\beta}|y] \right),$$

and so in this section we focus on the recovery of $v_{0,0}^{\alpha}[n]$, $v_{0,0}^{\beta}[n]$.

Suppose for a moment that $v^{x_g} = 0$. Then $\{v_{i,j}^y[n]\}$ is an overdetermined representation of $v_{0,0}^{\alpha}[n]$, $v_{0,0}^{\beta}[n]$. In fact, the optimal linear recovery of $v_{0,0}^{\alpha}[n]$, $v_{0,0}^{\beta}[n]$ from $\{v_{i,j}^y[n]\}$ is a pseudoinverse operator — this is equivalent to treating $\{v_{i,j}^y[n]\}$ as “a noisy

overcomplete representation” of $v_{0,0}^{\alpha}[n]$, $v_{0,0}^{\beta}[n]$. This perspective is surprisingly useful for designing a nonlinear demosaicking method. As is by now standard in image processing, suppose we assume that $v_{i,j}^{x_g}[n]$ is sparse. Then $v_{i,j}^y = 4^{-I} v_{i,j}^{c_r}[n] v_{0,0}^{\alpha}[n] + 4^{-I} v_{i,j}^{c_b}[n] v_{0,0}^{\beta}[n]$ “except when $v_{i,j}^{x_g}$ is active.” Under this scenario, the regression of $v_{i,j}^y$ on $v_{0,0}^{\alpha}$ and $v_{0,0}^{\beta}$ is a problem of outlier detection, where an outlier stems from the nonzero $v_{i,j}^{x_g}$ coefficient indicating a presence of a particular image feature. A variety of statistical tools exist for robust regression, including M-estimation, minimum L^1 error estimate, total least squares regression, and Bayesian mixture models. For the results we describe here, we employed iteratively-reweighted least-squares, as implemented by the Matlab function `robustfit`, though we note that any of the variety of methods described above could equally well be employed.

Additional *a priori* knowledge about the image can be used to improve robust regression performance. For instance, suppose a certain type of image feature involving a coefficient $v_{i,j}^{x_g}$ is detected. Then we declare $v_{i,j}^y$ as an outlier, thereby “pruning” the overcomplete representation. To demonstrate the utility of this, we implemented a simple edge detector (similar to [8]) based on filterbank subbands, which we describe below.

Recall that edge detection typically involves highpass filtering followed by thresholding. However, this approach alone will not provide useful information in demosaicking, as CFA sampling causes α and β to modulate to higher frequencies — ordinary highpass filter will not be able to distinguish x_g from $c_r \alpha$ and $c_b \beta$. Instead, we draw edge information from filterbank subbands (i, j) where $v_{i,j}^{c_r}[n] = v_{i,j}^{c_b}[n] = 0$ — i.e. $v_{i,j}^y[n] = v_{i,j}^{x_g}[n]$. By comparing the magnitudes of *highest frequency filterbank subbands* (subject to $v_{i,j}^y[n] = v_{i,j}^{x_g}[n]$) in horizontal and vertical directions, we arrive at

Table 1. Reconstruction performance (PSNR, dB) for the multispectral image dataset described at <http://color.eecs.harvard.edu/rgbw>

CFA	Demosaicking Method	Image Index					
		1	2	3	4	5	6
Bayer [1]	[4]	40.53	33.63	42.23	38.36	32.90	41.93
	[6]	40.68	32.65	41.80	39.43	32.79	41.70
	Proposed	41.11	33.90	42.06	38.42	32.92	41.67
Kodak [2]	[6]	39.50	32.50	40.84	31.55	27.98	37.83
	Proposed	41.06	32.40	41.25	37.78	33.02	41.62
Hirakawa [3]	[6]	42.20	34.64	43.20	39.27	33.91	42.67
	Proposed	42.24	35.34	43.84	39.94	34.81	43.69

a simple image feature orientation detector for CFA data.

4. EXPERIMENTAL RESULTS

Here we report results based on a 2-level filterbank implementation. The proposed demosaicking method is configured to treat “unfiltered” pixels in the Kodak CFA pattern of Figure 1(b) as a fourth color (instead of “white”) — that is, \mathbf{x} is now $\mathbf{x} = (x_r, x_g, x_b, x_u)$ where x_u is the unfiltered channel. This ensures that correlation across color channels are exploited without introducing color shifts [11]. In this case, (1) is generalized to incorporate $x_u - x_g$ as the third lowpass chrominance type.

We provide experimental results using the multispectral image dataset described at <http://color.eecs.harvard.edu/rgbw>, captured by a CRI Nuance System. This is a time-multiplexing device that captures spectrally narrowband images, so that complete RGB images (i.e., ground truth) can be constructed by integrating over the color matching functions. The advantage of using this system is that we can precisely simulate filtered and unfiltered pixels.

A ground truth image and its CFA sampled versions are shown in Figure 3. The demosaicking results, shown in Figures 3(e-k), indicate that the performance of the proposed method is comparable to the contemporary demosaicking methods for Bayer CFA pattern [4] and the universal demosaicking method of [6]. Comparing reconstructed images from Bayer and Kodak CFA patterns, the latter is seen to yield better smooth and simple edges, as the sampling pattern provides additional equations for robust regression scheme. Intuitively, the unfiltered pixels in the Kodak CFA pattern emphasize shape information, which provides the enhancement on smooth surface reconstruction but less accuracy on textures compared to the Bayer CFA pattern. For example, the surface of the car and the window are better reconstructed in Figure 3(j) than in Figure 3(i), whereas the artifacts of the snow on the bush in Figure are highly noticeable. Reconstruction from the CFA pattern of [3] suggests a better balance between color, shape, and texture information. Table 1 shows a numerical performance summary of peak signal-to-noise ratio (PSNR) for six reconstructed images. By this metric, the universal demosaicking method presented here attains comparable performance on Kodak and Bayer CFA patterns.

5. CONCLUSION

This article has developed a novel universal demosaicking strategy to work with any color filter array pattern. This strategy is based on spatio-spectral sampling theory and a filterbank-based treatment of color image sampling. Under the assumption of sparsity of high frequency filterbank coefficients, it was shown that reconstruction of

chrominance information can be reinterpreted as a robust regression scheme, regardless of the color filter array sampling pattern. Experimental results performed on multispectral images have demonstrated the effectiveness of the proposed method.

6. ACKNOWLEDGMENT

The authors give thanks to A. Chakrabarti and T. Zickler of Harvard University for providing the multispectral image database, and L. Condat of GREYC Lab for providing code to the method of [6].

7. REFERENCES

- [1] B. E. Bayer, “Color imaging array,” US Patent 3 971 065, 1976.
- [2] J. Compton and J. Hamilton, “Image sensor with improved light sensitivity,” US Patent 20 070 024 031A1, 2005.
- [3] K. Hirakawa and P. J. Wolfe, “Spatio-spectral color filter array design for optimal image recovery,” *IEEE Trans. Image Process.*, no. 10, pp. 1876–90, 2008.
- [4] B. K. Gunturk, Y. Altunbasak, and R. M. Mersereau, “Color plane interpolation using alternating projections,” *IEEE Trans. Image Process.*, vol. 11, pp. 997–1013, 2002.
- [5] R. Lukac and K. N. Plataniotis, “Universal demosaicking for imaging pipelines with an RGB color filter array,” *Pattern Recognition*, vol. 38, no. 11, pp. 2208–2212, Nov. 2005.
- [6] L. Condat, “A generic variational approach for demosaicking from an arbitrary color filter array,” in *Proc. IEEE Internat. Conf. Image Process.*, 2009.
- [7] D. Alleysson, S. Süsstrunk, and J. Hérault, “Linear demosaicking inspired by the human visual system,” *IEEE Trans. Image Process.*, vol. 14, pp. 439–449, 2005.
- [8] E. Dubois, “Filter design for adaptive frequency-domain Bayer demosaicking,” in *Proc. IEEE Internat. Conf. Image Process.*, 2006, pp. 2705–2708.
- [9] K. Hirakawa and P. J. Wolfe, ““Rewiring” filterbanks for local Fourier analysis: Theory and practice,” submitted manuscript, 2009, DOI arXiv:0909.1338.
- [10] K. Hirakawa, X.-L. Meng, and P. J. Wolfe, “A framework for wavelet-based analysis and processing of color filter array images with applications to denoising and demosaicking,” in *Proc. IEEE Internat. Conf. Acoust. Speech Signal Process.*, 2007, vol. 1, pp. 597–600.
- [11] M. Kumar, E. Morales, J. Adams, and W. Hao, “New digital camera sensor architecture for low light imaging,” in *Proc. IEEE Internat. Conf. Image Process.*, 2009.