# POCS-BASED RESTORATION OF BAYER-SAMPLED IMAGE SEQUENCES

*Murat Gevrekci, Bahadir K. Gunturk**

Louisiana State University
Electrical and Computer Engineering
Baton Rouge, LA 70803
lgevre1@lsu.edu, bahadir@ece.lsu.edu

*Yucel Altunbasak*

Georgia Institute of Technology
Electrical and Computer Engineering
Atlanta, GA 30332
yucel@ece.gatech.edu

## ABSTRACT

Spatial resolution of digital images are limited due to optical/sensor blurring and sensor site density. In single-chip digital cameras, the resolution is further degraded because such devices use a color filter array to capture only one spectral component at a pixel location. The process of estimating the missing two color values at each pixel location is known as demosaicking. Demosaicking methods usually exploit the correlation among color channels. When there are multiple images, it is possible not only to have better estimates of the missing color values but also to improve the spatial resolution further (using super-resolution reconstruction). Previously, we have proposed a demosaicking algorithm based on the projection onto convex sets (POCS) technique. In this paper, we improve the results of that algorithm adding a new constraint set based on the spatio-intensity neighborhood. We extend the algorithm to image sequences for *multi-frame demosaicking* and super resolution.

***Index Terms***— Super resolution, demosaicking, POCS

## 1. INTRODUCTION

To represent a color image, a sensor needs at least three color samples at every pixel location. However, putting three separate sensors at each camera is not a preferred approach because of cost and design difficulties. Instead, manufacturers use a color filter array (CFA) with a single sensor to capture only one color sample at pixel location. The color samples obtained with such a CFA must then be interpolated to estimate the missing samples. Because of the mosaic pattern of the color samples, this CFA interpolation problem is also referred to as "demosaicking" in the literature [1].

The most commonly used CFA is the Bayer pattern. The Bayer pattern includes twice as many green samples as red/blue samples, taking advantage of the fact that human eye is more sensitive to changes in green than in red/blue. Standard interpolation techniques (such as bilinear and bicubic interpolation) do not produce good results in the demosaicking problem. When the correlation among color channels are not taken into account, visible artifacts occur. There is a variety of demosaicking algorithms in the literature, including edge-directed [2, 3], constant-hue-based [4], Laplacian as correction terms [5], alias canceling [6], projections onto convex sets based [7], and homogeneity-directed [8] interpolation methods. A comprehensive survey/comparison of these algorithms is provided in [1].

When there are multiple images, spatial resolution can be improved beyond the physical limits of a sensor chip through subpixel registration. This multi-frame resolution enhancement is known as super-resolution (SR) reconstruction. A recent IEEE Signal Processing Magazine provides a comprehensive survey on SR image reconstruction [9]. Although there have been a significant amount of work on SR reconstruction, there are only few recent efforts modeling CFA sampling and addressing issues related to image reconstruction from CFA-sampled data.

In [10], data fidelity and regularization terms are combined to produce high-resolution images. The data fidelity term is based on a cost function that is composed by the sum of residual differences between actual observations and high-resolution image projected onto observations ($simulated$ observations). Regularization functions are added to this cost function to eliminate color artifacts and preserve edge structures. A similar algorithm is presented in [11].

Among different demosaicking approaches, the POCS-based demosaicking algorithm has one of the best performance in terms of mean square error [7], [1]. One advantage of the POCS framework is that new constraints can be easily incorporated into a POCS-based algorithm. Here, we add a new contraint set to this framework. This new constraint set is based on the spatio-intensity neighborhood of a pixel.

We also modify this algorithm for super-resolution reconstruction of CFA sampled image sequences. We investigate two approaches: *sequential* and *joint*. In sequential approach, we separate the restoration process into two parts. The first part is multi-frame demosaicking, where missing color samples are estimated. The second part is super-resolution reconstruction, where sub-pixel level resolution is achieved. In joint approach, super-resolution and demosaicking steps are merged into the same framework.

In Section 2, we present an imaging model that includes CFA sampling. We introduce our POCS-based *multi-frame demosaicking* algorithm in section 3. The Section includes three constraint sets and corresponding projection operators for multi-frame demosaicking. Single-frame demosaicking is a subset of this algorithm. In Section 4, we explain how to achieve sub-pixel resolution. Experimental results and discussions are provided in Section 5.

## 2. IMAGING MODEL

We use an image acquisition model that includes CFA sampling in addition to motion and blurring. The model has two steps: the first step models conversion of a high-resolution full-color image into a low-resolution full-color image. The second step models color filter sampling of full-color low-resolution image.

Let $\mathbf{x}_S$ be the unknown high-resolution image with color channels red ($\mathbf{x}_R$), green ($\mathbf{x}_G$), and blue ($\mathbf{x}_B$). The $i$th observation, $\mathbf{y}_S^{(i)}$, is obtained from this high-resolution image through spatial warping, blurring, and downsampling operations:

$$\mathbf{y}_S^{(i)} = DCW^{(i)}\mathbf{x}_S, \quad \text{for } S = R, G, B, \text{ and } i = 1, 2, ..., K, \tag{1}$$

*This work was supported by the National Science Foundation under Grant No 0528785.

where $K$ is the number of input images, $W^{(i)}$ is the warping operation (to account for the relative motion between observations), $C$ is the convolution operation (to account for the point spread function of the camera), and $D$ is the downsampling operation (to account for the spatial sampling of the sensor).

The full-color image $\mathbf{y}_S^{(i)}$ is then converted to a mosaicked observation $\mathbf{z}^{(i)}$ according to a CFA sampling pattern:

$$\mathbf{z}^{(i)} = \sum_{S=R,G,B} M_S \mathbf{y}_S^{(i)} \tag{2}$$

where $M_S$ takes only one of the color samples at a pixel according to the pattern. For example, at red pixel location, $[M_R,\ M_G,\ M_B]$ is $[1,\ 0,\ 0]$.

We investigate two approaches to obtain $\mathbf{x}_S$. In the sequential approach, each full-color low-resolution image $\mathbf{y}_S^{(i)}$ is estimated from the multiple observations $\mathbf{z}^{(i)}$, and then SR reconstruction is applied to these estimates $\mathbf{y}_S^{(i)}$ to obtain $\mathbf{x}_S$. In the joint approach, $\mathbf{x}_S$ is obtained directly from $\mathbf{z}^{(i)}$ without estimating $\mathbf{y}_S^{(i)}$.

## 3. MULTI-FRAME DEMOSAICKING

In this section, we introduce multi-frame demosaicking, where multiple images are used to estimate missing samples in a Bayer-sampled color image. The process is based on the POCS technique, where an initial estimate is projected onto convex constraint sets iteratively to reach a solution that is consistent with all constraints about the solution.

This algorithm is based on the demosaicking approach presented in [7]. In [7], we define two constraint sets. The first constraint set stems from inter-channel correlation. It is based on the idea that high-frequency components of red, green, and blue channels should be similar in an image. This constraint set is also used in this paper.

The second constraint set ensures that the reconstructed image is consistent with the observed data of the same image. This constraint set can be improved when there are multiple observation of the same scene. A missing sample (due to the CFA sampling) in an image could have been captured in another image due to the relative motion between observations. It is possible to obtain a better estimate of a missing sample by taking multiple images into account. Therefore, we define a constraint set based on the samples coming from multiple images.

In addition to previous constraint sets, we propose a third constraint set in this paper to achieve color consistency. Consistent neighbors of a pixel is decided according to both spatial and intensity closeness. Intensity of a pixel is updated using the average value of its consistent neighbors. These constraint sets are examined in following sections.

### 3.1. Detail Constraint Set

Let $\mathcal{W}_k$ be an operator that produces the $k$th subband of an image. There are four frequency subbands ($k = LL, LH, HL, HH$) corresponding to low-pass filtering and high-pass filtering permutations along horizontal and vertical dimensions.

The "detail" constraint set ($C_d$) forces the details (high frequency components) of the red and blue channels to be similar to the details of the green channel at every pixel location $(n_1, n_2)$, and is defined as follows:

$$C_d = \left\{ \begin{array}{l} \mathbf{y}_S^{(i)}(n_1,n_2): \\ \left| \left(\mathcal{W}_k \mathbf{y}_S^{(i)}\right)(n_1,n_2) - \left(\mathcal{W}_k \mathbf{y}_G^{(i)}\right)(n_1,n_2) \right| \le \mathbf{T}_d(n_1,n_2) \\ \quad \forall\,(n_1,n_2),\quad for\ k = LH, HL, HH\ \ and\ \ S = R,B \end{array} \right\} \tag{3}$$

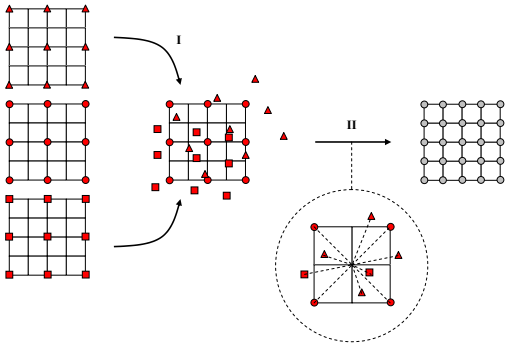


**Fig. 1**. I: Input images are warped onto reference frame. II: Weighted average of samples are taken to find values on the reference sampling grid.

where $\mathbf{T}_d(n_1, n_2)$ is a positive threshold that quantifies the "closeness" of the detail subbands to each other. For details and further discussion on this constraint set, we refer the reader to [7].

### 3.2. Observation Constraint Set

The interpolated color channels must be consistent with the color samples captured by the digital camera for all images. Even if a color sample does not exist in an image as a result of Bayer sampling, that particular sample could have been captured in another frame (due to motion). By warping all captured samples onto the common frame to be demosaicked, we can obtain a good estimate of the missing samples.

Figure 1 illustrates the sampling idea. In the figure, red channels of three Bayer-sampled images are shown. The grid locations with "triangles", "circles", and "squares" are the ones that have red samples. We would like to estimate the missing samples in the middle image. The other two images warped onto the middle image. The estimation problem now becomes an interpolation problem from a set of nonuniformly sampled data. After the interpolation, we obtain an "observation" image, $\bar{\mathbf{y}}_S^{(i)}$. The demosaicked samples $\mathbf{y}_S^{(i)}$ should be consistent with this observation image $\bar{\mathbf{y}}_S^{(i)}$. To make sure that, we can define the "observation" constraint set $C_o$ as follows:

$$C_o = \left\{ \begin{array}{l} \mathbf{y}_S^{(i)}(n_1,n_2): \\ \left| \mathbf{y}_S^{(i)}(n_1,n_2) - \bar{\mathbf{y}}_S^{(i)}(n_1,n_2) \right| \le \mathbf{T}_o(n_1,n_2) \\ \forall\,(n_1,n_2)\ \ and\ S = R,G,B \end{array} \right\}, \tag{4}$$

where $\mathbf{T}_o(n_1, n_2)$ is a positive threshold to quantify the closeness of the estimated pixels to the averaged data.

The observation image $\bar{\mathbf{y}}_S^{(i)}$ should be interpolated from nonuniform samples. There are different approaches to do nonuniform interpolation. One approach is to take a weighted average of all samples within a neighborhood of a pixel in question. For example, if a sample is $d$ away from the pixel in question, its weight could be set to be proportional to $e^{-(d/\sigma)^2}$, where $\sigma$ is a constant parameter. This requires accurate geometric registration as weight of each sample is determined according to the spatial closeness to the sampling location. In our experiments, we used simple bilinear interpolation to estimate the missing samples. That is, all frames are warped onto the reference frame using bilinear interpolation. These warped images are then averaged to obtain $\bar{\mathbf{y}}_S^{(i)}$.

One benefit of using this constraint set is noise reduction. When the input data is noisy, constraining the solution to be close to the

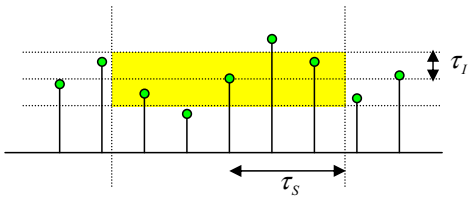


**Fig. 2**. Spatio-intensity neighborhood of a pixel is illustrated on a one-dimensional image. The yellow region is the neighborhood of the pixel in the middle.

observation image $\bar{\mathbf{y}}_S^{(i)}$ prevents amplification of noise and color artifacts.

### 3.3. Color Consistency Constraint Set

The third constraint set is the color consistency constraint set. It is reasonable to expect pixels with similar green intensities have similar red and blue intensities within a small spatial neighborhood. Therefore, we define *spatio-intensity neighborhood* of a pixel. Suppose that green channel $\mathbf{x}_G$ of an image is already interpolated and we would like to estimate the red value at a particular pixel $(n_1, n_2)$. Then, the spatio-intensity neighborhood of the pixel $(n_1, n_2)$ is defined as

$$\mathcal{N}(n_1,n_2) = \left\{ \begin{array}{l} (m_1,m_2) : \|(m_1,m_2)-(n_1,n_2)\| \le \tau_S \\ \quad and \ \left| \mathbf{y}_G^{(i)}(m_1,m_2) - \mathbf{y}_G^{(i)}(n_1,n_2) \right| \le \tau_I \end{array} \right\} \tag{5}$$

where $\tau_S$ and $\tau_I$ are spatial and intensity neighborhood ranges, respectively. Figure 2 illustrates spatio-intensity neighborhood for a one-dimensional signal. In that figure, we would like to determine the spatio-intensity neighbors of the pixel in the middle. The spatial neighborhood is determined by $\tau_S$; and the intensity neighborhood is determined by $\tau_I$. The yellow region shows the spatio-intensity neighborhood.

The spatio-intensity neighbors of a pixel should have similar color values. One way to measure color similarity is to inspect color differences between red and green channels (and blue and green channels). These differences are expected to be similar within the spatio-intensity neigborhood $\mathcal{N}(n_1, n_2)$. Therefore, the color consistency constraint set can be defined as follows:

$$C_c = \left\{ \begin{array}{l} \mathbf{y}_S^{(i)}(n_1,n_2) : \left| \left( \mathbf{y}_S^{(i)}(n_1,n_2) - \mathbf{y}_G^{(i)}(n_1,n_2) \right) \right. \\ \quad \left. - \overline{\left( \mathbf{y}_S^{(i)}(n_1,n_2) - \mathbf{y}_G^{(i)}(n_1,n_2) \right)} \right| \le \mathbf{T}_c(n_1,n_2) \\ \forall (n_1,n_2) \ and \ S = R, B \end{array} \right\} \tag{6}$$

where $\overline{(\cdot)}$ is averaging within the neighborhood $\mathcal{N}(n_1, n_2)$, and $\mathbf{T}_c(n_1, n_2)$ is a positive threshold.

An issue to be considered with all these constraint sets is the selection of the threshold values $\mathbf{T}_d(n_1, n_2)$, $\mathbf{T}_o(n_1, n_2)$, and $\mathbf{T}_c(n_1, n_2)$. The detail constraint set should be tight (i.e. $\mathbf{T}_d(n_1, n_2)$ should be small) when the inter-channel correlation is high [7]. When the images are noisy, a strict observation constraint that forces the solution towards the average image $\bar{\mathbf{y}}_S^{(i)}$ would prevent color artifacts. Therefore, noise variance and the number of images used to obtain $\bar{\mathbf{y}}_S^{(i)}$ should be considered during the selection of $\mathbf{T}_o(n_1, n_2)$. If the images were noise-free and the warping process was perfect, then we would keep $\mathbf{T}_o(n_1, n_2)$ very small. The color consistency threshold $\mathbf{T}_c(n_1, n_2)$ determines how much the color difference smoothness is enforced. The selection of the spatio-intensity neighborhood is also critical in this constraint set. When the spatial range $\tau_S$ too large, local spatial neighborhood assumption could be violated. When $\tau_S$ is too small, noise might become an issue. On the other hand, the intensity range $\tau_I$ should not be too large not to take unrelated pixels into account. In our experiments, we determined optimal parameters by trial-and-error. We leave the robust and automatic parameter selection issue as a future work.

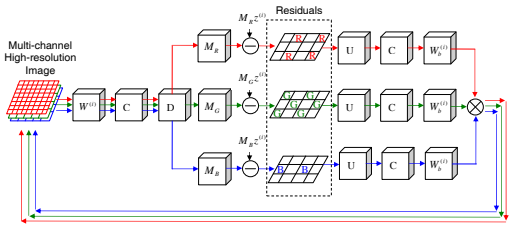


**Fig. 3**. Joint reconstruction approach is illustrated. The algorithm starts with a high-resolution image estimate. Simulated observations are obtained by forward imaging operations, including the CFA sampling. The residuals are computed on Bayer pattern samples for each channel, and then back-projected. The notation in the figure is as follows. $W^{(i)}$: Spatial warping onto $i$th observation, $C$: Convolution with the PSF, $D$: Downsampling by the resolution enhancement factor, $U$: Upsampling by zero insertion, $W_b^{(i)}$: Back-warping to the reference grid.

### 3.4. Projection Operations

Because of the limited space, we do not include the projection operators in this paper. The projection operators for the detail and observation constraint sets are given in [7]. The projection operator for the color constraint set can be derived similarly.

## 4. ACHIEVING SUBPIXEL RESOLUTION

### 4.1. Sequential Approach

In the sequential approach, demosaicking is first performed on each frame, and then super-resolution reconstruction is applied to the demosaicked images. For demosaicking, either single-frame demosaicking or multi-frame demosaicking (as presented in this paper) can be applied. For super-resolution reconstruction, each color channel can be treated separately using any of the super-resolution algorithms in the literature. We use the iterated back projections algorithm in our experiments [12]. Starting with an initial estimate for the high-resolution image, warping, blurring and downsampling operations are applied to produce a $simulated$ observation. The residual between the $simulated$ observation and the actual observation $\mathbf{y}_S^{(i)}$ is then back-projected. The back-projection operation includes upsampling by zero insertion, blurring, and back-warping. This is done for each channel separately. The process is repeated until convergence.

### 4.2. Joint Approach

The other approach is to perform demosaicking and super-resolution reconstruction jointly. That is, a high-resolution image is reconstructed from the input data $\mathbf{z}^{(i)}$ without estimating full-color low-resolution images $\mathbf{y}_S^{(i)}$. The reconstruction process is again based on the iterated back-projection algorithm [12]. Figure 3 shows an illustration of this approach.

**Fig. 4**. Three of the Bayer-pattern input images


**Fig. 5**. Single- and multi-frame demosaicking results. (a) Single frame demosaicking using [3]. (b) Single frame demosaicking using [7]. (c) Multi-frame demosaicking with detail and observation constraints used ($\mathbf{T}_d$=1, $\mathbf{T}_o$=20). (d) Multi-frame demosaicking with all constraints applied ($\mathbf{T}_d$=1, $\mathbf{T}_o$=20, $\tau_S$=4, $\tau_I$=6, $\mathbf{T}_c$=2). The iteration number for multi-frame demosaicking is five. (e) Sequential SR result using images multi-frame demosaicking with all constraint sets. (f) Joint SR result. The enhancement factor for SR is two.

## 5. EXPERIMENTS

We captured an image sequence of length 21. These images are then sampled with the Bayer pattern to produce the input data. The algorithms are applied on this data. Figure 4 shows three of these input data. The motion parameters are estimated using feature-based image registration technique [13]. The motion parameters are estimated from the interpolated green channels due to its high sampling rate.

Due to the limited space, we provide selected results from our experiments. Figure 5 includes results of single-frame demosaicking algorithms [3] and [7], the proposed multi-frame demosaicking algorithm with and without the spatio-intensity constraint set, and super-resolution results. As seen in the figure, the spatio-intensity constraint set improves results. SR reconstruction increases resolution further. The sequential SR performs better than the joint SR in terms of color artifacts, as we did not include any inter-channel constraint in the joint SR approach.

## 6. CONCLUSIONS

In this paper, we present a POCS-based framework for resolution enhancement of Bayer-sampled video sequences. We defined three constraint sets to perform demosaicking using multiple images. This multi-frame demosaicking algorithm performed much better than single-frame demosaicking algorithms. We also investigated two different approaches for super-resolution reconstruction. The sequential approach is performed on demosaicked images. When multi-frame demosaicking is used, the super-resolution result has less noise and color artifacts. The joint approach performed worse than the sequential approach because no inter-channel correlation was utilized. In these initial experiments, the reconstruction parameters were chosen by trial-and-error. As a future work, we will look into the parameter selection issue in more detail.

## 7. REFERENCES


[1] B. K. Gunturk, J. Glotzbach, Y. Altunbasak, R. W. Schafer, and R. M. Mersereau, "Demosaicking: color filter array interpolation," *IEEE Signal Processing Magazine*, vol. 22, no. 1, pp. 44–54, January 2005.

[2] R. H. Hibbard, "Apparatus and method for adaptively interpolating a full color image utilizing luminance gradients," *U.S. Patent 5,382,976*, 1995.

[3] C. A. Laroche and M. A. Prescott, "Apparatus and method for adaptively interpolating a full color image utilizing chrominance gradients," *U.S. Patent 5,373,322*, 1994.

[4] J. E. Adams, "Interactions between color plane interpolation and other image processing functions in electronic photography," in *Proc. SPIE*, vol. 2416, 1995, pp. 144–151.

[5] J. E. Adams and J. F. Hamilton, "Design of practical color filter array interpolation algorithms for digital cameras," in *Proc. SPIE*, vol. 3028, 1997, pp. 117–125.

[6] J. W. Glotzbach, R. W. Schafer, and K. Illgner, "A method of color filter array interpolation with alias cancellation properties," in *Proc. IEEE Int. Conf. Image Processing*, vol. 1, 2001, pp. 141–144.

[7] B. K. Gunturk, Y. Altunbasak, and R. M. Mersereau, "Color plane interpolation using alternating projections," *IEEE Trans. Image Processing*, vol. 11, no. 9, pp. 997–1013, Sept. 2002.

[8] K. Hirakawa and T. W. Parks, "Adaptive homogeneity-directed demosaicing algorithm," *Proc. IEEE Int. Conf. Image Processing*, vol. 3, pp. 669–672, 2003.

[9] "Super-resolution image reconstruction," *IEEE Signal Processing Magazine*, vol. 20, no. 3, pp. 21–86, May 2003.

[10] S. Farsiu, D. Robinson, M. Elad, and P. Milanfar, "Advances and challenges in super-resolution," *International Journal of Imaging Systems and Technology*, vol. 14, no. 2, pp. 47–57, August 2004.

[11] T. Gotoh and M. Okutomi, "Direct super-resolution and registration using raw cfa images," in *Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition*, vol. 2, July 2004, pp. 600–607.

[12] M. Irani and S. Peleg, "Improving resolution by image registration," *CVGIP: Graphical Models and Image Processing*, vol. 53, pp. 231–239, May 1991.

[13] D. Capel and A. Zisserman, "Super-resolution enhancement of text image sequences," in *Proc. IEEE Int. Conf. Pattern Recognition*, vol. 1, 2000, pp. 600–605.