Compressive Sensing Multi-spectral Demosaicing from Single Sensor Architecture

Hemant Kumar Aggarwal and Angshul Majumdar

Indraprastha Institute of Information Technology – Delhi

ABSTRACT

This paper addresses the recovery of multi-spectral images from single sensor cameras using compressed sensing (CS) techniques. It is an exploratory work since this particular problem has not been addressed before. Thus, we do not attempt to 'compete' and 'outperform' any prior work. We considered two types of sensor arrays - uniform and random; and two recovery approaches - Kronecker CS and group-sparse reconstruction. Experiments on real multi-spectra images show interesting results.

Index Terms— Multi-spectral Imaging, Demosaicing, Compressed Sensing.

1. INTRODUCTION

Multi-spectral imaging finds usage in a variety of applications in scientific, industrial and agricultural domains. However, multi-spectral imaging is an expensive venture; the cost associated with this imaging technique is the main deterrent in its widespread application in developing nations.

To have an idea about the cost of multi-spectral imaging consider the five sensor Condor-1000 MS5-VNN-285 which is a five band camera, with three bands as RGB and two bands in the infrared region of electromagnetic spectrum. It costs around 59000 dollars and the resolution is only 1360 × 1024. The cost of a single-sensor RGB camera with an order of magnitude more resolution is less than a 1000 dollars. In developing countries of south-east Asia, major parts of Africa and Latin America, such huge costs associated with multi-spectral acquisition devices hinder widespread applications of this technology. The objective of this work is to bring down the cost of such multi-spectral cameras.

Luckily for RGB cameras, the sensors are based on low-cost silicon technology. Till recently, the Moore's law for compacting silicon transistors held true, and RGB camera manufacturers were able to squeeze in more and more sensors per unit area thereby increasing the resolution of these cameras to astonishing heights. Unfortunately, the silicon sensor technology does not work outside the optical range. Thus multi-spectral sensor arrays are expensive and is the main contributing factor for the increased cost of such cameras.

Most multi-spectral cameras have a separate sensor array for each band of the EM-spectrum. Image acquisition using such separate sensor arrays requires large number of optical and mechanical parts which contributes to the increased cost and size of the cameras. Owing to such separate sensor arrays, there arises the problem of pixel registration.

In this work we propose a single sensor architecture for addressing the issue of cost, size and pixel registration. We draw inspiration from single-sensor architecture of modern RGB cameras in which only a single band is sampled at a particular location of the array. To get the full multi-spectral image one will need to interpolate the missing (unsampled) pixel values in different bands. This is generally referred to as the 'demosaicing problem'.

Demosaicing for RGB cameras is almost a solved problem [1]. Commercial RGB cameras uses a single sensor architecture on the Bayer pattern. The Bayer pattern has 50% sensors for the green band, and 25% sensors for red and blue bands. Since the Bayer pattern is uniform, linear interpolation techniques are quiet successful in solving the demosaicing problem. There are a handful of studies on the topic of compressed sensing (CS) based RGB demosaicing [2-5]. The uniform Bayer pattern is theoretical unsuitable for CS recovery; thus they propose replacing the traditional Bayer filter array by a panchromatic filter array. They show that such a filter array is amenable for color CS based demosaicing. The prior studies [2-5] are pedagogic exercises with limited practicality. Both the camera manufacturers and the consumers are satisfied with the current Bayer pattern and fast linear interpolation techniques and it is unlikely that the existing technology will be discarded.

The idea of employing single sensor architecture for multi-spectral cameras is borrowed from commercial RGB cameras. But there are only a handful of studies in this area [6-11]. All of them use uniform sampling patterns and use linear or kernel methods for demosaicing. We will briefly review them in the following section.

In this work we investigate the possibility of using uniform and randomized single sensor architecture for multi-spectral image acquisition and subsequent reconstruction employing CS based techniques. Such an investigative study has not been done before.

The rest of the paper is organized into several sections. The review of prior works in multi-spectral demosaicing is discussed Section 2. Section 3 proposes the CS based reconstruction techniques for uniform and randomized sampling patterns. The experimental results are discussed in Section 4. Finally the conclusions of the work are drawn in Section 5.

2. LITERATURE REVIEW

In the Bayer pattern, the Green channel is sampled more compared to the red and the blue channels; the logic being that the human eye is more receptive towards green compared to the other colors. Multi-spectral images are not acquired for the human eye; thus the question of importance sampling does not arise here. However in [6] a multispectral filter array design was proposed based on probability of appearance of each band for a target recognition problem. This design takes into consideration both spatial uniformity and spectral consistency during sampling process. Based on above filter array design, a generic demosaicing algorithm is proposed in [7]; this is called Binary Tree based Edge Sensing method. This method accounts for edge correlations among the multi-spectral channels while interpolating the missing band values.

However the applicability of the proposed filter array is limited. It is tailored for target detection problems. Moreover it requires prior knowledge regarding the importance of different bands. Even if we assume that this prior knowledge is available, the filter array is designed for splitting the sampling space in a dyadic scale; thus even the splitting is restrictive and cannot handle arbitrary prior probabilities. In spite of these shortcomings, in practice, this technique [6, 7] yields very good demosaicing results.

The studies [8, 9] attempt to solve a very specific demosaicing problem - that of Multi-spectral Color Filter Arrays (MFCA). It extends the same basic Bayer CFA (that can sample RGB channels) to incorporate two more channels - Orange and Cyan. Thus it can sample only 5 specific channels - Red, Green, Blue and Orange, Cyan. It should be noted that this architecture cannot be used for acquiring ANY 5 spectral channels.

In [8] a guided filter is used for demosaicing. The guided filter requires reference guide image and uses this image to interpolate the various channel images. In [8], the guide image is obtained from the most densely sampled channel, i.e. the green channel. This technique gives good results for their specific problem; but such tailored approaches are hard to be generalized.

In [9], the problem remains the same, but the reconstruction algorithm is slightly more generalized. It uses a Kernel interpolation method to find estimate the missing pixel values in different bands. It yields good results for their specific problem. In principle this approach can be generalized to arbitrary uniform sampling patterns.

3. BRIEF REVIEW OF COMPRESSED SENSING

Today Compressed Sensing (CS) is widely understood by the signal processing community and a discussion on this topic may be redundant; but for the sake of completion we briefly review this topic.

CS studies the recovery of sparse signals from its noisy lower dimensional measurements, i.e. solving a linear inverse problem of the following form

$$y_{m\times 1} = M_{m\times n}x_{n\times 1} + \eta_{m\times 1},\ \eta \sim N(0,\sigma^2) \qquad (1)$$

Here x is the signal of interest, M is the measurement operator and y is the sampled measurement vector. The measurement is supposed to be corrupted with Gaussian noise.

Natural signals are almost always never sparse in their physical domain, but most of the times they have a sparse representation in the transform domain, e.g. natural images are sparse in DCT, medical image are sparse in wavelet, seismic waves are sparse in curevelet, EEG signals are sparse in Gabor, etc. When the transform is an orthogonal or tight-frame, it is possible to express (1) in the following form,

$$y = M\Psi^T\alpha + \eta \qquad (2)$$

where Ψ is the sparsifying transform and α is the vector of sparse transform coefficients.

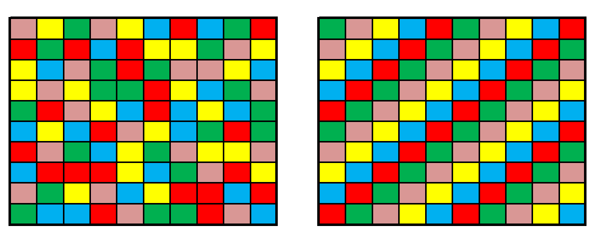
One can recover the sparse transform coefficients by solving the l_1-norm minimization problem.

$$\min_{\alpha} \|\alpha\|_1 \text{ subject to } \|y - M\Psi\alpha\|_2^2 \le \varepsilon \qquad (3)$$

where ε is the noise parameter.
Once α is obtained, the signal of interest can be recovered as follows:

$$x = \Psi^T\alpha \qquad (4)$$

The quality of CS reconstruction depends on two factors - i) the sparsity of the solution (α) and the incoherence between the measurement (M) and sparsifying basis (Ψ) [10]. Intuitively, we can understand that, sparser the solution better is the recovery. But the incoherence between the two basis should also be considered; a sparsity basis which is not incoherent with the measurement basis will ultimately lead to poor recovery.



(a) (b)

Fig. 1. (a) Random 5-channel Sampling Array; (b) Uniform 5-channel Sampling Array

4. PROPOSED DEMOSAICING METHOD

Our proposal has two parts: the first is the single array design and the second part is the CS recovery algorithms.

4.1. Filter Arrays

In the single sensor architecture, at each pixel location, only one of the channels is sampled. We propose two sampling patterns. The first one is a random sampling pattern. The sampling locations of each channel is distributed uniformly at random spatially. Fig. 1a. shows an example of such a 5-channel array. The design is shown in pseudo-colors and can refer to any spectrum channel.

The random array is conducive for CS recovery since the measurements are incoherent. The random array is easily extendable to any number of channels. This is a departure from prior studies [6-9], where the design was the array was not generalized and was designed to suit particular applications.

We also propose a uniform filter array. A five channel design is shown in Fig. 1b. In theory, such uniform sampling patterns are not conducive to CS recovery. But, since this is an exploratory work, we will try it nevertheless. The uniform filter design, shown here is also extendable to any number of channels.

4.2. Recovery Method

The acquisition process from such single sensor designs can be expressed as:

$$y_i = R_i x_i + \eta,\ i = 1...\text{C} \tag{5}$$

where C denotes the number of spectral channels

In multi-spectral imaging, the channels are correlated with each other. We intend to exploit the channel correlation during reconstruction.

4.1.1. Group Sparsity

In prior works [11-14], it was shown that the inter-channel correlation leads to row-sparsity for the color imaging problem. The same logic extends to multi-spectral imaging. Since the images are correlated, they have similar sparsity signature in the transform domain, i.e. the positions of the high valued sparse coefficients remain the same even the values of the coefficients might be different.

The data acquisition model (5) is modeled as following (6),

$$y_i = R_i \Psi \alpha_i + \eta,\ i = 1...\text{C} \tag{6}$$

This can be succinctly represented in the following fashion.

$$vec(y) = \Phi\ vec(\alpha) + \eta \tag{7}$$

where vec has the usual connotation $y = [y_1 | ... | y_C]$, $\alpha = [\alpha_1 | ... | \alpha_C]$ and $\Phi = BlockDiag(R_i \Psi), \forall i$.

According to the assumption of same sparsity signature, the solution α will be row-sparse. Thus the inverse problem (7) can be solved via,

$$\min_{\alpha} \|\alpha\|_{2,1} \text{ subject to } \|y - \Phi\alpha\|_2^2 \le \varepsilon \tag{8}$$

where $\|\alpha\|_{2,1} = \sum_j \|\alpha^{j\to}\|_2$, $\alpha^{j\to}$ denotes the j^{th} row of α.

The choice of such values for the $l_{2,1}$-norms can be understood intuitively. The l_2-norm over the rows enforces non-zero values on all the elements of the row vector whereas the summation over the l_2-norm enforces row-sparsity, i.e. the selection of few rows. Algorithm for solving this problem is available in [15].

The multi-spectral images are directly sub-sampled in the pixel domain; thus in this case, the measurement basis (R) is the Dirac / Identity basis. The main concern here is the choice of sparsifying transform Ψ. Wavelets lead to very sparse representation of natural scenes, but they are not very incoherent with the Dirac basis. The Fourier basis is maximally incoherent with the Dirac basis, but does not lead to a very sparse representation of natural scenes. We found that DCT is a good compromise. Since it is closely related to the Fourier basis, it is more incoherent with the Dirac basis than wavelets and leads to sufficiently sparse representations.

4.1.2. Kronecker Compressed Sensing [16]

There is yet another way to exploit the inter-channel correlation. Since, the channels are correlated, the variation at a particular pixel position is smooth along the channel. The smoothness leads to a compact representation in the Fourier domain. Thus, we can sparsify the multi-spectral image by exploiting intra-channel spatial redundancy and inter-channel correlation in the following form,

$$\beta = F^T \otimes \Psi\ vec(x) \tag{9}$$

where $x = [x_1 | ... | x_C]$.

The Kronecker product is a convenient notation to represent the operations along column and row directions. Since, both Ψ and the Fourier transform are orthogonal, (9) can be alternately represented as,

$$vec(x) = (F^T \otimes \Psi)^T \beta \tag{10}$$

Thus, we can represent (5) in the following form after incorporating the Kronecker product notation,

$$vec(y) = R\ vec(x) + \eta = R(F^T \otimes \Psi)^T \beta + \eta \tag{11}$$

where $R = BlockDiag(R_i), \forall i$.

The sparse transform coefficients can be solved via l_1-norm minimization (3). In this work, we actually use the Re-weighted l_1-norm [17] to achieve better results using the SPGL1 solver [18].

With the Kronecker Compressed Sensing (KCS) [16] formulation, we employ the Fourier sparsifying basis. Although Fourier basis alone does not yield a very sparse representation for each channel of the multi-spectral image, the Kronecker Fourier basis sufficiently sparsifies the full multi-spectral image. Also, the Fourier basis is an ideal choice in this scenario since it is maximally incoherent with the Dirac measurement basis.

5. EXPERIMENTAL EVALUATION

The experimental results were carried out on a well known multi-spectral imaging dataset [19]. The experiments were carried out on the four images - Balloon, Beads, Cloth and Flowers. The color image versions of these images are shown in Fig. 3.

Fig. 3. Left to Right - Balloon, Beads, Cloth and Flowers.

We carried out experiments for 3-channel and 4-channel reconstruction. We chose the first 3 and 4 channels respectively for this purpose. The accuracy for demosaicing is evaluated based on PSNR values between the original and the reconstructed. The results for 3-channel and 4-channel image demosaicing are shown in Tables 1 and 2 respectively. For random sampling, the sampling array was simulated 100 times for each experiment. The average recovery results are shown here. This needed to be done so as to get robust results.

Table 1. PSNR for 3-channel Reconstruction

Datasets	Random Pattern		Uniform Pattern	
	KCS	Group-sparse	KCS	Group-sparse
Balloon	38.18	38.05	39.04	31.01
Beads	28.78	25.78	29.7	22.74
Cloth	27.59	26.1	28.76	24.37
Flowers	32.85	32.48	33.14	28.43

Table 2. PSNR for 4-channel Reconstruction

Datasets	Random Pattern		Uniform Pattern	
	KCS	Group-sparse	KCS	Group-sparse
Balloon	37.21	35.82	34.34	13.88
Beads	24.54	22.91	25.25	15.96
Cloth	26.5	25.04	27.63	14.5
Flowers	31.33	30.25	28.27	18.3

The following conclusions can be drawn:

1. The demosaicing accuracy decreases as the number of channels increase - this is obvious. Since the sampling ratio decreases (1:4 from 1:3) with increasing number of channels; the recovery becomes progressively harder.
2. The KCS reconstruction always yields better reconstruction than group-sparse reconstruction. The difference between the two becomes more pronounced as the sampling ratio decreases.
3. The surprising observation is that, when the sampling ratio is relatively high (1:3), the uniform sampling array yields better results random sampling array for KCS reconstruction, but when the sampling ratio falls (1:4), the random array yields better results. This is not a fluke, since we simulated 100 random sampling masks for each experiment.
4. For group-sparse reconstruction, random array always yields better results than uniform arrays.
5. As the sampling ratio falls (from 1:3 to 1:4), group-sparse reconstruction for uniform sampling arrays yields very poor recovery.

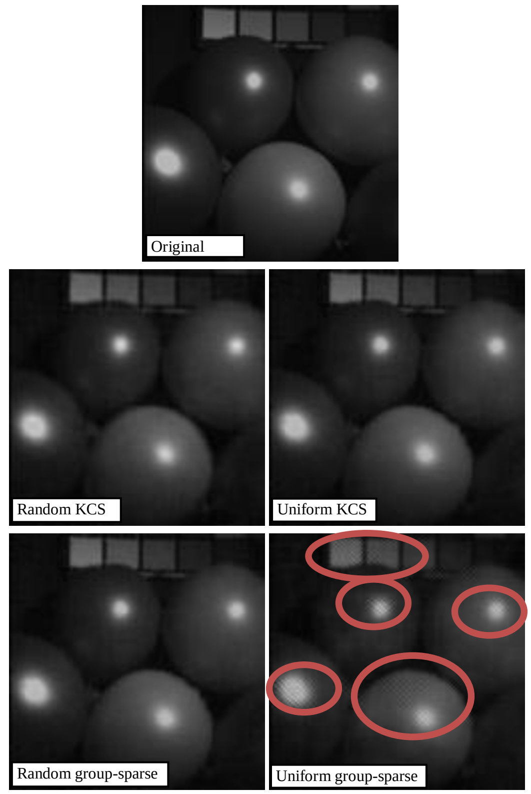


Fig. 4. Channel 2 - Original and Reconstructed

The original and reconstructed images of Channel 2 for 3-channel reconstruction is shown in Fig. 4. It can be seen that the 'Random KCS', 'Uniform KCS', 'Random group-sparse' almost yields similar reconstruction; there are nominal reconstruction artifacts. But 'Uniform group-sparse' reconstruction yields severe artifacts as has been outlined in the figure. The qualitative results corroborate the findings in Tables 1 and 2.

5. CONCLUSION

This is the first work that explores the problem of compressed sensing (CS) reconstruction of multi-spectral images acquired with a single sensor architecture. We propose two filter array designs - random and uniform. Both of them can be generalized to any number of channels. For reconstruction we propose two approaches; both of these exploit the inter-channel correlations while reconstruction. The first approach models the problem as a group-sparse optimization while the second approach proposed is a Kronecker CS formulation.

We believe in reproducible research. The code for reproducing the results is available at [20].

REFERENCES

[1] D. Menon and G. Calvagno, “Color image demosaicking: An overview,” Signal Processing: Image Communication, vol. 26, no. 89, pp. 518 –533, 2011.

[2] A. A. Moghadam, M. Aghagolzadeh, M. Kumar and H. Radha, "Compressive demosaicing," IEEE International Workshop on Multimedia Signal Processing (MMSP), October 2010.

[3] M. Aghagolzadeh, A. A. Moghadam, M. Kumar and H. Radha, "Bayer and panchromatic color filter array demosaicing by sparse recovery," IS&T/SPIE Electronic Imaging Conference, San Francisco, California, January 2011.

[4] A. A. Moghadam, M. Aghagolzadeh, M. Kumar and H. Radha, "Incoherent color frames for compressive demosaicing," in Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), May 2011.

[5] M. Aghagolzadeh, A. A. Moghadam, M. Kumar and H. Radha, "Compressive demosaicing for periodic color filter arrays," in Proceedings of International Conference on Image Processing (ICIP), September 2011.

[6] L. Miao and Q. Hairong. The design and evaluation of a generic method for generating mosaicked multispectral filter arrays. IEEE Transactions on Image Processing, 15(9):2780–2791, 2006.

[7] L. Miao, Q. Hairong, R. Ramanath, and W. E. Snyder. Binary tree-based generic demosaicking algorithm for multispectral filter arrays. IEEE Transactions on Image Processing, 15(11):3550–3558, 2006.

[8] Y. Monno, M. Tanaka and M. Okutomi, "Multispectral demosaicking using guided filter", Digital Photography VIII. Edited by Battiato, Sebastiano; Rodricks, Brian G.; Sampat, Nitin; Imai, Francisco H.; Xiao, Feng. Proceedings of the SPIE, Volume 8299, article id. 82990O, 7 pp. (2012)

[9] Y. Monno, M. Tanaka and M. Okutomi, "Multispectral demosaicking using adaptive kernel upsampling", IEEE ICIP, pp. 3157-3160, 2011.

[10] E. Candès and J. Romberg, "Sparsity and incoherence in compressive sampling", Inverse Problems, Vol. 23 (3), pp. 969 -, 2007.

[11] A. Majumdar, R. K. Ward and T. Aboulnasr, “Algorithms to Approximately Solve NP Hard Row-Sparse MMV Recovery Problem: Application to Compressive Color Imaging”, IEEE Journal on Emerging and Selected Topics in Circuits and Systems, Vol. 2 (3), pp. 362-369. 2013.

[12] A. Majumdar and R. K. Ward, “Compressed Sensing of Color Images”, Signal Processing, Vol. 90 (12), 3122-3127, 2010.

[13] A. Majumdar and R. K. Ward, “Compressive Color Imaging with Group Sparsity on Analysis Prior”, IEEE International Conference on Image Processing, pp. 1337-1340, 2010.

[14] A. Majumdar and R. K. Ward, “Non-convex Group Sparsity: Application to Color Imaging”, IEEE International Conference on Acoustics, Speech, and Signal Processing, pp. 469-472, 2010.

[15] A. Majumdar and R. K. Ward, “Synthesis and Analysis Prior Algorithms for Joint-Sparse Recovery”, IEEE International Conference on Acoustics, Speech, and Signal Processing, pp. 3421-3424, 2012.

[16] M. F. Duarte and R. G. Baraniuk, "Kronecker Compressive Sensing," IEEE Transactions on Image Processing, Vol.21 (2), pp.494,504, 2012.

[17] E. J. Candes, M. B. Wakin, and S. Boyd, "Enhancing Sparsity by Reweighted l1 Minimization", Journal of Fourier Analysis and Applications, Vol. 14(5), pp. 877-905, 2008.

[18] E. van den Berg and M. P. Friedlander, “Probing the Pareto frontier for basis pursuit solutions”, SIAM J. on Scientific Computing, Vol. 31(2), pp. 890-912, 2008.

[19] F. Yasuma, T. Mitsunaga, D. Iso, and S. Nayar, “Generalized Assorted Pixel Camera: Post-Capture Control of Resolution, Dynamic Range and Spectrum,” Department of Computer Science, Columbia University CUCS-061-08, Tech. Rep., Nov 2008.