Multi-Resolution Coded Apertures Based on Side Information for Single Pixel Spectral Reconstruction

Hans Garcia*, Claudia V. Correa$^{\diamond}$, Karen Sánchez*, Edwin Vargas* and Henry Arguello$^{\diamond}$
*Dept. of Electrical Engineering, $^{\diamond}$Dept. of Computer Science, Universidad Industrial de Santander, Colombia
E-mail: henarfu@uis.edu.co

***Abstract*—Compressive spectral imaging (CSI) architectures allow to reconstruct spectral images from a lower number of measures than the traditional scanning-based methods. In these architectures, the coded aperture design is critical to obtain high-quality reconstructions. The structure of coded apertures is traditionally designed without information about the scene, but recently side information-based architectures provide prior information of the scene, which enables adaptive coded aperture designs. This work proposes the development of an adaptive coded aperture design for spectral imaging with the single pixel camera, based on a multi-resolution approach. An RGB side image is used to define blocks of similar pixels, such that they can be used to design the coded aperture patterns. This approach improves the reconstruction quality in up to 23dB compared with traditional single pixel camera, and the computation time in up to 99.5% because it does not require an iterative algorithm.**

I. Introduction

Spectral imaging techniques capture a 3D data cube, with two spatial dimensions (x, y) and one spectral dimension (λ) along the electromagnetic spectrum. The spectral information enables object classification in areas such as artwork conservation [1], biomedical imaging [2] and food quality [3]. Conventional scanning-based spectral images (SI), such as whisk-broom or pushbroom spectrometers, involve massive amounts of data, which increase storing and processing costs. On the other hand, compressive spectral imaging (CSI) acquires the spatial/spectral information applying the principles established by compressive sensing (CS) [4], which allows SI retrieval from a lower number of measures than the traditional methods. Specifically, CSI establishes that it is possible to retrieve the spatial and spectral information from a small number of samples, under the assumption that it has a sparse representation in a given basis $\mathbf{\Psi}$. In particular, a SI $\mathbf{F} \in \mathbb{R}^{M\times N\times L}$, where M and N are the spatial dimensions and L represents the number of spectral bands, has a dispersion level S if its vector form $\mathbf{f} \in \mathbb{R}^{MNL}$ can be written as a linear combination of S vectors on any basis $\mathbf{\Psi}$, such that $\mathbf{f} = \mathbf{\Psi\theta}$ with $S \ll MNL$. This condition allows to recover $\mathbf{F}$ from $K \ll MNL$ projections.

CSI architectures often use 2D detectors [5], [6], [7], whose costs increase with the resolution. On other hand, the single pixel camera (SPC) [8] has been recently employed as a low-cost CSI sensor [9]. In particular, SPC uses a single point detector such as a whisk-broom spectrometer, and a coded aperture. Specifically, the SPC sensing process can be represented in matrix form as $\mathbf{g} = \mathbf{Hf}$, where $\mathbf{H}$ is the transfer function of the optical system [10], which is directly related to the coded aperture design.

Recent works have established the importance of coded aperture designs to retrieve high-quality SI using a small number of projections [10], [11]. More specifically, the traditional coded aperture (CA) design in [12] does not exploit prior information about the SI, so that only generic designs are possible, such as Hadamard, Bernoulli or uniform coded apertures [10]. On other hand, grayscale side information has been recently employed in coded aperture design for CASSI architecture[13]. This design improves the quality of the reconstructions because it takes into account the spatial distribution of the scene.

In this work, an adaptive coded aperture design for the single pixel camera based on an RGB image of the spectral scene is proposed, as illustrated in Fig. 1. Specifically, this coded aperture design takes advantage of the spatial similarities in the scene using a multi-resolution approach, and aims at obtaining high-quality and fast multi-resolution reconstructions.

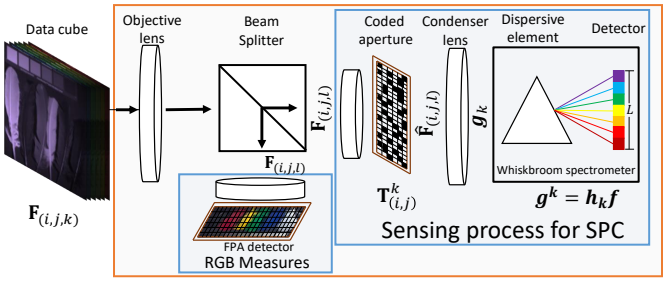

Fig. 1. Scheme of the SPC used to acquire the spatial and spectral information with RGB side information.

II. Side information-based SPC sensing model for Spectral Imaging

As illustrated in Fig. 1, the proposed system is composed by two arms, the first one related to the SPC and the other acquires the side information with an RGB sensor.

A. Single pixel camera sensing model

The single pixel architecture illustrated in Fig. 1 employs a coded aperture $\mathbf{T}^k_{(i,j)}$ that spatially modulates all the spectral bands from the input scene $\mathbf{F}_{(i,j,l)}$ with the same pattern, where i, j index the spatial coordinates, l accounts for the spectral band, and k indexes each captured snapshot. Specifically,

ISBN 978-90-827970-1-5

the coded aperture $\mathbf{T}^k_{(i,j)}$ is a binary pattern whose spatial distribution determines the quality of the reconstructed scene. In this case, the binary levels are either 1 or -1. In practical terms, the modulation effect caused by the -1 entries can be implemented by acquiring a measurement with an all ones coded aperture and subtract it from each captured snapshot. Mathematically, the coded aperture effect over the input scene is represented as

$$\hat{\mathbf{F}}^k_{(i,j,l)} = \mathbf{F}_{(i,j,l)}\mathbf{T}^k_{(i,j)}. \tag{1}$$

Then, the modulated scene $\hat{\mathbf{F}}$ is concentrated in a single spatial point by the condenser lens, and captured by a whisk-broom spectrometer. This detector splits the incoming light rays in a discrete measure of the spectrum for each band as

$$g^k_l = \sum_i \sum_j \mathbf{F}_{(i,j,l)}\mathbf{T}^k_{(i,j)}, \tag{2}$$

for $i = 1, 2, \ldots, M$, $j = 1, 2, \ldots, N$, and $l = 1, 2, \ldots, L$. The acquisition system can also be modeled as

$$g^k_l = \mathbf{h}^T_k \mathbf{f}_l, \tag{3}$$

where $\mathbf{h}_k$ is the vectorization of the coded aperture for each snapshot k, $\mathbf{f}_l$ is the vectorization of the l-th band of the SI $\mathbf{F}$. In general, the sensing model for all shots captured for the l-th band can be written as

$$\mathbf{g}_l = \mathbf{H}\mathbf{f}_l, \tag{4}$$

with $\mathbf{g}_l = \left[g^1_l, \ldots, g^K_l\right]^{\mathrm{T}}$, $\mathbf{H}$ is a $K \times MN$ matrix given by $\mathbf{H} = \left[\mathbf{h}^T_1, \ldots, \mathbf{h}^T_K\right]$, and K is the number of shots. It is important to remark that each shot employs a different coded aperture pattern. Furthermore, the sensing model for all the spectral bands and K shots is given by

$$\mathbf{g} = \hat{\mathbf{H}}\mathbf{f}, \tag{5}$$

where $\mathbf{g} = \left[(\mathbf{g}_0)^{\mathrm{T}}, \ldots, (\mathbf{g}_{L-1})^{\mathrm{T}}\right]^{\mathrm{T}}$, $\hat{\mathbf{H}}$ is the sensing matrix whose structure is illustrated in Fig. 2, which can be obtained as a block diagonal matrix

$$\hat{\mathbf{H}} = \mathbf{I}_L \otimes \mathbf{H}, \tag{6}$$

where $\mathbf{I}_L$ is an $L \times L$ identity matrix, such that the number of columns and rows of $\hat{\mathbf{H}}$ is MNL and γMNL, respectively, with $\gamma = \frac{K}{MN}$ as the compression ratio, taking values $\gamma \in [0, 1]$.

B. RGB Camera sensing model for spectral imaging

The model used for acquiring the RGB image assumes that the detector has an equal response for all spectral bands in each spectral channel (Red, green and blue). Therefore, the acquisition of the RGB image can be modeled as the linear system given by

$$\bar{\mathbf{g}} = \bar{\mathbf{H}}\mathbf{f}, \tag{7}$$

where $\bar{\mathbf{H}} = \mathbf{I}_3 \otimes \left(\mathbf{1}_{1\times L/3} \otimes \mathbf{I}_{MN}\right)$. Figure 3 shows an example of the structure of $\bar{\mathbf{H}}$, where white squares represent a 1 value for the color indicated by the border, and gray zones are zero. The acquired RGB image has an important role for this work, because it is used to design the coded aperture patterns. Furthermore, it can be also included in the reconstruction problem to improve the quality of the recovered data cube.

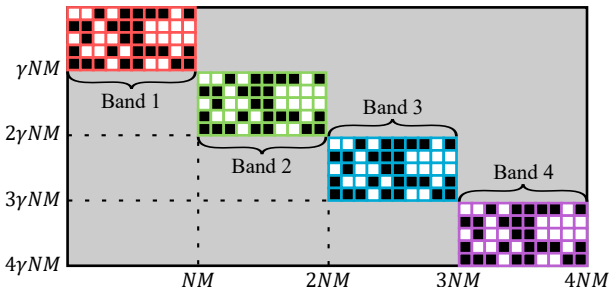

Fig. 2. SPC sensing matrix $\hat{\mathbf{H}}$, with $M = 2$, $N = 5$, $\gamma = 0.5$ and $L = 4$. White squares represent a positive value 1 according to the coded aperture design, black squares represent -1, and gray zones are 0.

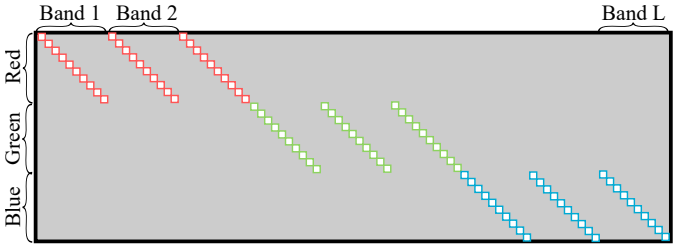

Fig. 3. RGB sensing matrix $\bar{\mathbf{H}}$, for $M = 2$, $N = 5$, and $L = 9$. White squares represent 1 and gray zones are 0.

III. Coded Aperture Design Based on Super-Pixels

The input for the coded aperture design is the RGB image captured as in (7). Specifically, it is used to identify the uniform zones in the scene, such that similar pixels can be grouped into a super-pixel. The resulting map of super-pixels is then used to design the coded aperture pattern. The main goal is to obtain multi-resolution (MR) coded apertures whose features match with the super-pixel map.

First, to determine the super-pixel map, a clustering method such as the SLIC algorithm [14] is employed. A simple example of this process is illustrated in Fig. 4, where each color represents a super-pixel. Note that the color super-pixel map is decomposed as β coded aperture patterns, where β is the desired number of super-pixels, and it is a required input for the SLIC method. To obtain the set of CA $\mathbf{T}$, the map of super-pixels is used to generate a multi-resolution decimation matrix $\mathbf{\Delta}$, as in [15], where each row of $\mathbf{\Delta}$ represents the integration of pixels into each super-pixel. It is worth noting that the size of $\mathbf{\Delta}$ depends on the β as $\mathbf{\Delta} \in \mathbb{R}^{\beta \times MN}$. Moreover, the coded aperture design is directly related to the design of the matrix $\mathbf{H}$ from (4) which, in turn, determines the sensing matrix $\hat{\mathbf{H}}$ in (6).

ISBN 978-90-827970-1-5 © EURASIP 2018

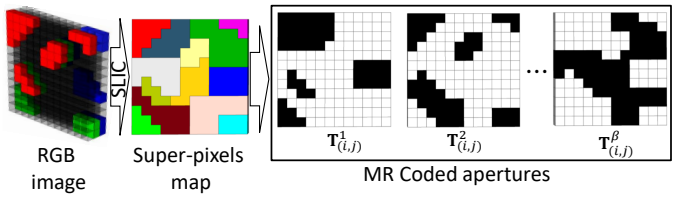

Fig. 4. Scheme of the process to generate MR coded apertures. Each super-pixel of the RGB scene is associated to a single coded aperture block.

Several SPC snapshots are required in order to provide good quality reconstructions. In this work, we propose to capture β measurements, in other words, the number of snapshots is equal to the number of desired super-pixels. This condition imposes that

$$rank(\mathbf{H}) = \beta. \tag{8}$$

Therefore, a full rank matrix $\mathbf{W}$ is used to design the sensing matrix for each spectral band as

$$\mathbf{H} = \mathbf{W}\mathbf{\Delta}. \tag{9}$$

Specifically, $\mathbf{W} \in \mathbb{R}^{\beta\times\beta}$ is a Hadamard matrix, because its inverse is the same transposed escalated matrix. This property enables the fast multi-resolution reconstruction approach proposed in Section V.

Coded apertures for each snapshot are obtained by rearranging each row of $\mathbf{H}$ in (9) as a matrix. The resulting MR coded apertures, whose block features match the super-pixels from $\mathbf{\Delta}$ are also illustrated in Fig. 4.

Furthermore Fig. 5(b) and (c) show two MR coded apertures obtained with the proposed design for the scene in Fig. 5(a).

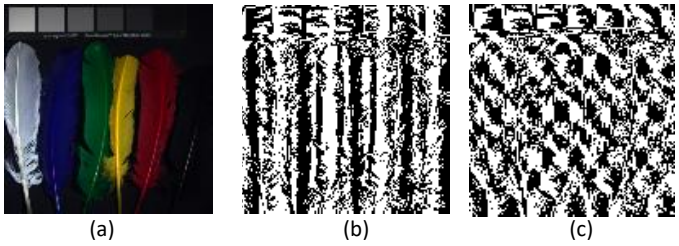

Fig. 5. Examples of the designed MR coded apertures for the spectral image in (a), using the proposed approach on (9).

IV. Reconstruction with side information

In order to reconstruct the spatio-spectral data cube, the SPC measurements from (5) are stacked with those from the RGB camera in (7), as well as the corresponding sensing matrices

$$\tilde{\mathbf{H}} = \begin{bmatrix}\hat{\mathbf{H}}\\ \bar{\mathbf{H}}\end{bmatrix}, \quad \tilde{\mathbf{g}} = \begin{bmatrix}\mathbf{g}\\ \bar{\mathbf{g}}\end{bmatrix}. \tag{10}$$

Note that $\hat{\mathbf{H}}$ is obtained as in (6), using the matrix $\mathbf{H}$ from (9).

Then, the spatio-spectral data cube can be obtained by solving the regularization problem given by

$$\tilde{\mathbf{f}} = \mathbf{\Psi}\{argmin_{\tilde{\theta}}||\tilde{\mathbf{H}}\mathbf{\Psi}\tilde{\theta} - \tilde{\mathbf{g}}||_2^2 + \tau||\tilde{\theta}||_1\}, \tag{11}$$

where τ is a regularization parameter, $\tilde{\theta}$ is the sparse representation of $\tilde{\mathbf{f}}$ in the basis $\mathbf{\Psi}$. It is important to remark that the new compression ratio $\bar{\gamma}$, taking into account the RGB image is given by

$$\bar{\gamma} = \frac{\gamma(L+3)}{L}, \tag{12}$$

where $\bar{\gamma} \approx \gamma$ as L increases.

V. Fast multi-resolution (FMR) reconstruction

Several applications require fast SI reconstructions, however, traditional reconstruction methods solve the $l_2 - l_1$ norm optimization in (11) whose complexity depends on the data dimensions. To date, fast reconstruction alternatives have been developed [16]. In this work, we propose to obtain a fast multi-resolution reconstruction based on the coded aperture design from section III. FMR reconstructions of the SI can be obtained without solving a minimization problem, this is possible by taking advantage of the designed coded aperture with a Hadamard structure. Using this approach, the complexity to obtain the SI is reduced only to one matrix product. First, taking into account that for some Hadamard matrix $\mathbf{W} \in \mathbb{R}^{\beta\times\beta}$

$$\mathbf{W}^T\mathbf{W} = \beta\mathbf{I}_\beta, \tag{13}$$

where $\mathbf{I}_\beta$ is a $\beta\times\beta$ identity matrix. If we multiply (4) by $\mathbf{W}^T$, and replacing $\mathbf{H}$ as in (9) we obtain that

$$\bar{\mathbf{f}}_l = (1/\beta)\mathbf{W}^T\mathbf{g}_l = (1/\beta)\mathbf{W}^T\mathbf{W}\mathbf{\Delta}\mathbf{f}_l. \tag{14}$$

Then, multiplying $\bar{\mathbf{f}}$ by some matrix $\hat{\mathbf{\Delta}}$, such that $\hat{\mathbf{\Delta}}\mathbf{\Delta} \approx \mathbf{I}$ yields

$$\tilde{\mathbf{f}}_l = \hat{\mathbf{\Delta}}\bar{\mathbf{f}}_l = (1/\beta)\hat{\mathbf{\Delta}}\mathbf{W}^T\mathbf{W}\mathbf{\Delta}\mathbf{f}_l \approx \mathbf{f}_l. \tag{15}$$

Thus, the best option is to have $\hat{\mathbf{\Delta}}$ as the transpose of $\mathbf{\Delta}$, with normalized rows. Then, a FMR reconstruction $\tilde{\mathbf{f}}$ of the SI is obtained with high quality at low computational complexity, when $\mathbf{W}$ is a Hadamard matrix and $\mathbf{H}$ follows the design in (9).

VI. Simulations and Results

Several simulations were realized to test the performance of the proposed approach when the number of bands L for the SPC measures is varied from 6 to 30, for the compression ratio $\gamma = \{0.03, 0.07, 0.25\}$, i.e. approximately $512, 1024$ and 4096 measures, respectively. Simulations use two data cubes with 128×128 pixels of spatial resolution [17]. Results are compared with respect to three different approaches. In the first approach, the RGB image is not stacked to the SPC measures, i.e. the classical SPC, using a traditional coded aperture design based on randomly permuted rows of the Hadamard transform with unit-norm columns [12]. The second approach uses the proposed architecture stacking the RGB image to the SPC measures (SPC + RGB) as in (10) with traditional Hadamard-based patterns. The third approach uses the coded aperture

ISBN 978-90-827970-1-5 © EURASIP 2018

design proposed in section III and the RGB image is stacked to the SPC measures as in (10) (RGB + SPC + CA design). These three approaches obtain the data cube reconstructions by solving (11) using the corresponding matrices. The fourth approach uses the same configuration as the third approach but recovers $\tilde{\mathbf{f}}$ with the fast multi-resolution (FMR) reconstruction proposed in (15).

For each simulation, SPC measures are obtained using the model in (5), and the RGB measures are obtained using the model in (7). In all cases (except for the case when the FMR reconstruction is obtained) the SI is recovered solving the minimization problem in (11) through the gradient projection for sparse reconstruction algorithm (GPSR) [18]. The basis representation used in this paper is $\mathbf{\Psi} = \mathbf{\Psi}_1 \otimes \mathbf{\Psi}_2$, where $\mathbf{\Psi}_1$ is a 2D wavelet Symmlet 8 basis, and $\mathbf{\Psi}_2$ is a 1D discrete cosine transform (DCT). The comparisons are expressed in terms of peak signal-to-noise ratio (PSNR). All simulations were conducted and timed using an Intel Core i7-6700 @3.40GHz processor, and 32GB RAM memory.

In order to illustrate the improvement of the proposed approach, Fig. 6 shows the quality of the reconstruction for each value of L and the two scenes, using each approach, where the quality improvements for the SPC+RGB+CA design are clearly noticeable. Moreover, the FMR reconstruction is better than the non-designed traditional sensing matrix for SPC+RGB in some cases. This is because the FMR ensures edge recovery.

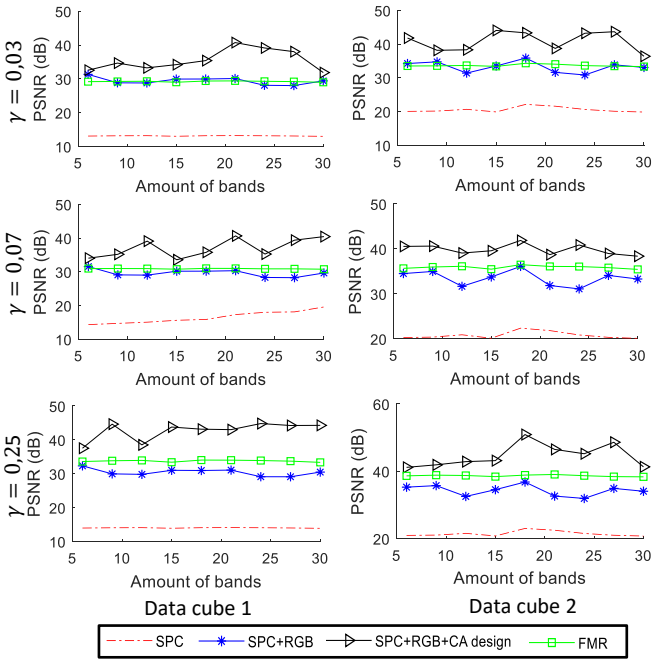

Fig. 6. Comparison of the reconstruction quality using the different proposed approaches and the traditional SPC for the two scenes, and three different values of γ.

On the other hand, when we focus on the computation time to recover the SI, Fig. 7 shows the comparison of time required by each approach including the calculation of the super-pixel map (when required). It is easy to see that the fast multi-resolution (FMR) approach is the faster option, although it does not provide the best reconstruction quality. However, it provides a good trade off between time and quality, being the faster approach and the second best quality. Furthermore, this approach uses the same measurements than the proposed approach, which obtains the highest PSNR (SPC + RGB + CA design). This is an interesting property, because with the same measurements it is possible to obtain fast, and good quality images when there is enough time for computation.

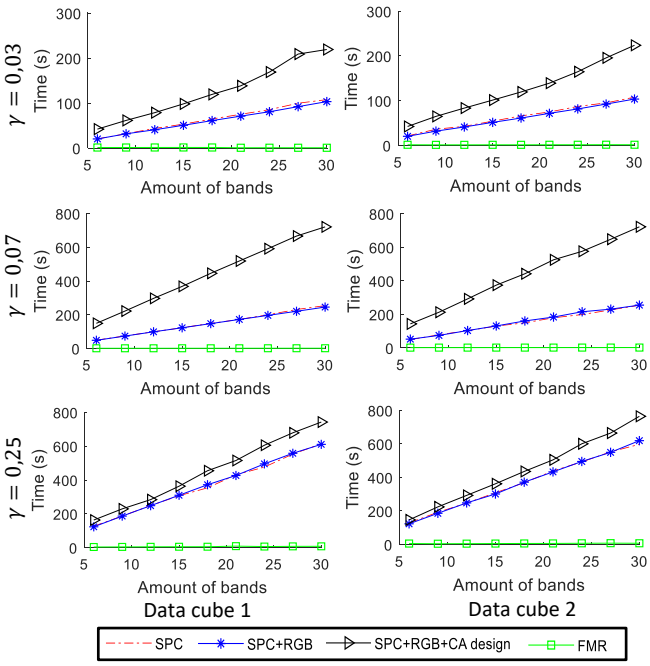

Fig. 7. Comparison of computation time to reconstruct the SI using the different proposed approaches and the traditional SPC for the two scenes, and three different values of γ.

Moreover, figures 8 and 10 compare the RGB mappings of the reconstructed data cubes, where it can be easily noticed that the proposed reconstruction approaches provide quality improvements. Furthermore, Fig. 9 shows a comparison between the spectral signatures of two points from data cube 1, for each reconstruction approach, to show that the proposed approaches reconstruct a closer spectral signature to the ground truth.

VII. Conclusion

This paper presented an adaptive coded aperture design for the single pixel camera. The proposed design uses a side RGB image of the scene to generate a super-pixel map that to generate multi-resolution coded apertures based on pixel spectral similarities. Simulations show that the proposed coded apertures improve the reconstruction quality in up to 23dB

ISBN 978-90-827970-1-5 © EURASIP 2018

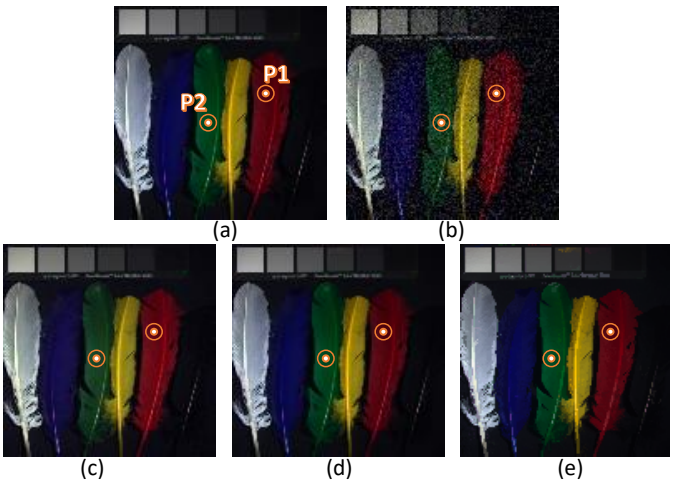

Fig. 8. RGB Comparison between the ground truth with the different reconstruction approaches for the data cube 1 with spatial dimensions 128×128, $L = 30$ and $\gamma = 0.07$. (a) Ground truth; and the reconstructions using (b)SPC, (c) SPC + RGB, (d) SPC + RGB + CA design, and (e) FMR.

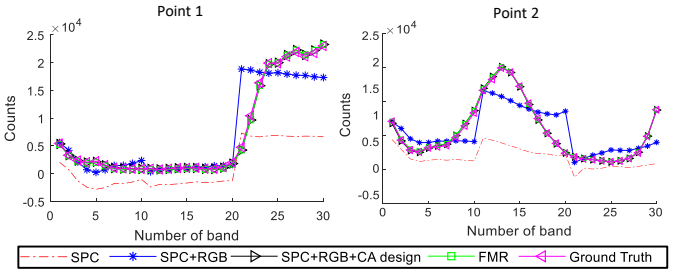

Fig. 9. Comparison of the spectral signatures of the recovered data cube 1 with $L = 30$ using the traditional and proposed approaches for the two spatial points highlighted in Fig. 8.

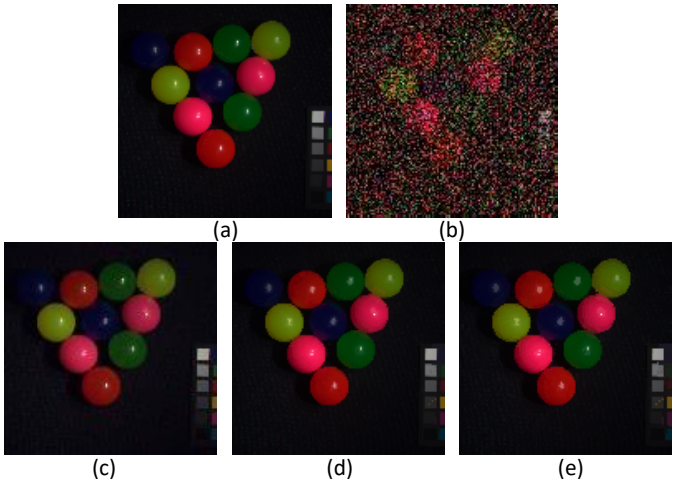

Fig. 10. RGB Comparison between the ground truth with the different reconstruction approaches for the data cube 2 with spatial dimensions 128×128, $L = 30$ and $\gamma = 0.07$. (a) Ground truth; and the reconstructions using (b)SPC, (c) SPC +RGB, (d) SPC + RGB + CA design, and (e) FMR.

of PSNR with respect to the traditional single pixel camera. In addition, a non-iterative fast multi-resolution reconstruction method, motivated by the proposed CA, has been proposed. Simulation results show that this non-iterative reconstruction is up to 99.5% faster than traditional iterative methods, at the expense of lower image quality.

REFERENCES

[1] A. Pelagotti, A. D. Mastio, A. D. Rosa, and A. Piva, "Multispectral imaging of paintings," *IEEE Signal Processing Magazine*, vol. 25, no. 4, pp. 27–36, 2008.

[2] G. Lu and B. Fei, "Medical hyperspectral imaging: a review." *Journal of biomedical optics*, vol. 19, no. 1, p. 10901, 2014.

[3] N. N. Wang, D. W. Sun, Y. C. Yang, and H. Pu, "Recent Advances in the Application of Hyperspectral Imaging for Evaluating Fruit Quality." *Food Anal Meth*, vol. 8 SRC - G, pp. 1–14, 2015.

[4] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[5] X. Lin, Y. Liu, J. Wu, and Q. Dai, "Spatial-spectral encoded compressive hyperspectral imaging," *ACM Trans. Graph.*, vol. 33, no. 6, pp. 233:1–233:11, 2014.

[6] X. Lin, G. Wetzstein, Y. Liu, and Q. Dai, "Dual-Coded Compressive Hyper-Spectral Imaging," *Classical Optics 2014, OSA Technical Digest (online) (Optical Society of America, 2014)*, vol. 39, no. 7, pp. 2044–2047, 2014.

[7] X. Cao, H. Du, X. Tong, Q. Dai, and S. Lin, "A prism-mask system for multispectral video acquisition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 12, pp. 2423–2435, 2011.

[8] M. Duarte, M. Davenport, D. Takhar, J. Laska, T. S. T. Sun, K. Kelly, and R. Baraniuk, "Single-Pixel Imaging via Compressive Sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 1–19, 2008.

[9] H. F. Rueda, C. A. Vargas, and H. Arguello, "Single-pixel optical sensing architecture for compressive hyperspectral imaging Arquitectura óptica de único pixel para el sensado compresivo de imágenes hiperespectrales," *Revista Facultad de Ingeniería Universidad de Antioquia*, vol. 75, pp. 134–143, 2014.

[10] C. V. Correa, H. Arguello, and G. R. Arce, "Spatiotemporal blue noise coded aperture design for multi-shot compressive spectral imaging," *J. Opt. Soc. Am. A*, vol. 33, no. 12, pp. 2312–2322, 2016.

[11] H. Arguello and G. R. Arce, "Rank minimization code aperture design for spectrally selective compressive imaging," *IEEE Transactions on Image Processing*, vol. 22, no. 3, pp. 941–954, 2013.

[12] G. Warnell, S. Bhattacharya, R. Chellappa, and T. Basar, "Adaptive-Rate Compressive Sensing Using Side Information," *IEEE Transactions on Image Processing*, vol. 24, no. 11, pp. 3846 – 3857, 2014.

[13] L. Galvis, D. Lau, X. Ma, H. Arguello, and G. R. Arce, "Coded aperture design in compressive spectral imaging based on side information," *Applied Optics*, vol. 56, no. 22, p. 6332, 2017.

[14] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Susstrunk, "SLIC Superpixels," Tech. Rep. June, 2010.

[15] H. Garcia, C. V. Correa, O. Villarreal, S. Pinilla, and H. Arguello, "Multi-Resolution Reconstruction Algorithm for Compressive Single Pixel Spectral Imaging," in *25th European Signal Processing Conference (EUSIPCO)*, Kos Island, Greece, 2017, pp. 498–502.

[16] H. Arguello, C. V. Correa, and G. R. Arce, "Fast lapped block reconstructions in compressive spectral imaging," *Appl. Opt.*, vol. 52, no. 10, pp. D32–D45, 2013.

[17] F. Yasuma, T. Mitsunaga, D. Iso, and S. K. Nayar, "Generalized Assorted Pixel Camera : Postcapture Control of Resolution , Dynamic Range , and Spectrum," *IEEE Transactions on Image Processing*, vol. 19, no. 9, pp. 2241–2253, 2010.

[18] M. A. T. Figueiredo, R. D. Nowak, and S. J. Wright, "Gradient Projections For Sparse Reconstruction: Application to Compressed Sensing and Other Inverse Problems," *J. Selected Topics in Signal Processing, IEEE*, vol. 1, no. 1, pp. 586–598, 2007.

ISBN 978-90-827970-1-5