COMPRESSIVE RECONSTRUCTION IN RAMAN HYPERSPECTRAL IMAGING USING PIXEL UNDERSAMPLING

Arijit Pramanik[1]*, Sudip K. Deb*[2] *and Ajit Rajwade*[1]

[1]Department of Computer Science and Engineering, IIT Bombay
[2]Centre for Research in Nanotechnology and Science, IIT Bombay
arijitp@cse.iitb.ac.in; sudip.deb@iitb.ac.in; ajitvr@cse.iitb.ac.in

ABSTRACT

Raman hyperspectral imaging has a number of applications in diverse fields such as material science, medicine and geology. However the image acquisition of Raman hyperspectral images is known to be a sequential and extremely time-consuming process, depending upon the detector efficiency. In this paper we describe a straightforward, yet effective method that can potentially enable much faster acquisition. The method involves measuring only a subset of the pixels of the complete image, where the subset is chosen either at uniform intervals or else randomly. The reconstruction of the complete image can then proceed by means of an on-the-fly dictionary-based inpainting algorithm, where the dictionary and the sparse codes are inferred on the fly from the undersampled acquisitions. Our proposed architecture is considerably simpler than existing architectures developed for compressive Raman image acquisition. We demonstrate encouraging reconstruction results on several real datasets.

Index Terms— Raman Hyperspectral Imaging, Compressive Sensing, Spatial Undersampling

1. INTRODUCTION

Raman spectroscopy is a non-destructive, non-invasive method that can provide important information about various materials ranging from biological tissues [1] to materials in pharmacy, chemistry or materials science [2], [3], [4], [5], to art preservation [6]. Moreover, Raman imaging does not require usage of any additional chemical agents for the acquisition, and provides very high spatial resolution. Raman spectroscopy is also employed in the field of geology [7] for determining mineral composition of rocks as well as in the field of semiconductors [8] for probing into their structure and examining the presence of traces of other materials integrated into it. However, the Raman signal is usually weak (one in one hundred million incident photons). The signal strength can be very low for some sample types and 'requires powerful excitation compatible with the sample damage threshold' [9]. The absence of powerful enough illumination must be compensated for by significantly higher acquisition times for Raman spectra, so as to avoid the inherent shot noise which can lower the signal to noise ratio (SNR). This is particularly the case if the CCD detectors have low efficiency, and higher efficiency detectors significantly raise the instrument cost.

There are two main approaches to Raman spectra image acquisition - serial and direct (also called snapshot). In the former case, the images are acquired sequentially in a point-wise manner (i.e. the entire spectrum is acquired one pixel at a time), or using line mapping where a 2D spatio-spectral signal is acquired across a single line. In the point mapping approach, the acquisition time is especially high. For a typical time of 1 second per spectrum, a 1 megapixel image can take days for acquisition. In the direct approach, all spatial points are simultaneously imaged for a given wavelength, which is more efficient. In this case, however, the acquisition time is directly proportional to the number of wavelengths.

Compressive imaging has emerged as a flourishing subfield of signal and image processing in recent times [10]. It involves acquisition of images directly in a compressed format, followed by conversion of the compressive measurements to the conventional image format typically via efficient convex optimization procedures. These procedures typically exploit the inherent sparsity or compressibility of many modalities of images in well-known orthonormal bases such as the wavelet or discrete cosine transform bases [11]. The emphasis of compressive imaging is on saving acquisition time. The time-intensive acquisition of Raman images, as well as their inherent smoothness as signals, renders Raman imaging as a potential area for applications of compressed sensing.

Contributions and Paper Organization: In this paper, we propose an idea for a compressive Raman imaging architecture, which involves measuring the spectra at only a subset of the total pixels of the image. The subset can be chosen randomly or in a structured fashion. In either case, this has the potential to immediately improve acquisition time. The missing pixels can be filled in, using an application of blind compressed sensing (BCS) [12] to implement an inpainting procedure. We present this formulation and the algorithm in Sec. 2. Some prior work in the area of compressive Raman imaging is summarized in Sec. 3. Experimental results on

real datasets are presented in Sec. 4. We conclude in Sec. 5.

2. PROBLEM FORMULATION

Let the image of interest be denoted as H, having size $N_x \times N_y \times N_\lambda$, where N_λ is the number of wavelengths. As per our architecture, the complete spectrum (of N_λ values) will be measured at only a fraction of the $N_x N_y$ pixels. No measurements are made at other pixels. Let the acquired incomplete image be denoted as G. The missing pixels in G need to be estimated via an efficient algorithm. For this purpose, we consider dividing H (and correspondingly G) into a number of overlapping patches, each of size $p \times p$. Let $\boldsymbol{h_i}, \boldsymbol{g_i}$ be the patches in H and G respectively, at location indexed as i, expressed as $p^2 \times 1$ vectors. Then we have the following:

$$\boldsymbol{g_i} = \boldsymbol{\Phi_i} \boldsymbol{h_i} + \boldsymbol{\eta_i}, \tag{1}$$

where $\boldsymbol{\eta_i}$ is the noise vector at location i and $\boldsymbol{\Phi_i}$ is a $p^2 \times p^2$ diagonal sensing matrix such that $\Phi_{i,jj}$ (the j^{th} diagonal element in $\boldsymbol{\Phi_i}$) contains 1 if g_{ij} is measured and 0 otherwise. If we follow a random sampling pattern, then the sensing matrices for every patch will be different. However if the sampling pattern bears more regularity, i.e. if it measures only every l^{th} pixel ($l > 0$) in both directions (often termed 'decimation'), then there will be only p^2 different sensing matrices. In this case, the reconstruction is akin to an image super-resolution problem.

Let n be the total number of patches thus considered. The task now reduces to estimating $\{\boldsymbol{h_i}\}_{i=1}^n$ given $\{\boldsymbol{h_i}, \boldsymbol{\Phi_i}\}_{i=1}^n$. We now frame this as a BCS problem. For this, we express each patch $\boldsymbol{h_i}$ as a sparse linear combination of columns of a dictionary matrix $\boldsymbol{A}$ of size $p^2 \times K$, i.e. $\boldsymbol{h_i} = \boldsymbol{A s_i}$ where $\boldsymbol{s_i}$ is a vector of sparse coefficients. Typical choices of dictionaries would include the wavelet or discrete cosine transform, since image patches are sparse (or approximately sparse) in these bases. However, given the inherently non-negative nature of the data, we impose the constraint that both $\boldsymbol{A}$ and $\boldsymbol{s_i}$ are both element-wise non-negative. With this in mind, we now seek to minimize the following objective function:

$$\begin{aligned} J(\boldsymbol{A}, \{\boldsymbol{s_i}\}_{i=1}^n) = \|\boldsymbol{g_i} - \boldsymbol{\Phi_i A s_i}\|^2 + \lambda \|\boldsymbol{s_i}\|_1, \\ \text{such that } \boldsymbol{A} \succeq \boldsymbol{0}, \forall i \boldsymbol{s_i} \succeq \boldsymbol{0}, \\ \forall j \in \{1, ..., K\}, \|\boldsymbol{A}_{\boldsymbol{.,j}}\|_2^2 = 1 \end{aligned} \tag{2}$$

where $\lambda > 0$ is a sparsity-promoting parameter, $\boldsymbol{A}_{\boldsymbol{.,j}}$ is the j^{th} column of $\boldsymbol{A}$, $\boldsymbol{0}$ represents a zero-valued matrix or vector, and $\succeq$ represents an element-wise 'greater than' inequality. This objective function seeks to solve a compressive version of the popular non-negative sparse coding (NNSC) algorithm [13]. In this work, we choose $K \ll p^2$ since a higher K increases the number of degrees of freedom, and the size of most available Raman spectral images is usually very small.

We implement the minimization of the function in Eqn. 3 using alternating minimization on the dictionary and sparse codes, starting from a random non-negative dictionary. Each step of the minimization is performed using projected gradient descent with adaptive step-size. That is, the stepsize of the gradient descent is adaptively chosen to ensure decrease of the objective function after imposition of all the constraints in Eqn. 3. The procedure is iterated till convergence, which is guaranteed due to the biconvex nature of the objective function. The algorithm is summarized in Alg. 1. Once the individual patches $\boldsymbol{h_i} = \boldsymbol{A s_i}$ are reconstructed, an estimate of H is assembled by sliding window averaging.

2.1. Dictionary Inference for Regular Sampling

This problem is equivalent to performing super-resolution on the acquired image. In this case, we do not initialize $\boldsymbol{A}$ randomly. A much better initial guess is required in this case, because the sensing matrices for applications such as super-resolution tend to have high coherence with typical dictionaries (due to the regularity of the sampling patterns). Rather, we first perform bicubic interpolation on the image G to yield $\tilde{G}$. We then infer the dictionary $\boldsymbol{A_{\tilde{G}}}$ from the patches of $\tilde{G}$ by optimizing the following objective function:

$$\begin{aligned} J_2(\boldsymbol{A_{\tilde{G}}}, \{\boldsymbol{\tilde{s_i}}\}_{i=1}^n) = \|\boldsymbol{\tilde{g_i}} - \boldsymbol{A_{\tilde{G}} \tilde{s_i}}\|^2 + \lambda \|\boldsymbol{\tilde{s_i}}\|_1, \\ \text{such that } \boldsymbol{A_{\tilde{G}}} \succeq \boldsymbol{0}, \forall i \boldsymbol{\tilde{s_i}} \succeq \boldsymbol{0}, \\ \forall j \in \{1, ..., K\}, \|\boldsymbol{A_{\tilde{G}_{.,j}}}\|_2^2 = 1. \end{aligned} \tag{3}$$

This now acts as an initial guess for inferring the actual dictionary $\boldsymbol{A}$ from H via Eqn. 3.

Algorithm 1 INFER $\boldsymbol{A}$,$\boldsymbol{S} = \{\boldsymbol{s_i}\}_{i=1}^n$ from $\boldsymbol{G}$

Require: $\boldsymbol{\mu, \lambda, K, p, n, \{\Phi_i\}_{i=1}^N, \epsilon, \alpha}$
- **if** $\boldsymbol{G}$ obtained from $\boldsymbol{H}$ by random sampling **then**
  - $\boldsymbol{A} \leftarrow random(p^2 N_\lambda, K)$
- **else**
  - Initialize $\boldsymbol{A} = \boldsymbol{A_{\tilde{G}}}$ inferred from $\boldsymbol{\tilde{G}}$
- **end if**
- $\boldsymbol{S} \leftarrow random(K, n)$

Ensure: $\boldsymbol{A} \succeq \boldsymbol{0}$ **and** $\|\boldsymbol{A_j}\|_2^2 = 1 \forall j \in \{1, .., K\}$ **and** $\boldsymbol{S} \succeq \boldsymbol{0}$
- $\boldsymbol{A^{old}} \leftarrow \boldsymbol{A}$, $\boldsymbol{S^{old}} \leftarrow \boldsymbol{S}$
- $\boldsymbol{J} \leftarrow \sum_{i=1}^n \|\boldsymbol{g_i} - \boldsymbol{\Phi_i A s_i}\|^2 + \lambda \sum_{i=1}^K \sum_{j=1}^n |\boldsymbol{s_{ij}}|$
- **while** $\boldsymbol{\Delta J} > \epsilon$ **do**
  - $\boldsymbol{A} \leftarrow \boldsymbol{A} - \mu(\sum_{i=1}^n \boldsymbol{\Phi_i^t}(\boldsymbol{\Phi_i A s_i} - \boldsymbol{g_i}) \boldsymbol{s_i^t}$
  - Set all negative entries of $\boldsymbol{A}$ to 0

Ensure: $\|\boldsymbol{A_j}\|_2^2 = 1 \, \forall j \in \{1, ..., K\}$ **and** $\boldsymbol{A} \succeq \boldsymbol{0}$
- $\boldsymbol{S} \leftarrow ((\boldsymbol{\Phi_i A})^{\boldsymbol{t}} \boldsymbol{g_i}) ./ ((\boldsymbol{\Phi_i A})^{\boldsymbol{t}} (\boldsymbol{\Phi_i A}) \boldsymbol{s_i} + \boldsymbol{\lambda})$
- $\boldsymbol{J'} \leftarrow \sum_{i=1}^n \|\boldsymbol{g_i} - \boldsymbol{\Phi_i A s_i}\|^2 + \lambda \sum_{i=1}^K \sum_{j=1}^n |\boldsymbol{s_{ij}}|$
- **if** $\boldsymbol{J'} > \boldsymbol{J}$ **then**
  - $\mu = \mu \times \alpha$ ($\alpha < 1$ is a reduction factor)
  - $\boldsymbol{A} \leftarrow \boldsymbol{A^{old}}$, $\boldsymbol{S} \leftarrow \boldsymbol{S^{old}}$
- **end if**
- **end while**

3. RELATED WORK

There exists relatively less work in the field of compressive Raman imaging. The work in [14] uses the Coded aperture snapshot spectral imager (CASSI) architecture [15] for Raman hyperspectral imaging. In this architecture, the light reflected by the scene of interest is modulated by a coded aperture and passed through a prism. The modulated and dispersed components of light then impinge upon a single 2D detector array which records a snapshot image of the hyperspectral datacube H underlying the scene. The snapshot can be viewed as a superposition of modulated and spatially shifted wavelength-dependent slices of H. Multiple snapshots need to be acquired by changing the aperture code, especially if N_λ is high. The compressive architecture in [16] computes a score function at every pixel, i.e. it computes (via hardware) a dot product between a random binary code and the spectrum at that pixel. Compared to both these architectures, the proposed architecture if successfully implemented by programming the spectrometer to skip pixels as per a pre-written pattern, has the advantage of simplicity in the hardware as well as the reconstruction algorithm. Our architecture can be incorporated in the point-wise serial as well as snapshot-based modes of acquisition. The idea of image reconstruction for grayscale images by skipping random subsets of pixels has recently been successfully implemented using matched wavelet transforms in [17], where the sensing matrix is termed as 'partial canonical identity' (PCI) matrix. However, this idea is new (and very relevant) in the context of Raman spectral imaging.

4. RESULTS

We now present our results on real samples for whom we have acquired Raman spectra using a Horiba Jobin Yvon HR800 Confocal Laser Raman Spectrometer. We performed three experiments with λ, the sparsity promoting parameter as 2×10^{-4} and number of dictionary columns $K = 12$ for each of these datasets where 20%, 50% and 80% of the total pixels in the image have been measured. The values of λ and K have been empirically chosen by cross validation on pure silicon datasets. We first show reconstruction results for acquired spectra of pure silicon (Si), silicon with traces of gallium nitride (Si+GaN) and pure paraffin (Par) with random sampling. We also present results for structured sampling for Si and Par. In each case, we present reconstructions for a particular slice in this paper. In addition, in the supplemental material, we show reconstructed spectral plots for missing pixels averaged across a 3×3 window, and videos to show reconstructions over the entire spectrum. We evaluate our results using the Root Mean Squared Error (RMSE) metric.

Random Sampling: For a pure of size 41×41 with 226 bands, we encounter a counting time of 2 seconds per pixel, i.e. an acquisition time of 1 hour. In Fig. 1, we show a comparison for the original image with the under-sampled and reconstructed image for different sampling percentages, viz. 80%, 50% and 20%, at the 128^{th} spectral band (where a spectral peak occurs). In Fig. 2, we present reconstruction results for the acquired Raman image of Par of size 51×51 with 208 spectral bands with a counting time of around 20 seconds per pixel, i.e. acquisition time of 7.5 hours. We show results for different sampling percentages at the 101^{st} spectral band, where one of the spectral peaks occur. Similarly, we provide results for Raman spectral acquisition of Si+GaN which again takes a per pixel counting time of around 2 seconds, i.e. acquisition time of 3.5 hours. Given the original image of size 41×61 with 179 spectral bands, we present the reconstructed spectral slice at the 82^{nd} band in Fig. 4. However, the reconstruction results are inferior compared to the others owing to the presence of complex textures arising from the distribution of GaN in the Si sample.

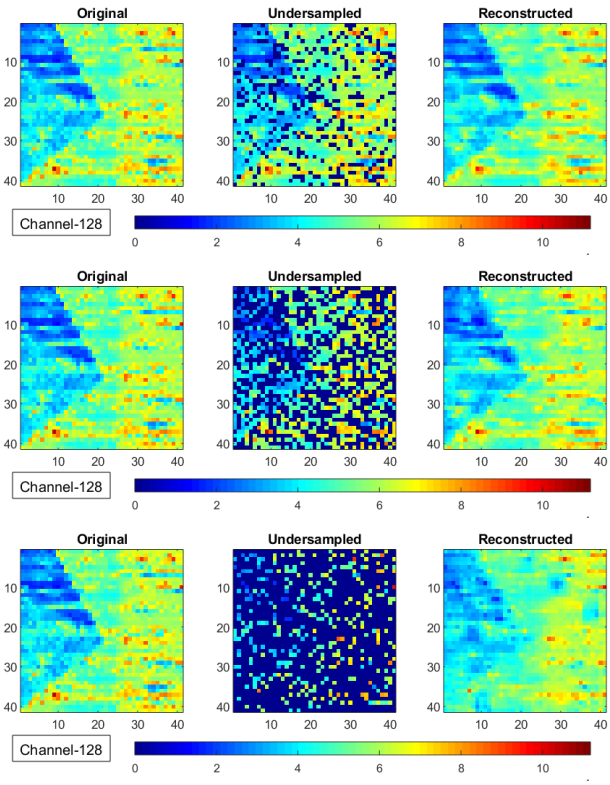

Fig. 1. Top to bottom : 80%, 50%, 20% sampling with **RMSE : 0.07405**, **0.08940** and **0.14299** respectively. Left to right per row: ground truth, **randomly** sampled and reconstructed Raman spectral image of pure silicon of size $41 \times 41 \times 226$ at 128^{th} spectral band. See supp. mat. for video results.

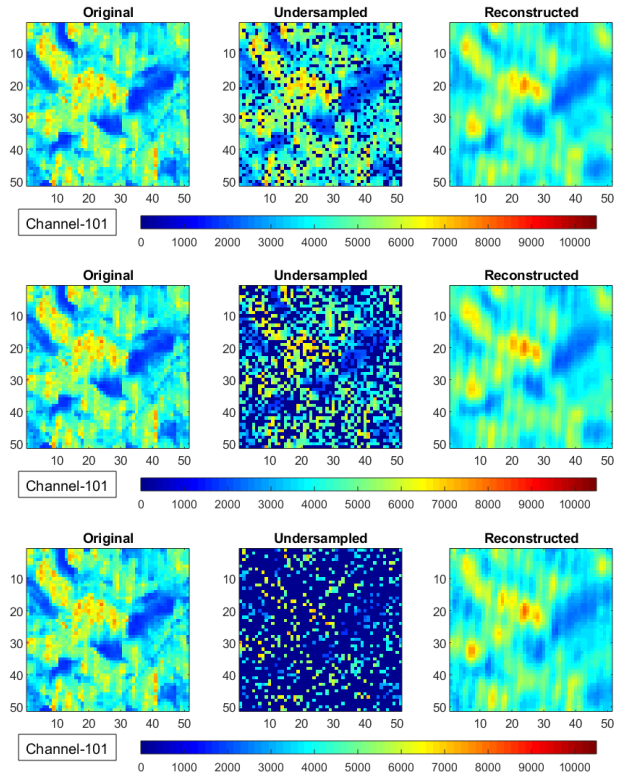

Fig. 2. Top to bottom: 80%, 50%, 20% sampling with **RMSE : 0.08269**, **0.08467** and **0.09388** respectively. Left to right per row: ground truth, **randomly** sampled and reconstructed image of pure paraffin of size $51 \times 51 \times 208$ at 101^{st} spectral band. See supp. mat. for video results.

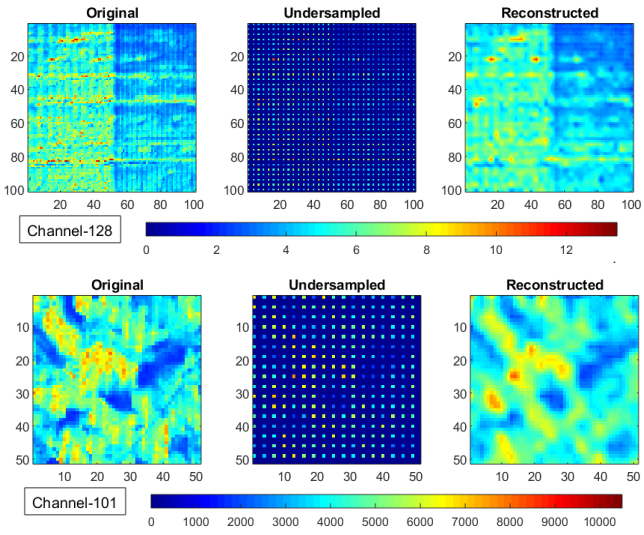

Fig. 3. Top to bottom : (a) pure silicon of size $101 \times 101 \times 226$ at 128^{th} spectral band with **RMSE : 0.152** (b) pure paraffin of size $51 \times 51 \times 208$ at 101^{st} spectral band with **RMSE : 0.103**. Left to right per row: ground truth, **structurally** sampled with 4^{th} pixel in both directions and reconstructed image. See supp. mat. for video results.

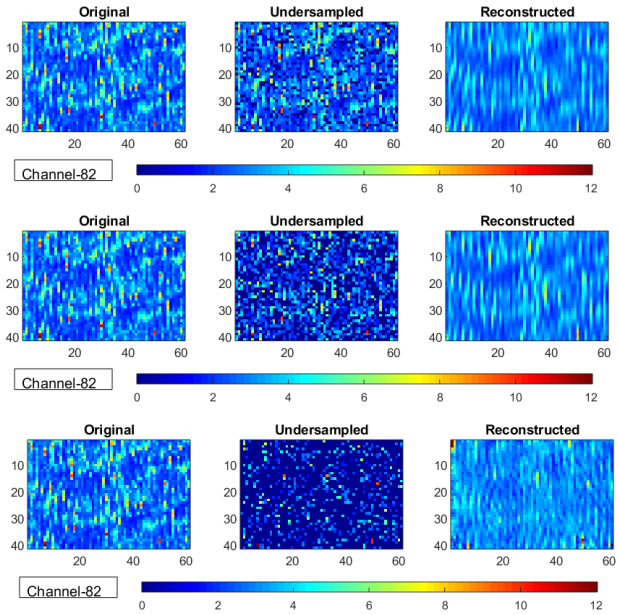

Fig. 4. Top to bottom: 80%, 50%, 20% sampling with **RMSE : 0.2547**, **0.2686** and **0.3359** respectively. Left to right per row: ground truth, **randomly** sampled and reconstructed image of Si + GaN of size $41 \times 61 \times 179$ at 82^{th} spectral band. See supp. mat. for video results.

Structured sampling: We now show reconstruction results for Si and Par acquired as before, but for structural undersampling with every 4^{th} pixel measured in both directions. The results for the pure ($101 \times 101 \times 226$) are as follows. Compared to the image obtained by bicubic interpolation $\tilde{\boldsymbol{G}}$ (RMSE: **0.165**) and the image reconstructed from the dictionary $\boldsymbol{A}_{\tilde{G}}$ inferred from $\tilde{\boldsymbol{G}}$ (RMSE: **0.17**), we observe a superior performance for the reconstructed image for which $\boldsymbol{A}$ is further evolved using $\boldsymbol{G}$ with the initial dictionary as $\boldsymbol{A}_{\tilde{G}}$ (RMSE: **0.152**). See Fig. 3. Using the same paraffin sample of size $51 \times 51 \times 208$, our results demonstrated in Fig. 3 (RMSE: **0.103**) outperform bicubic interpolated image $\tilde{\boldsymbol{G}}$ and that reconstructed using $\boldsymbol{A}_{\tilde{G}}$ inferred from $\tilde{\boldsymbol{G}}$ with RMSE values of **0.1141** and **0.1145** respectively.

5. CONCLUSIONS

We have presented a method of compressive acquisition and reconstruction for Raman hyperspectral images, which is very simple and effective, and can lead to a hardware prototype for efficient image acquisition. The results are validated on actual Raman spectral images with a large number of wavelengths at very low sampling ratios. One important line of future work will involve development of such a hardware prototype. The reconstruction algorithm can also be further improved using deep learning techniques [18] or tensor factorizations [19].

6. REFERENCES

[1] M. Schaeberle et al, "Raman chemical imaging: histopathology of inclusions in human breast tissue," *Anal. Chem*, 1996.

[2] D. Wetzel and S. LeVine, "Imaging molecular chemistry with infrared microscopy," *Science*, 1999.

[3] Y. Zhang Y, H. Hong, and W. Cai, "Imaging with raman spectroscopy," *Current Pharmaceutical Biotechnology*, 2010.

[4] S. Kawata et al, "Near-field optics and spectroscopy for molecular nano-imaging," *Sci. Prog*, 2004.

[5] J. Ling, S. Weitman, V. Miller, R. Moore, and A. Bovik, "Direct raman imaging techniques for study of the subcellular distribution of a drug," *Applied Optics*, 2002.

[6] D. Lauwers, Ph. Brondeel, L. Moens, and P. Vandenabeele, "In situ raman mapping of art objects," *Philosophical Transactions of the Royal Society A*, 2016.

[7] L. Haskin et al, "Raman spectroscopy for mineral identification and quantificafionor in sire planetary surface analysis: A point count method," *Journal of Geophysical Research*, 1997.

[8] A. Tiberj and J. Camassel, *Raman Imaging in Semiconductor Physics: Applications to Microelectronic Materials and Devices*, 2012.

[9] Chandrasekhara Venkata Raman and Kariamanikkam Srinivasa Krishnan, "A new type of secondary radiation," *Nature*, vol. 121, no. 3048, pp. 501, 1928.

[10] E. Candes and M. Wakin, "An introduction to compressive sampling," *IEEE Sig. Proc. Mag.*, 2008.

[11] J. Romberg, "Imaging via compressive sampling," *IEEE Sig. Proc. Mag.*, 2008.

[12] A. Rajwade, D. Kittle, T.-H. Tsai, D. Brady, and L. Carin, "Coded hyperspectral imaging and blind compressive sensing," *SIAM J. Imaging Sciences*, 2013.

[13] P. O. Hoyer, "Non-negative sparse coding," in *IEEE Workshop on Neural Networks for Signal Processing*, 2002.

[14] D. G. Carreno and H. Arguello Fuentes, "Transmittance analysis in coded aperture compressive raman spectroscopy imaging," in *Symposium of Signals, Images and Artificial Vision*, 2013.

[15] D. Kittle, K. Choi, A. Wagadarikar, and D. J. Brady, "Multiframe image estimation for coded aperture snapshot spectral imagers," *Appl. Opt.*, 2010.

[16] B. Davis et al, "Multivariate hyperspectral raman imaging using compressive detection," *Analytical Chemistry*, 2011.

[17] N. Ansari and A. Gupta, "Image reconstruction using matched wavelet estimated from data sensed compressively using partial canonical identity matrix," *IEEE Trans. Image Processing*, 2017.

[18] A. Lucas, M. Iliadis, R. Molina, and A. Katsaggelos, "Using deep neural networks for inverse problems in imaging," *IEEE Signal Processing Magazine*, 2018.

[19] P. Zhou, C. Lu, Z. Lin, and C. Zhang, "Tensor factorization for low-rank tensor completion," *IEEE Transactions on Image Processing*, 2018.