Abstract

Raman spectroscopy is a non-invasive technique that characterizes the biochemical components of a sample using the Raman scattering effect. A Raman data cube is a three-dimensional set comprising two spatial and one spectral dimensions. Typically, thousand bands per pixel are captured which considerably increases acquisition times. Furthermore, many pixels capture information from a mixture of several components, thus decreasing the accuracy of substances classification. For this reason, unmixing methods emerge as a solution to this problem. Traditionally, the SUnSAL algorithm is used to solve the spectral mixing problem on linear mixing assumption [1]. Additionally, the recent theory of compressive sensing (CS) provides a solution to the excessive size of the data cube and large acquisition times. An efficient CS optical implementation is the coded aperture snapshot spectral imaging (CASSI) system which capture spatial and spectral information of scene. This paper presents the spectral unmixing process on CASSI measurements of Raman spectral images. The SUnSAL algorithm is modified to be used with CASSI images and a signature re-scaling step is added to take in account the variability of signatures by different capture set-up. Extensive simulations with real and synthetic Raman images show the accuracy of this algorithm.

1. Introduction

Raman spectroscopy is a powerful technique to analyse the chemical composition of different compounds. This technique uses the Raman scattering effect, which occurs due to the interaction between the sample and the photons generated by a laser, such that, when photons strike the molecules of the sample, a few of them (about 1 in $10^{11}$) are scattered with changes in their energy, these changes provide composition information of the scene. Traditional point spectroscopy approaches provide limited information of the sample, since just one pixel of the scene is measured. However, in Raman spectroscopy, an image can be created by measuring several adjacent sample points (pixels), which are then used to build a data cube that provides much more information. A Raman image consists of a large amount of spatial information across a multitude of wavenumbers, in which the wavenumber represents the frequency changes of scattered photons [2] [3] [4].

There are two general approaches to obtain Raman images, serial and direct imaging. The serial imaging approach measures a spatial (one-dimensional) region across all the spectrum, this region can be a point (whiskbroom spectrometer) or a line (pushbroom spectrometer). On the other hand in the direct imaging approach all points in the Raman image in a fixed wavenumber are captured simultaneously [5] [6] [7]. Despite the advantages of Raman images, their great acquisition time (about 1 second per spectrum) and, their huge sizes (about 1000 bands per pixel), make this technique impractical when large images or many images need to be captured.

The recent theory of compressive sensing (CS) [8] [9] arises as a solution to excessive acquisition times and huge sizes of the image, by exploiting the assumption that natural scenes can be represented in some basis $\Psi$ with just a few non-zero coefficients, known as sparsity. In addition, it is possible to design sensing protocols to capture the essential information with far fewer measurements that those dictated by Nyquist-Shannon theorem. In CS, sensing matrices $\Phi$, must be highly incoherent with representation basis $\Psi$ [8] [9] [10] [11]. Finally the signal of interest is recovered by computationally solving an optimization problem.

The coded-aperture snapshot spectral imager (CASSI) is an efficient implementation of compressive spectral imaging (CSI) sensor. Figure 1 shows the elements in CASSI system architecture. CASSI measurements are $y = Hf$, where $H \in \mathbb{R}^{N(N+L-1) \times N \cdot N \cdot L}$ is the sensing matrix that accounts for the effects of the coded aperture and dispersive
element; and \( f \in \mathbb{R}^{N \times L} \) is the spectral scene. The number of samples required to recover the data cube is proportional to non-zero coefficients of the underlying signal in the basis \( \Psi \). Formally, a signal \( F \in \mathbb{R}^{N \times N \times L} \), or its vector representation \( f \) has an \( S \)-sparse representation in some basis \( \Psi \). Usually a kronecker basis given by \( \Psi_{2D} = \Psi_2 \otimes \Psi_1 \) is used to represent the spatio.spectral signal, where \( \Psi_2 \) is often the 2D-Wavelet Symmlet 8 basis and \( \Psi_1 \) is the discrete cosine basis. Then, given the set of measurements \( y \), the data cube is reconstructed by solving an undetermined linear system given by \( \arg \min_\theta ||y - A\theta||_2^2 + \tau||\theta||_1 \), where \( \theta \) is an \( S \)-sparse representation of \( f \) in the basis \( \Psi \) and, \( \tau \) is a regularization parameter \( [10] \).

Additionally to the size problem of Raman images, many pixels capture information from a mixture of several components, which lead to difficulties in the process of classification or identification of compounds, since these measured pixel can be described as a linear combination of endmembers, which refers to the spectra of pure materials. Then, unmixing is a technique used to decompose mixed pixel in its constituent materials or endmembers, by finding the abundances or percentages of pure materials present in the pixel \([11][12]\). Frequently, the signatures of pure materials (library) are known, and can be found in databases such as RRUFF Database (Raman spectroscopy, X-ray) \([13]\), USGS Digital Spectral Library \([14]\). The linear model (LMM) is often used due its simplicity and accurate results, this model assumes that a pixel is a linear combination of a few endmembers of the entire spectral library. SUnSAL is an algorithm based on the alternating direction method of multipliers (ADMM) that solves the mixing problem \([1]\). This algorithm uses the sparsity assumption to find the abundances.

### 2. Spectral unmixing

Unmixing refers to any process that separates a pixel in its constituent spectra and/or a set of fractional abundances. Unmixing algorithms rely on a mixing model, that can be linear or non-linear; most common unmixing approaches rely on the linear model (LMM), which assumes that the scattered photons interact with just one pixel in the scene before reflecting, then the captured spectra can be expressed as a linear combination of materials present in the underlying pixel as follows \([12][15][16][19]\):

\[
f_j = Ma_j = M[a_{j1}, a_{j2}, ..., a_{jN_e}]^T + n
\]

where \( f_j \in \mathbb{R}^L \) is the observed spectra in the \( j \)-th pixel, \( L \) is the number of spectral bands, \( M \in \mathbb{R}^{L \times N_e} \) is the spectral library, \( N_e \) corresponds to the number of endmembers, \( a \in \mathbb{R}^{N_e} \) is the abundances vector and \( n \) is the noise. Finally the abundances vector entries are assumed non-negative, \( a_{ji} \geq 0 \). Further, the sparse linear (SLM) model has been recently explored \([17][18]\). In this approach each pixel is seen as a sparse linear mixture from the spectral dictionary. This concept is shown in Fig 2.

Where the data cube \( F \in \mathbb{R}^{N \times N \times L} \), with \( N \times N \) spatial dimensions and \( L \) spectral bands, or its vector representation \( f = [f_1^T, f_2^T, ..., f_L^T]^T \in \mathbb{R}^{N^2 \times L} \) with \( f_i^T \) that corresponds to a column vector containing the values of all pixels in the \( i \)-th band. Then data cube can be written as

\[
f = Ma
\]

where \( M = (M \otimes I) \in \mathbb{R}^{N^2 L \times N^2 N_e} \) is an identity matrix, \( \otimes \) is the Kronecker product operation of two matrices, \( a = [a_1^T, a_2^T, ..., a_N^T]^T \) is the collection of abundances for all pixels in the scene. Thus the unmixing optimization problem can be defined as \([1]\).

\[
\min ||a||_1 \quad \text{subject to} \quad ||f - Ma||_2^2 < \delta
\]

Exploiting CS theory, the abundance collection \( a \) described in equation \([3]\) can be obtained directly from compressed...
measurements, in CASSI these measurements can be written as \( y = Hf + n = H \mathbf{M} a + n \). Then the goal is to find the abundances vector from CASSI measurements by solving the optimization problem given by:

\[
\hat{a} = \min \frac{1}{2} \| y - H \mathbf{M} a \|_2^2 + \tau \| a \|_1
\]  

(4)

\[
(\mathbf{I} \otimes \Omega_{2D}^T) \beta = \arg \min \frac{1}{2} \| y - H \mathbf{M} \beta \|_2^2 + \tau \| \beta \|_1
\]

(7)

subject to \( \{(\mathbf{I} \otimes \Omega_{2D}^T) \beta \}_i \geq 0 \) for \( i = 1, \ldots, N^2 N_e \) where \( \mathbf{M}' = (\mathbf{M} \otimes \Omega_{2D}^T) \). When pure pixel assumption holds, the \( v_1, v_2, \ldots, v_{N_e} \) can be obtained from \( \hat{\beta}' \) (4) rewrite as:

\[
\hat{\beta}' = \arg \min \frac{1}{2} \| y - H \mathbf{M}' \beta' \|_2^2 + \tau \| \beta' \|_1
\]

(5)

subject to \( \{(\mathbf{I} \otimes \Omega_{2D}^T) \beta' \}_i \geq 0 \) for \( i = 1, \ldots, N^2 N_e \), where \( \mathbf{M}' = (\mathbf{M} \otimes \Omega_{2D}^T) \). When pure pixel assumption holds, the \( v_1, v_2, \ldots, v_{N_e} \) can be obtained from \( \hat{\beta}' \).

\[
\hat{\alpha}' = [a_{11}', a_{12}', \ldots, a_{N_e1}', a_{N_e2}']^T
\]

(8)

where \( e_{jk} \) is an unitary vector in the \( k-th \) position of \( \hat{\alpha}' \), and non-zero coefficient in the \( j-th \) position of \( e_{jk} \), \( \forall j \in (1, 2, \ldots, N_e) \), which means that in \( k-th \) pixel there is a pure pixel of \( j-th \) endmember.

**3. Inverse Problem**

SUUnSAL is an algorithm based on the alternating direction method of multipliers (ADMM), this algorithm solves the problem given in Eq. (7) as shown in algorithm [1]. where \( G = (\mathbf{I} \otimes \Omega_{2D}^T) \), \( \mathbf{A} \) is the same \( \mathbf{M}' \) defined in Eq. (7) the estimated sparse abundances vector can be recovered.
Algorithm 1 SUnSAL

1: \( k = 0, \mu > 0, u_0, \) and \( d_0, B = A^T A + \mu I \)
2: do
3: \( w \leftarrow A^T y + \mu G^T (u_k + d_k) \)
4: \( \beta_{k+1} \leftarrow B^{-1} w \)
5: \( u_{k+1} \leftarrow \max\{0, \text{soft}((G\beta_{k+1} - d_k), \lambda/\mu)\} \)
6: \( d_{k+1} \leftarrow d_k - (G\beta_{k+1} - u_{k+1}) \)
7: \( k \leftarrow k + 1 \)
8: while stopping criterion is not satisfied

by \( \hat{a}' = ((I \otimes \Omega_2^T)\hat{a}') \), finally the values \( v_1, v_2, \ldots, v_{N_e} \)
are extracted from \( \hat{a}' \), in order to do that let define \( b \) as:

\[
B = \{\hat{a}'_1, \hat{a}'_2, \ldots, \hat{a}'_{N_e}\}^T = \{b_1, b_2, \ldots, b_{N_e}\} \quad (9)
\]

where \( \hat{a}'_i \) has been defined in [8] then \( v_i \) is extracted by \( v_i = \max(b_i)\mu \), with \( \mu \) a parameter to take in account noise.

4. Computer Simulations

The proposed unmixing algorithm is evaluated using synthetic and real images using a fixed number of iterations as stop criterion. Synthetic data cubes are created with endmembers randomly selected from the RRUFF database [13] containing information in the range 326-1296 cm\(^{-1}\), all signatures in synthetic images are re-sampled to 256 bands. Thus, images of 16 \times 16 and 32 \times 32 pixels of spatial resolution and 256 spectral bands are created. The abundance map are shown in Fig. 4. Multishot CASSI model, described in [10], is used to compress the images and then the unmixing technique is used on the compressed signal. CASSI measurements were obtained by varying number of shots from 4 to 14, and the results are compared with those obtained when the image is reconstructed before using unmixing algorithm, the performance is evaluated using signal to reconstruction error (SRE), defined as

\[
\text{SRE} = 10 \log_{10} \frac{||a||_2^2}{||a - \hat{a}||_2^2}.
\]

An coded aperture analysis was performed to find the optimal value using Bernoulli random distribution to create this apertures. Figure 5 shows the behaviour of SRE with different transmittances, it can be noticed that 60% is the optimal transmittance.

Results of proposed unmixing process and traditional unmixing on reconstructed data cube are shown in Tables 1 and 2 note that with just 17% of information the unmixing algorithm on compressed image exhibit mean SRE of 5 db or more, which is defined as minimum requiring accuracy by Ramirez et al. [11]. Further, apply the unmixing process directly on compressed images exhibit better results than those achieved by using this technique on reconstructed data cube.

Real data cubes of a pill, containing caffeine, aspirin and paracetamol, was captured on laboratory using Horiba.
Table 2: SRE (in decibels) for the $16 \times 16$ spatial resolution synthetic data cube.

<table>
<thead>
<tr>
<th>shots</th>
<th>SRE (db)</th>
<th>proposed</th>
<th>SUnSAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 (13%)</td>
<td>2.3918</td>
<td>0.5774</td>
<td></td>
</tr>
<tr>
<td>3 (20%)</td>
<td>5.5039</td>
<td>0.6768</td>
<td></td>
</tr>
<tr>
<td>4 (26%)</td>
<td>7.2878</td>
<td>1.3954</td>
<td></td>
</tr>
<tr>
<td>5 (33%)</td>
<td>7.5347</td>
<td>2.0492</td>
<td></td>
</tr>
<tr>
<td>6 (39%)</td>
<td>9.2887</td>
<td>2.1867</td>
<td></td>
</tr>
<tr>
<td>7 (46%)</td>
<td>9.9718</td>
<td>2.7555</td>
<td></td>
</tr>
<tr>
<td>8 (53%)</td>
<td>10.1815</td>
<td>3.4501</td>
<td></td>
</tr>
</tbody>
</table>

Table 3: SRE (in decibels) for the $32 \times 32$ spatial resolution real data cube.

<table>
<thead>
<tr>
<th>shots</th>
<th>SRE (db)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (6%)</td>
<td>2.447</td>
</tr>
<tr>
<td>2 (13%)</td>
<td>2.7402</td>
</tr>
<tr>
<td>2 (19%)</td>
<td>2.9005</td>
</tr>
<tr>
<td>4 (25%)</td>
<td>3.4821</td>
</tr>
<tr>
<td>5 (31%)</td>
<td>3.6517</td>
</tr>
<tr>
<td>6 (38%)</td>
<td>4.4685</td>
</tr>
<tr>
<td>7 (44%)</td>
<td>5.091</td>
</tr>
<tr>
<td>8 (50%)</td>
<td>5.3762</td>
</tr>
</tbody>
</table>

Table 4: SRE (in decibels) for the $16 \times 16$ spatial resolution real data cube.

<table>
<thead>
<tr>
<th>shots</th>
<th>SRE (db)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 (9%)</td>
<td>4.6120</td>
</tr>
<tr>
<td>4 (12%)</td>
<td>5.6688</td>
</tr>
<tr>
<td>5 (16%)</td>
<td>6.1781</td>
</tr>
<tr>
<td>6 (19%)</td>
<td>7.0650</td>
</tr>
<tr>
<td>7 (22%)</td>
<td>7.1501</td>
</tr>
<tr>
<td>8 (25%)</td>
<td>8.1799</td>
</tr>
<tr>
<td>9 (28%)</td>
<td>9.9379</td>
</tr>
<tr>
<td>10 (32%)</td>
<td>10.1130</td>
</tr>
<tr>
<td>11 (35%)</td>
<td>10.230</td>
</tr>
<tr>
<td>12 (38%)</td>
<td>10.3066</td>
</tr>
<tr>
<td>13 (41%)</td>
<td>10.5694</td>
</tr>
<tr>
<td>14 (45%)</td>
<td>10.6435</td>
</tr>
</tbody>
</table>

Figure 6: estimated pure pixels.

LabRAM spectrometer, this data cubes exhibit either $16 \times 16$ or $32 \times 32$ pixels of spatial resolution, and 1024 spectral bands. The dictionary was obtained form RRUFF database [13]. figure 6 shows a pixel of underlying data cube which was identified as a pure pixel by the proposed algorithm, and the respective sign of this component from the spectral library. Additionally unmixing process was performed on full data cube and the obtained abundance matrix was used as reference data to prove the accuracy of proposed algorithm. the results are shown in table 3.

5. Future work and conclusions

A compressive spectral unmixing method has been evaluated in Raman images, avoiding the expensive task of reconstruct the Raman data cube. Further a re-scaling step has been added to take in account the variability of Raman signatures. the abundances array is recovered by solving a optimization problem using a ADMM based algorithm. simulations shows that in synthetic images the set of abundances can be estimated correctly using only 17% of the data cube information. For real data cubes around 12%, in $32 \times 32$ data cube, and 44% in $16 \times 16$ data cube, is needed in order to estimate abundances vector accurately; however this results are based in a estimation from full data cube. Future work points to the design of coded apertures which is extremely important in CASSI, since it can lead to a significant improvement of results obtained.
6. Acknowledges

The authors gratefully acknowledge the Vicerrectoría de Investigación y Extensión of Universidad Industrial de Santander for supporting this work registered under the project title “EXTRACCIÓN Y SEPARACIÓN DE INFORMACIÓN ESPECTRAL EN IMÁGENES OBTENIDAS DE FORMA REMOTA USANDO MUESTREO COMPRESIVO Y APERTURAS CODIFICADAS DE COLOR”, with VIE code 1804.

References


