



HAL
open science

Fast reconstruction of hyperspectral images from coded acquisitions using a separability assumption

Elizabeth Hemsley, Ibrahim Ardi, Tony Rouvier, Simon Lacroix, Hervé Carfantan, Antoine Monmayrant

► **To cite this version:**

Elizabeth Hemsley, Ibrahim Ardi, Tony Rouvier, Simon Lacroix, Hervé Carfantan, et al.. Fast reconstruction of hyperspectral images from coded acquisitions using a separability assumption. *Optics Express*, 2022, 30 (5), pp.8174-8185. 10.1364/OE.448893 . hal-03610209v2

HAL Id: hal-03610209

<https://hal.science/hal-03610209v2>

Submitted on 28 Jul 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Fast Reconstruction of Hyperspectral Images from Coded Acquisitions using a Separability Assumption

July 28, 2022

Elizabeth Hemsley^{1,*}, Ibrahim Ardi^{1,2}, Tony Rouvier², Simon Lacroix¹, Hervé Carfantan², Antoine Monmayrant¹

¹LAAS-CNRS, Université de Toulouse, CNRS, 7 Avenue du Colonel Roche, 31400 Toulouse, France

²IRAP, Université de Toulouse, CNRS, CNES, 14 Avenue Édouard Belin, 31400 Toulouse, France

*ehemsley@laas.fr

Abstract

We present a fast reconstruction algorithm for hyperspectral images, utilizing a small amount of data without the need for any training. The method is implemented with a dual disperser hyperspectral imager, and makes use of spatial-spectral correlations by a so-called *separability assumption* which assumes that the image is made of regions of homogenous spectra. The reconstruction algorithm is simple and ready-to-use, and does not require any prior knowledge of the scene. A simple proof-of-principle experiment is performed, demonstrating that only a small number of acquisitions are required, and the resulting compressed data-cube is reconstructed near instantaneously.

1 Introduction

Hyperspectral (HS) imagers provide precise spectral information for every pixel in a scene, and are useful for a range of applications - *e.g.* deep space or earth observation, material analysis, gas detection, etc. The time taken to obtain the 3D HS data-cube can be long, particularly for scanning type imagers, necessitating several acquisitions with a 2D sensor, and the volume of data produced is large. Often much of the information in the HS datacube is redundant, given that a real world scene contains a considerable amount of spectral-spatial correlations [1, 2]. These correlations lead to the field of compressed sensing, which reduces the amount of data needed to reconstruct the HS datacube, based on the assumption of an underlying sparse dictionary [3, 4, 5, 6, 7]. More recently, neural networks have been trained to reconstruct a HS cube using a similarly small amount of data [8, 9, 10]. But these methods either require significant time during the reconstruction, or a learning phase which make them not straightforward and ready-to-use. In this paper we present an alternative method requiring no training or prerequisite knowledge of the scene. The amount of data required for this algorithm is similarly small, and the reconstruction time is very short. A fundamental point is that we do not attempt to reconstruct the entire hyperspectral datacube, that is to reconstruct the spectrum for each pixel of the scene. The key idea of our approach is to observe that nearby pixels often share the same spectrum with only a small panchromatic intensity

36 difference, for example due to slightly different lighting conditions. If two pixels have differing
 37 spectra, the difference in the panchromatic intensity will likely be more significant. Therefore the
 38 scene can be divided into spectrally homogenous regions using information from the panchromatic
 39 image, and we need only assign a spectrum to each region.

40 **Paper outline:** the next section depicts how the reconstruction problem for a dual disperser HS
 41 imager is given by a least square minimization, augmented with a spectral regularization. Section 3
 42 presents the experimental setup, how the scene ground truth is recovered, and the metric used to
 43 asses the performance of our approach. Section 4 analyzes experimental results for one simple and
 44 one more complex scene, and in Section 5 we discuss how to remove reliance on segmentation of the
 45 panchromatic image.

46 2 Methodology

Due to spectral-spatial correlations within a HS scene, adjacent pixels are likely to share the same spectra. Following this logic, we can separate any hyperspectral scene into spectrally homogenous regions, defined as spatially connected areas where the spectra for every pixel is the same, apart from a multiplication factor which depends on the intensity. The whole homogenous region can thus be described by one single spectrum and an intensity map given by the panchromatic intensity. We denote this assumption as the *separability assumption* (SA). If a homogenous region in the scene consists of K pixels indexed by $k \in \{1 \dots K\}$, the HS information \mathbf{o}_k of pixel k in the region is given by:

$$\mathbf{o}_k = P_k \mathbf{s}, \tag{1}$$

47 where \mathbf{s} is the normalized spectrum of the region made of W spectral bands, P_k the panchromatic
 48 intensity of that pixel, and the index k corresponds to all the pixels of the region. Therefore, the
 49 hyperspectral information \mathbf{o} of the whole region can be written:

$$\mathbf{o} = \begin{bmatrix} \mathbf{o}_1 \\ \mathbf{o}_2 \\ \vdots \\ \mathbf{o}_K \end{bmatrix} = \begin{bmatrix} P_1 \mathbf{s} \\ P_2 \mathbf{s} \\ \vdots \\ P_K \mathbf{s} \end{bmatrix} = \begin{bmatrix} P_1 \mathbf{I} \\ P_2 \mathbf{I} \\ \vdots \\ P_K \mathbf{I} \end{bmatrix} \mathbf{s} \tag{2}$$

50 where \mathbf{I} is the identity matrix of size W .

51 The measurement is implemented using a dual disperser HS imager [11], illustrated in Figure 1.
 52 For this architecture each camera pixel is co-registered with a scene pixel, so the resulting measure-
 53 ment value on each camera pixel is a linear combination of the spectral bands present at the scene
 54 pixel. The weights in this linear combination vary across the scene and are determined by a spatial
 55 filter placed in the imaging plane between the two dispersive 4-f lines, as explained in [12]. Dual
 56 disperser imagers are often used with reconfigurable spatial masks, for example digital micromirror
 57 devices (DMDs), so with different mask configurations, we make N coded measurements of the
 58 scene.

59 For each region in the scene, a measurement \mathbf{m}_n corresponding to the values of the camera

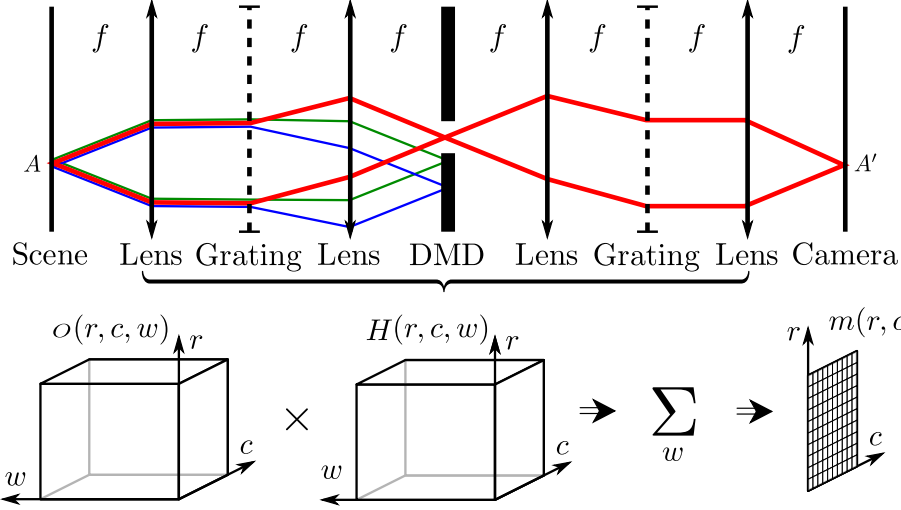


Figure 1: Simplified schematic of the dual disperser architecture and the programmable filtering \mathbf{H} implemented by the DMD mask.

60 pixels in the region can be written as

$$\mathbf{m}_n = \mathbf{H}_n \mathbf{o} = \mathbf{H}_n \begin{bmatrix} P_1 \mathbf{I} \\ P_2 \mathbf{I} \\ \vdots \\ P_K \mathbf{I} \end{bmatrix} \mathbf{s} = \mathbf{G}_n \mathbf{s} \quad (3)$$

61 where the measurement matrix \mathbf{H}_n of dimension $K \times W$ is given by the spatial filter for acquisition
 62 $n \in \{1 \dots N\}$, and the measurement \mathbf{m}_n is a vector of length K . In other words the measurement
 63 value corresponding to pixel k is a linear combination of the wavelength bands w for that pixel,
 64 with weights defined by the k^{th} row of \mathbf{H}_n and multiplied the panchromatic intensity P_k .

65 For N such acquisitions we can concatenate the results, giving

$$\mathbf{m} = \begin{bmatrix} \mathbf{m}_1 \\ \mathbf{m}_2 \\ \vdots \\ \mathbf{m}_n \end{bmatrix} = \begin{bmatrix} \mathbf{H}_1 \\ \mathbf{H}_2 \\ \vdots \\ \mathbf{H}_n \end{bmatrix} \mathbf{o} = \begin{bmatrix} \mathbf{H}_1 \\ \mathbf{H}_2 \\ \vdots \\ \mathbf{H}_n \end{bmatrix} \begin{bmatrix} P_1 \mathbf{I} \\ P_2 \mathbf{I} \\ \vdots \\ P_K \mathbf{I} \end{bmatrix} \mathbf{s} = \begin{bmatrix} \mathbf{G}_1 \\ \mathbf{G}_2 \\ \vdots \\ \mathbf{G}_n \end{bmatrix} \mathbf{s} = \mathbf{G} \mathbf{s} \quad (4)$$

66 We obtain the panchromatic information P_k via one acquisition with the spatial filter completely
 67 open. The spectrum \mathbf{s} for the homogeneous region can then be reconstructed using the measure-
 68 ments \mathbf{m} and this panchromatic information P_k with a simple least squares procedure. However,
 69 as matrix \mathbf{G} can be ill-conditioned, we propose to use a Tikhonov regularization along the spectral
 70 dimension [13, 14]. The estimated spectrum for the region is thus calculated by

$$\hat{\mathbf{s}} = \arg \min_{\mathbf{s}} \|\mathbf{m} - \mathbf{G} \mathbf{s}\|_{\Gamma^{-1}}^2 + \mu_{\lambda} \|\mathbf{D}_{\lambda} \mathbf{s}\|^2 \quad (5)$$

$$= (\mathbf{G}^t \Gamma^{-1} \mathbf{G} + \mu_{\lambda} \mathbf{D}_{\lambda}^t \mathbf{D}_{\lambda})^{-1} \mathbf{G}^t \Gamma^{-1} \mathbf{m} \quad (6)$$

71 where \mathbf{D}_λ is a first order difference operator, μ_λ is a regularization parameter and the weighting
72 matrix $\mathbf{\Gamma} = \text{diag}\{\mathbf{m}\}$ accounts for the noise variance which is proportional to the intensity for
73 a Poisson counting noise. As the system of Eq. (6) is overdetermined when $K \times N > W$, the
74 regularization parameter μ_λ can be set to zero if matrix $\mathbf{G}^t \mathbf{\Gamma}^{-1} \mathbf{G}$ is well-conditioned. Note that for
75 a constant regularization parameter μ_λ , the regularization is naturally weaker for large regions than
76 for small regions as the number of elements in the first term of Eq. (5) is larger, while the number
77 of elements in the second term (ensuring the regularization) stays constant. Finally, the value of
78 the estimated spectrum can take on negative values, which if necessary can be avoided by imposing
79 positive values using a non-negative least squares procedure, at an additional computational cost.

80 From a computational point of view, the solution can be directly computed using Eq. (6), as
81 the linear system to solve only has W unknowns. Note that solution of the linear system has a
82 constant computation cost regardless of the number of pixels K in the region and the number of
83 acquisitions N . Moreover, the computation costs of the matrix $\mathbf{G}^t \mathbf{\Gamma}^{-1} \mathbf{G}$ and the vector $\mathbf{G}^t \mathbf{\Gamma}^{-1} \mathbf{m}$
84 are linear with respect to K and N .

85 Additionally, as each region can be computed independently, the algorithm is paralellizable,
86 giving the potential for extremely fast reconstruction overall.

87 One important consideration for this algorithm is how the image can be divided into homo-
88 geneous regions, given no prior knowledge of the scene. We can consider that if two pixels have
89 different spectra, there will likely be a difference in the panchromatic intensity, if we assume the
90 presence of adjacent metamers is unlikely. For two adjacent but spectrally distinct regions, the dif-
91 ference in intensity is visible on the panchromatic as an edge, and thus some form of segmentation
92 algorithm applied on the panchromatic image can be used to divide the image into regions. As
93 segmentation is an ill-defined and non-trivial problem, we expect that this method will be most
94 reliable for simple scenes that are easily segmented and obey the separability assumption *e.g.* a
95 scene that is not highly textured, blurred and does not contain adjacent metamers or smooth spatial
96 changes in spectra.

97 3 Experiment

98 The algorithm was tested experimentally using the dual disperser system described in [12], with an
99 additional lens before the input. The scene has $W = 110$ spectral bands in the range 425-650 nm,
100 with 400 by 400 spatial pixels, and the spatial filtering is implemented using a DMD. Due to the
101 orientation of the DMD, as well as size mismatch between the DMD mirror and pixel, the matrix
102 \mathbf{H} contains non-binary values. Whilst one would imagine that a fully 2D randomized DMD mask
103 gives the best conditioning of equation (6), in practice the geometry mismatches lead to significant
104 mixing of the data resulting in a less accurate reconstruction. To compensate for this effect, we
105 are obliged to use a simpler mask design with only a 1D variation in the direction of dispersion.
106 Consequently more acquisitions will be required than for an ideal 2D randomized mask, as there
107 is less variation in the measurement data. The mask is designed according to [15], where long
108 sequences of on or off mirrors are avoided to maximize variation in the information and improve
109 conditioning of equation (6). This paper does not perform an extensive study into the ideal ratio
110 of open mirrors (ROM), but we obtained good results with using $ROM = 0.1$, with some scope for
111 further optimization. The regularization parameter was chosen to be $\mu_\lambda = 10^3$. A typical procedure
112 to implement this algorithm consists of an acquisition of the panchromatic image, followed by one
113 or more coded acquisitions with a 1D random mask.

114 3.1 Slit Scanning

115 To evaluate the reconstruction method, ideally one would compare to the underlying ‘Ground Truth’
116 of the scene. However, for an experimental system this is not possible, so instead we compare to
117 the reconstruction obtained via a straightforward slit scanning approach.

118 The data for both Slit Scanning (SS) and the Separability Assumption (SA) approaches are
119 obtained with the same exposure time on the camera. The gray level on the DMD for the SS data
120 was at its maximum value of 255. As SA combines several bands, the acquired images are brighter
121 than for SS, so for the SA data the gray level of the DMD was reduced so as not to saturate the
122 camera. We chose to keep the acquisition time constant to ease the comparison between SS and
123 SA and avoid any nonlinearity at the camera level. In practice, the camera exposure time could
124 be reduced for the SA data acquisition which would provide a further bonus with a reduction in
125 terms of acquisition time. The SS reconstruction is limited by the dynamic range of the camera
126 and its accuracy is limited for darker regions within the scene, where the signal to noise ratio is low,
127 or where the signal is below the background level of the camera. Therefore, the SS data used for
128 comparison to SA combines two data sets, with slit width of 1 mirror and slit width of 2 mirrors, to
129 increase the dynamic range of the measurement. When the panchromatic intensity of the scene falls
130 below a given threshold, the acquisition data from the wider slit is used to reconstruct the spectrum
131 (further details in Supplemental Document). The reconstruction with slit width of 2 mirrors will
132 incidentally reduce the spectral resolution.

133 Similarly to compressed sensing approaches, the nature of the data acquisition also gives us
134 the Fellgett’s advantage: compared to the SS acquisition, the SA acquisition has a higher SNR in
135 darker regions.

To compare the two reconstructions we use the Spectral Angle Mapper (SAM), which measures
the similarity of two spectra independent of their amplitude:

$$\text{SAM}(r, c) = \left| \arccos \left(\frac{\mathbf{o}(r, c)^T \cdot \hat{\mathbf{o}}(r, c)}{\|\mathbf{o}(r, c)\| \cdot \|\hat{\mathbf{o}}(r, c)\|} \right) \right| \quad (7)$$

136 At a given position (r, c) in the scene, the SAM is equal to zero when the spectra are identical,
137 and equal to $\pi/2$ when the spectra are exactly orthogonal. The SAM is an ideal metric for this
138 approach as it measures the spectral correspondence, rather than the spatial correspondence which
139 is easily preserved by the geometry of the system.

140 3.2 Segmentation of the panchromatic image

141 Many segmentation algorithms are available, the latest of state of the art resorting to neural net-
142 works [16, 17, 18, 19]. However, the focus of this paper is not on the segmentation algorithm itself,
143 and we would ideally like that the method is robust enough that obtaining the exact segmentation is
144 not crucial for the quality of the results (see Section 5). In practice, the panchromatic image is seg-
145 mented into homogeneous closed regions using the classical watershed algorithm [20], implemented
146 using MATLAB functions `watershed`. To alleviate noise, we first used an anisotropic diffusion algo-
147 rithm [21], implemented using MATLAB function `anisodiff2D`, with gradient modulus threshold
148 $\kappa = 2 \times 10^{-3}$, and 50 iterations. Anisotropic diffusion denoises the image, while preserving the con-
149 tours, followed by a morphological gradient to improve the region-oriented segmentation, removing
150 the gradients below a threshold T , which varies the segmentation sensitivity. The segmentation al-
151 gorithm returns the detected homegenous regions in the scene, which are bounded by a contour that

152 is one pixel wide. The contours are likely to contain a spectral mixture of the two adjacent regions,
153 but to focus primarily on the separability assumption they are discarded during the reconstruction
154 of the HS image.

155 4 Results

156 4.1 Simple Scene

157 We performed initial tests on a simple and easily segmented scene, consisting of a plastic brick wall
158 illuminated with a white LED, shown in Figure 2a. The RGB image from the slit scanning (SS)
159 reconstruction is shown in Figure 2b, and from the SA algorithm in Figure 2c, for $N = 5$ coded
160 acquisitions, the region contour highlighted in white. The DMD gray level was 255 for the SS
161 acquisition and 140 for the SA acquisition. The panchromatic image was normalized between 0
162 and 1 and then segmented with a threshold of $T = 1 \times 10^{-2}$. Two examples of the reconstructed
163 spectra are shown in Figure 3, showing a good correspondence especially for the bright regions
164 (Figure 3a). Figure 4 displays the histogram and average of the pixel-by-pixel SAM values in the
165 scene, dependent on the number of acquisitions. As one can see, only a few acquisitions are required
166 for the results to converge to a good accuracy, with a low average SAM.

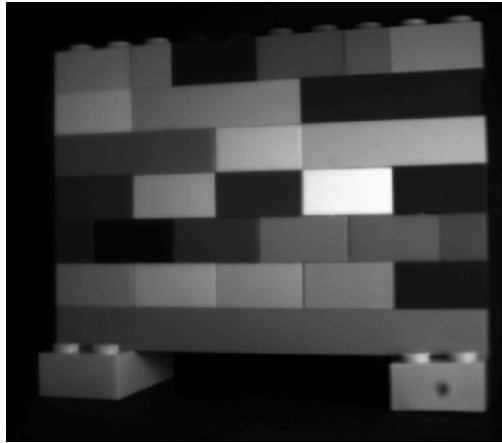
167 The reconstructed data cube has a total of $Q = 360$ regions, which means that the number
168 of data points required to represent the scene is $(Q \times W) + (R \times C) = 2 \times 10^5$, whereas the full
169 datacube has size to $R \times C \times W = 1.76 \times 10^7$, pixels. This gives a total compression factor of 88
170 for the hyperspectral data cube. In terms of data acquisitions, the SA method requires $N + 1 = 6$,
171 compared to a normal slit scan with $W = 110$ acquisitions (i.e. not high dynamic range), so the
172 compression factor for the acquisition data is $W/(N + 1) \simeq 18$. In terms of overall acquisition time
173 for the whole sequence of images SA is 18 times faster. The reconstruction time for the SA method
174 was 15 seconds, using non-optimized Matlab code.

175 This example also illustrates well the two main failure modes of the SA; firstly via regions that
176 are too small to contain enough information, and secondly via adjacent metamers.
177 For very small regions (such as the bottom left of Figure 2c), we can see on the RGB that the
178 reconstructed spectra are inaccurate, as the reconstruction problem is under-determined, despite
179 the regularization.

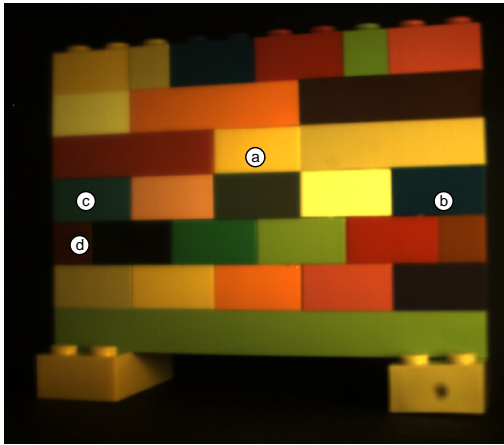
180 For adjacent metamers, Figure 5 shows spectra for two different pixels in the same region,
181 although by looking at the RGB image we can see that they are actually from different colored
182 bricks (points (c) and (d) in Figure 2b). When the scene contains two spectrally different zones,
183 which are combined into a single region during segmentation, the reconstructed spectrum is a
184 mixture of the spectra of the two individual zones. This is clear when we compare the spectrum
185 constructed by SA (dotted red line in Figure 5) to the average spectral value for that region found
186 by SS (solid red line in Figure 5).

187 4.2 Complex Scene

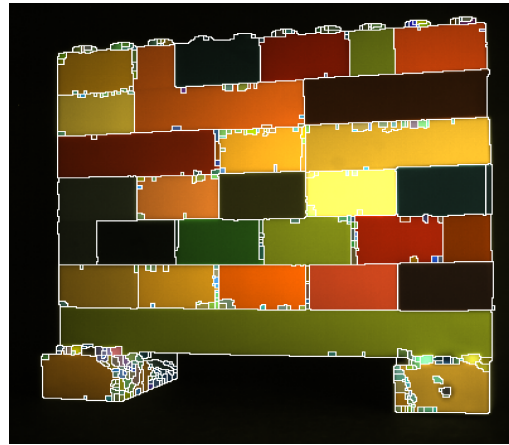
188 Now we examine the performance with a more complex scene. The scene consists also of plastic
189 bricks but has a 3D structure, gradients in light and dark due to shadows, and specular reflections,
190 so is difficult to segment accurately. The DMD gray level was 255 for the SS acquisition and 125
191 for the SA acquisition. The RGB reconstruction of the scene using the slit scanning data is shown



(a)



(b)



(c)

Figure 2: (a) Panchromatic image, and RGB image for (b) slit scanning and (c) the SA method with $N = 5$ acquisitions for the simple brick scene.

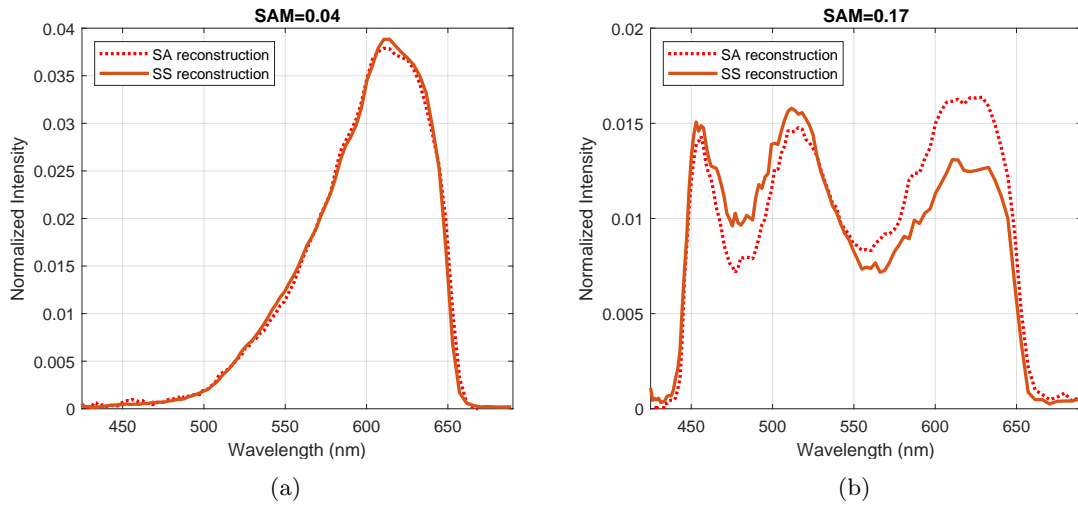


Figure 3: Comparison of reconstructed SA spectra to the SS spectra, for pixels labelled (a) and (b) in Figure 2c.

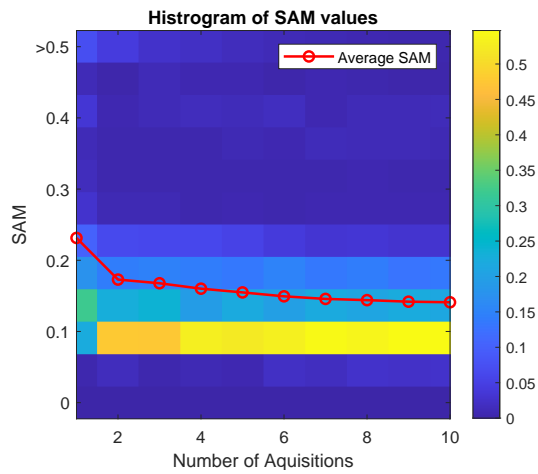


Figure 4: Histogram of SAM values with number of acquisitions.

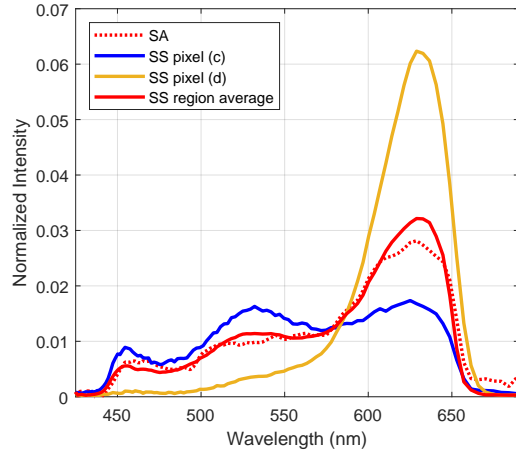


Figure 5: Comparison of reconstructed SA spectra to the SS spectra for pixels (c) and (d) labeled in Figure 2c, and for the average SS spectrum over the region.

192 in Figure 6a. Figures 6b, 6c and 6d show the reconstructed RGB results for the SA algorithm with
 193 varying segmentation thresholds.

194 Figure 7a shows the SAM histogram and average SAM as a function of the number of acquisitions,
 195 similar to the simple scene, although now as there are more small regions the convergence
 196 with N is somewhat slower. For $N = 5$, Figure 7b shows that the reconstruction accuracy does
 197 not change much with the segmentation threshold, and there is a broad optimum region between
 198 $T = 0.001$ and $T = 0.015$. For an over segmented scene we would expect more small regions which
 199 do not contain enough information to be reconstructed accurately, whereas for the under segmented
 200 scene we are more likely to have two different regions merged into one, so the resulting spectra is
 201 a mixture of the two. Figure 7c shows how the average SAM of the region depends on the size of
 202 the region K , showing a large increase in the value and the variation in SAM as the region gets
 203 smaller. For a fully 2D mask we know the reconstruction problem can be over determined when
 204 $K > W/N = 22$, i.e. for regions with more that 22 pixels we should have enough information to
 205 accurately reconstruct the spectrum. However, as we are using only a 1D mask, we should expect
 206 the limit on K to be somewhat higher, depending on the shape of the region. Of course, regular-
 207 ization over the spectrum reduces the limit, depending on the smoothness of the spectra and the
 208 regularization parameter chosen. These results indicate that obtaining perfect scene segmentation
 209 is not vital to the success of the algorithm.

210 For the best segmentation, the reconstructed data cube has a total of 1138 regions, which means
 211 that the number of data points required to represent the scene is $(Q \times W) + (R \times C) = 2.85 \times 10^6$,
 212 giving a total compression factor of 6.2 for the hyperspectral data cube. As for the simple scene,
 213 SA method permits a $\times 18$ reduction in the amount of data needed for reconstruction compared to
 214 the SS method. The reconstruction time for the SA method was 11 seconds (Matlab code).

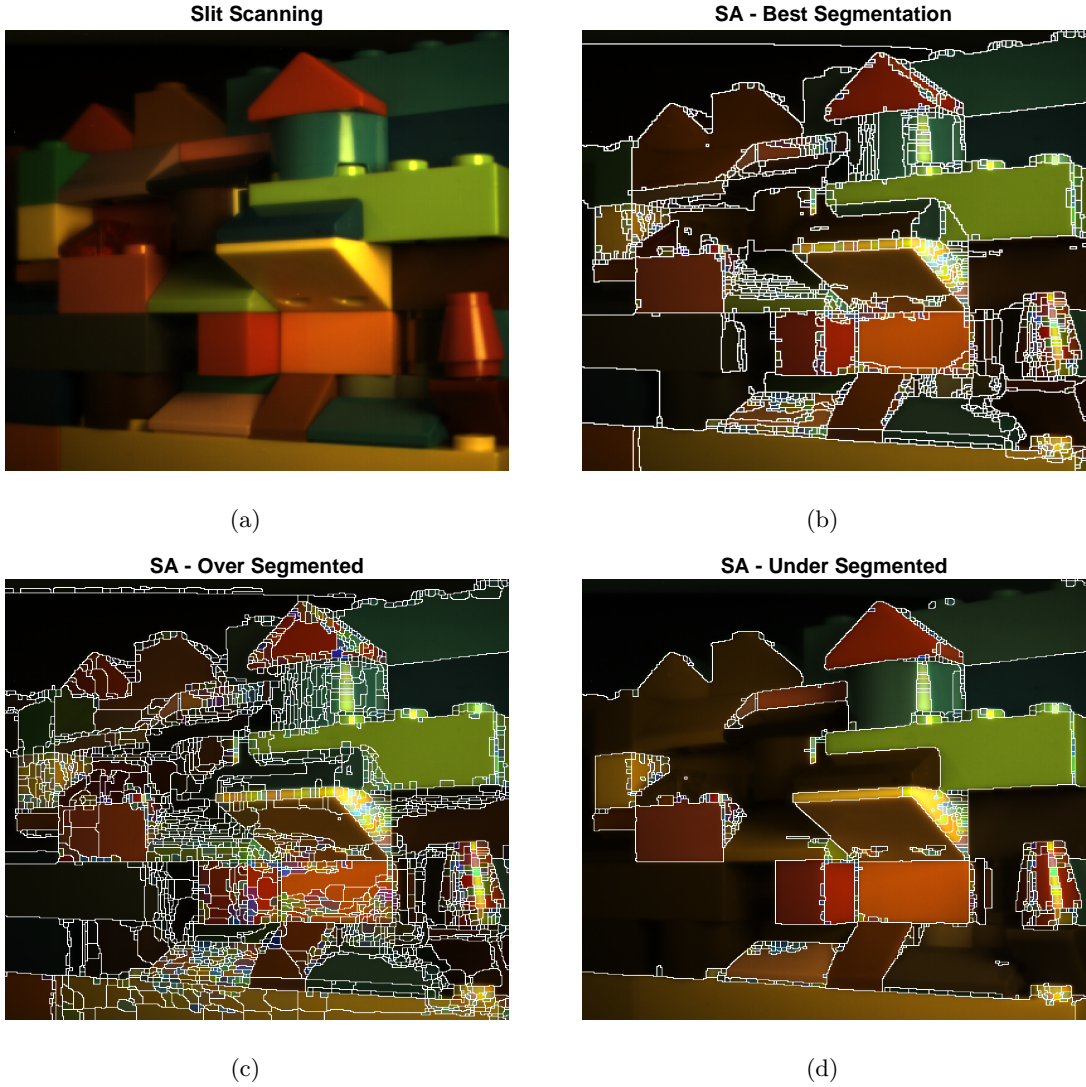


Figure 6: RGB images of the complex brick scene using (a) SS method, and (b-c) the SA method with $N = 5$ acquisitions and segmentation threshold (b) $T = 0.008$ (best segmentation in terms of average SAM), (c) $T = 0.001$ (over-segmented), and (d) $T = 0.02$ (under segmented).

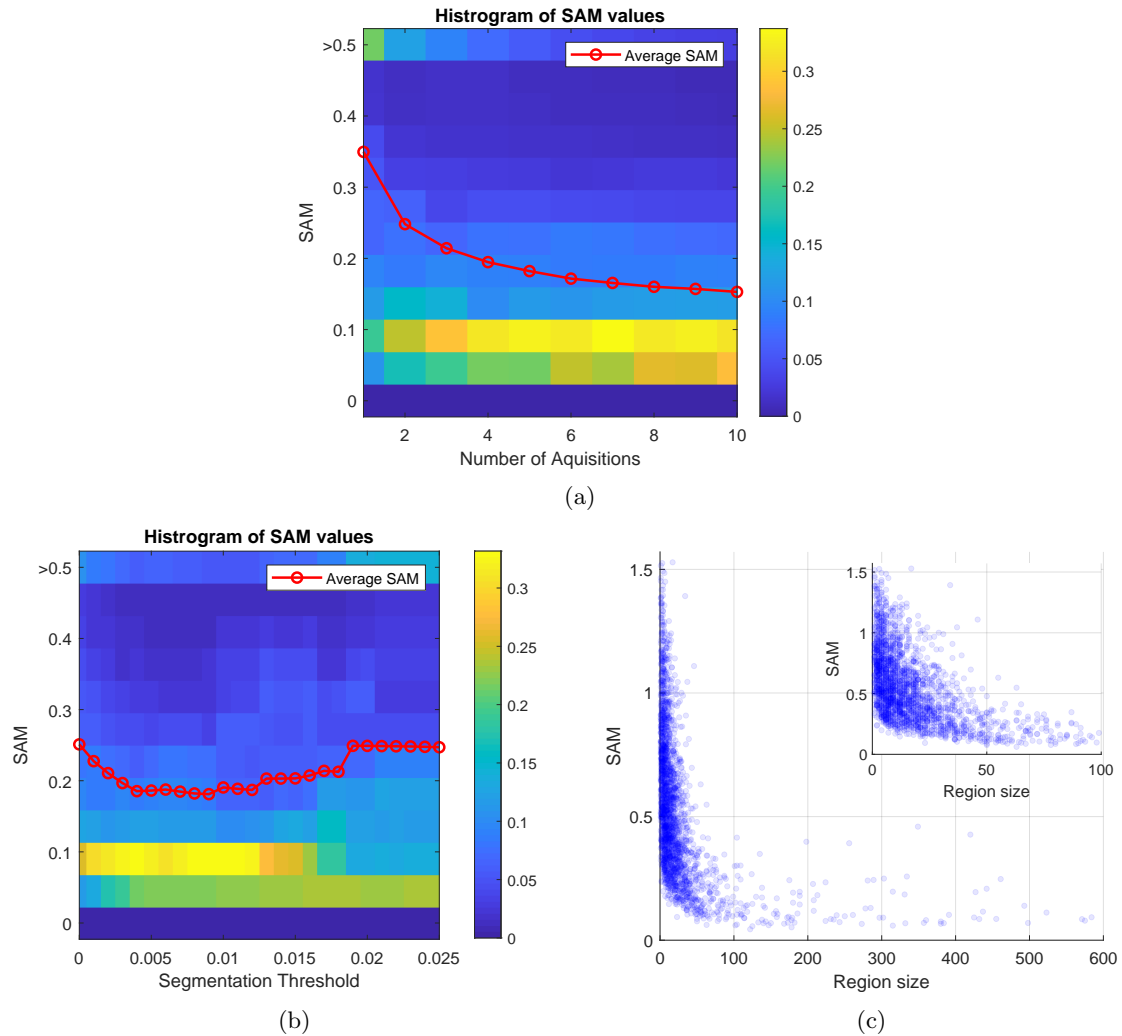
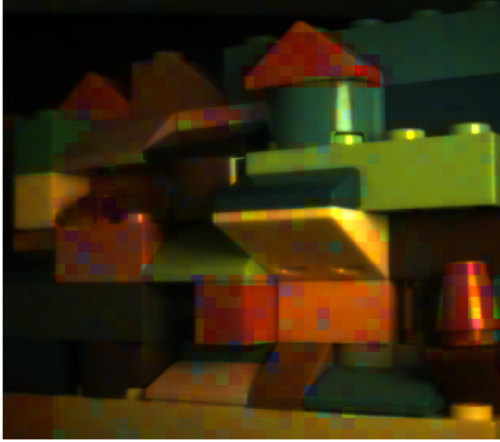


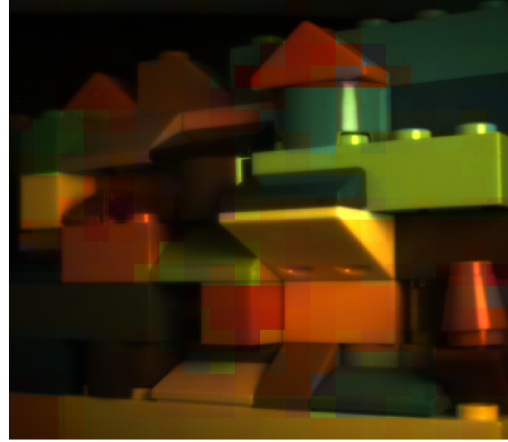
Figure 7: Histogram of SAM values: (a) as a function of the number of acquisitions N for $T = 0.008$, and (b) as a function of the segmentation threshold for $N = 5$; (c) For $N = 5$, average SAM over a region as a function of the region size.

Region size = 10 x 10 pixels, average SAM= 0.24

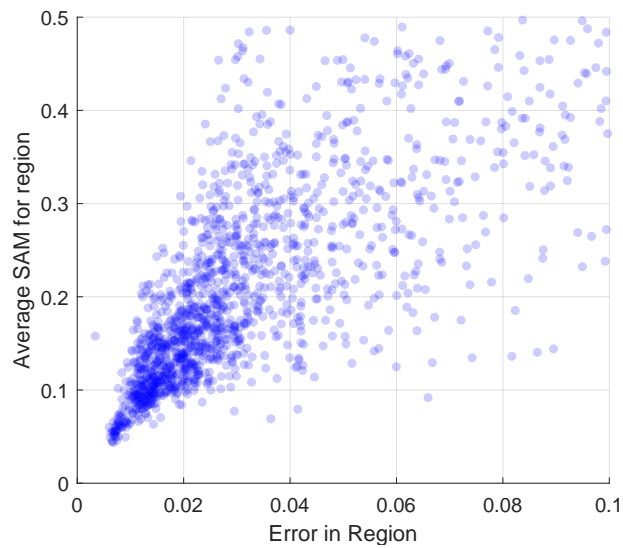


(a)

Region size = 20 x 20 pixels, average SAM= 0.18



(b)



(c)

Figure 8: RGB reconstruction using SA methods on square regions or 'superpixels', region size (a) 10×10 pixels, and (b) 20×20 pixels, and $N = 5$ acquisitions. (c) the average SAM plotted against the residual error R of a region, for 10×10 regions.

215 5 Beyond Panchromatic Segmentation

216 As we have seen, the results are somewhat insensitive to the segmentation, and it may be possible
217 to avoid panchromatic segmentation entirely. For example, Figures 8a and 8b show the results for
218 the SA approach when we arbitrarily divide the scene into 10×10 or 20×20 square regions. We
219 can gain information on how well each region is reconstructed spectrally using the residual error,
220 given by

$$R = \left\langle \left| \frac{\mathbf{m} - \mathbf{G}\hat{\mathbf{s}}}{\mathbf{m} + \mathbf{G}\hat{\mathbf{s}}} \right| \right\rangle \quad (8)$$

221 Where the average $\langle \cdot \rangle$ is over all the pixels of the region and over all the acquisitions. As Figure 8c
222 shows, the residual gives us information on how well a region is reconstructed, although there is a
223 large variation in the results. We could use the residual error to post-process the reconstruction
224 and improve the segmentation and overall accuracy of the reconstruction. We could also rely on
225 the residual error to devise a split/merge strategy to refine the SA reconstructed HS cube, starting
226 from some arbitrary division into regions. This paves the way towards implementation of SA that
227 do not suffer from the limitations induced by segmentation of the panchromatic image.

228 6 Conclusion

229 The reconstruction algorithm is based on the particular properties of the dual disperser hyperspec-
230 tral imager, namely the property of co-registration, and relies on spectral-spatial correlations in the
231 scene. The algorithm reconstructs simple real-world scenes accurately and quickly using a small
232 amount of data, also with an acquisition time much shorter than for a slit scanning acquisition.
233 Full parallelization of the computation could theoretically reduce the computation time by a fac-
234 tor equal to the number of regions. Experimental improvements allowing a 2D mask, as well as
235 investigation into the optimum ROM could reduce the number of acquisitions necessary or allow
236 smaller regions to be reconstructed. Currently, due to the reliance on panchromatic segmentation,
237 this method may not be well suited to challenging scenes that are difficult to segment *e.g.* highly
238 textured scenes or with lots of metamers. However, preliminary tests using arbitrary division of the
239 scene into regions show good results, and demonstrates that we can use the residual error as a met-
240 ric on reconstruction accuracy. Future work involves avoiding reliance on imperfect segmentation
241 algorithms, instead utilizing the residual errors to improve the reconstruction and the segmentation
242 itself.

243 **Funding.** This work was funded by the Agence Nationale de la Recherche (ANR-18-ASTR-0012-01
244 (ImHypAd project))

245 **Disclosures.** The authors declare no conflicts of interest.

246 **Data availability.** Data underlying the results presented in this paper are not publicly available at this
247 time but may be obtained from the authors upon reasonable request.

248 **Supplemental document.** See Supplement 1 for supporting content.

References

- [1] Sérgio MC Nascimento, Kinjiro Amano, and David H Foster. Spatial distributions of local illumination color in natural scenes. *Vision Research*, 120:39–44, 2016.
- [2] Ayan Chakrabarti and Todd Zickler. Statistics of real-world hyperspectral images. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 193–200, 07 2011.
- [3] Ying Fu, Yinqiang Zheng, Imari Sato, and Yoichi Sato. Exploiting spectral-spatial correlation for coded hyperspectral image restoration. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3727–3736, 2016.
- [4] Ashwin Wagadarikar, Renu John, Rebecca Willett, and David Brady. Single disperser design for coded aperture snapshot spectral imaging. *Applied Optics*, 47(10):B44–B51, 2008.
- [5] Ashwin A Wagadarikar, Nikos P Pitsianis, Xiaobai Sun, and David J Brady. Video rate spectral imaging using a coded aperture snapshot spectral imager. *Optics Express*, 17(8):6368–6388, 2009.
- [6] David Kittle, Kerkil Choi, Ashwin Wagadarikar, and David J Brady. Multiframe image estimation for coded aperture snapshot spectral imagers. *Applied Optics*, 49(36):6824–6833, 2010.
- [7] Gonzalo R Arce, David J Brady, Lawrence Carin, Henry Arguello, and David S Kittle. Compressive coded aperture spectral imaging: An introduction. *IEEE Signal Processing Magazine*, 31(1):105–115, 2013.
- [8] Ziyi Meng and Xin Yuan. Perception inspired deep neural networks for spectral snapshot compressive imaging. In *2021 IEEE International Conference on Image Processing (ICIP)*, pages 2813–2817, 2021.
- [9] Lizhi Wang, Tao Zhang, Ying Fu, and Hua Huang. Hyperreconnet: Joint coded aperture optimization and image reconstruction for compressive hyperspectral imaging. *IEEE Transactions on Image Processing*, 28(5):2257–2270, 2018.
- [10] Hao Zhang, Xu Ma, Xianhong Zhao, and Gonzalo R Arce. Compressive hyperspectral image classification using a 3d coded convolutional neural network. *Optics Express*, 29(21):32875–32891, 2021.
- [11] Michael E Gehm, Renu John, David J Brady, Rebecca M Willett, and Timothy J Schulz. Single-shot compressive spectral imaging with a dual-disperser architecture. *Optics Express*, 15(21), 2007.
- [12] Elizabeth Hemsley, Simon Lacroix, Hervé Carfantan, and Antoine Monmayrant. Calibration of programmable spectral imager with dual disperser architecture. *Optics Communications*, page 125767, 2020.
- [13] Jérôme Idier. *Bayesian approach to inverse problems*. John Wiley & Sons, 2013.
- [14] Andrey N. Tikhonov and Vasilii Y. Arsenin. *Solutions of ill-posed problems*. V. H. Winston & Sons, 1977.

- 286 [15] Elizabeth Hemsley, Ibrahim Ardi, Simon Lacroix, Hervé Carfantan, and Antoine Monmayrant.
287 Optimized coded aperture for frugal hyperspectral image recovery using a dual-disperser sys-
288 tem. *JOSA A*, 37(12):1916–1926, 2020.
- 289 [16] Nikhil R Pal and Sankar K Pal. A review on image segmentation techniques. *Pattern Recog-*
290 *nition*, 26(9):1277–1294, 1993.
- 291 [17] HP Narkhede. Review of image segmentation techniques. *International Journal of Science and*
292 *Modern Engineering*, 1(8):54–61, 2013.
- 293 [18] Catalin Amza. A review on neural network-based image segmentation techniques. *De Montfort*
294 *University, Mechanical and Manufacturing Engg., The Gateway Leicester, LE1 9BH, United*
295 *Kingdom*, pages 1–23, 2012.
- 296 [19] Nima Tajbakhsh, Laura Jeyaseelan, Qian Li, Jeffrey N Chiang, Zhihao Wu, and Xiaowei
297 Ding. Embracing imperfect datasets: A review of deep learning solutions for medical image
298 segmentation. *Medical Image Analysis*, 63:101693, 2020.
- 299 [20] J. B. T. M. Roerdink and A. Meijster. The watershed transform: Definitions, algorithms and
300 parallelization strategies. *Fundamenta Informaticae*, 41(1, 2):187–228, January 2000.
- 301 [21] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE*
302 *Transactions on Pattern Analysis and Machine Intelligence*, 12(7):629–639, 1990.