Compressive Hyperspectral Imaging Mask Optimization

Binbin Lv
Department of Automation
Hangzhou Dianzi University
Hangzhou, China
161060026@hdu.edu.cn

Jiamin Wu
Department of Automation
Tsinghua University
Beijing, China
wujm14@mails.tsinghua.edu.cn

Chenggang Yan
Department of Automation
Hangzhou Dianzi University
Hangzhou, China
cgyan@hdu.edu.cn

Xinhong Hao
Science and Technology on Mechatronic Dynamic Control Laboratory, Beijing Institute of Technology
Beijing, China
haoxinhong@bit.edu.cn

Guiguang Ding
School of Software, Tsinghua University
Beijing, China
dinggg@tsinghua.edu.cn

Yongdong Zhang
School of Information Science and Technology, University of Science and Technology of China
Beijing, China
zhyd@ict.ac.cn

Qionghai Dai
Department of Automation,
Tsinghua University
Beijing, China
qhdai@tsinghua.edu.cn

ABSTRACT

Hyperspectral imaging is a hot topic nowadays. It is an urgent problem to be solved how to achieve swift hyperspectral imaging. In this thesis, our primary purpose is to further optimize how to place a mask in front of a sensor in order to achieve compressed Hyperspectral imaging. We apply optimized projection matrix, matrix differential, projection analysis and other related knowledge to optimizing this realistic matter. After we simply introduce the background of hyperspectral imaging, we will firstly present the basic principle of compressed hyperspectral imaging based on mask, and then mainly analyze the way to achieve projection matrix optimizing algorithm and the challenges these sort of realistic problems face. Finally, we compare the experiment results of these two methods, and the rebuilding results before and after the optimizing.

KEYWORDS

hyperspectral imaging, optimized mask, compressive sensing

ACM Reference Format:
Binbin Lv, Jiamin Wu, Chenggang Yan, Xinhong Hao, Guiguang Ding, Yongdong Zhang, and Qionghai Dai. 2018. Compressive Hyperspectral Imaging Mask Optimization. In *Proceedings of ACM Conference (ICIMCS'18).* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3240876.3240921

ICIMCS'18, August 17-19,2018, Nanjing, China

ACM ISBN 978-1-4503-6520-8/18/08…$15.00
https://doi.org/10.1145/3240876.3240921

1 INTRODUCTION

With the increasing requirement of the hyperspectral data in various applications such as remote sensing[19] and computer vision[4, 7, 12, 15, 17], lots of high-speed hyperspectral imaging techniques have been proposed during the past decades. The primary job of hyperspectral imaging is to efficiently collect and acquire a 3D data cube transformed from 2D space and 1D spectral transformation. To achieve this objective, conventional methods record the data through mechanical scanning along spatial dimension[1] and spectral dimension[18], respectively. Snapshot hyperspectral imaging collects the whole 3D data through a single image, which has a significant advantage over scanning-based methods for dynamic scene or aerial photography. Conventional snapshot hyperspectral methods multiplex high-dimensional signals separately into a 2D sensor. Therefore, the total pixel number of the 3D data cube is sacrificed. Four-dimensional imaging spectrometer (4DIS) [6] and computed tomography imaging spectrometer (CTIS) [9] are two typical examples. Recently, coded aperture snapshot spectrometer (CASSI) [8] applies compressive sensing to encode the 3D data cube into a 2D image and reconstructs with the prior of the data sparsity. This computational method overcomes the compromise between the spatial resolution and spectral resolution with a much higher light efficiency. CASSI system uses different fixed coded masks. Another more flexible choice is to apply digital mirror device (DMD) to change the modulation pattern[10]. And different kinds of coding frameworks[11, 16] are then proposed to increase the reconstruction performance

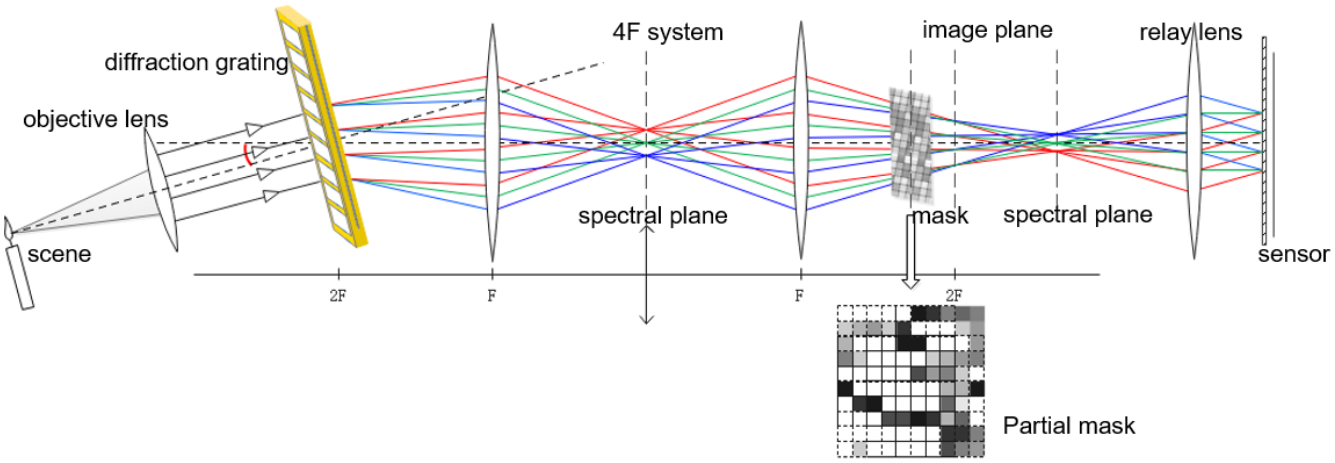

Figure 1: Schematic of the spatial-spectral encoded compressive hyperspectral imaging.

and acquisition speed. In the mean time, many works have shown that the optimization of the coding pattern will greatly improve the restoration performance[13]. However, all these compressed spectral imaging methods use random mask or empirically chosen masks for modulation and very few quantitative analysis is provided for the influence of the mask. In this paper, on the basis of a typical compressive hyperspectral imaging technique[10], we exploit the effect of the mask with matrix analysis and projection optimization and try to optimize the mask for better compressive reconstruction, which may also be applied to other high-dimensional multimedia acquisition framework and inspire more applications. Simulation results on the public hyperspectral dataset[3] prove that the performance can be enhanced with the optimized mask.

2 MATHEMATICAL MODEL

We used the compressive hyperspectral imaging setup as depicted in Fig. 1[10]. Fig. 1 shows the schematic of the system. A diffraction grating was placed at the first image plane to disperse the spectral information into different angles, mapping the spectral information of the scene into angular dimension. Then we placed a mask in front of the image plane and relayed the coded image plane to a sensor for detection. Based on the derivation of geometrical optics, each spectral band of the image corresponds to a specific different coding pattern. Due to the different angles of the spectral band, the coding pattern for different spectrums has a translational shift. The optimized mask is shown in Fig.1, the mask is not binarized, but has a gradual transmittance.

A mathematical model can be established for the imaging process. The coded image i on the sensor can be represented as:

$$i(x,y) = \int_{\Omega_\lambda} h(x,y,\lambda)d\lambda \tag{1}$$

where, h here represents for the 3D hyperspectral data we want to retrieve, and x,y corresponds to the spatial domain and λ corresponds to the spectral domain. The shift of the pattern related to the distance between the mask and the sensor is represented by s.

One can easily recognize that the shape of an object is very similar in different spectral bands while their reflectance may be different. This kind of sparsity can be further exploited by learning an over-complete dictionary[14] based on a large hyperspectral dataset. By employing KSVD method, we obtain the over-complete dictionary on which the hyperspectral information can be represented by several bases, as shown below:

$$h = D\alpha = \sum_{i=1}^{q} d_j \alpha_j \tag{2}$$

The α here meets the sparse constraint and the D here means the over complete dictionary. Then we can further solve the optimizing problem as the later formula to reconstruct the encoded hyperspectral data

$$min_\alpha \|\alpha\|_1 \qquad s.t. \qquad \|i - \Phi D\alpha\|_2^2 \leq \varepsilon \tag{3}$$

Here Φ corresponds to the projection matrix, as illustrated by formula 2. Our projection matrix P is then presented as formula 5:

$$P = \Phi D \tag{4}$$

As we can see the projection matrix above largely depends on mask pattern and the over-complete dictionary. It indicates that different over-complete dictionaries may have different performance with the fixed mask pattern. And a bad mask pattern may make the projection matrix ill-posed and lose the robustness of noise. Though the dimension of the projection matrix is large, the unknown parameters are very limited, which is equal to the size of the mask. For reconstruction process, we usually do it in a patch by patch mode, which is more suitable for parallel computing. To make the masks we designed accessible for the reconstruction, the pattern itself should be periodical. Then we do not need to optimize the

Table 1: Quantitative evaluation of the performance enhancement with an optimized mask

Image Set	Optimization Method	MSE	HS Image PSNR(dB)	Main Angle Error(MAE)
'Chair' Data	Random mask	8.758	38.707	3.321°
	Optimized mask	7.687	39.273	3.181°
'Door' Data	Random mask	62.488	30.173	7.846°
	Optimized mask	51.051	31.051	7.672°

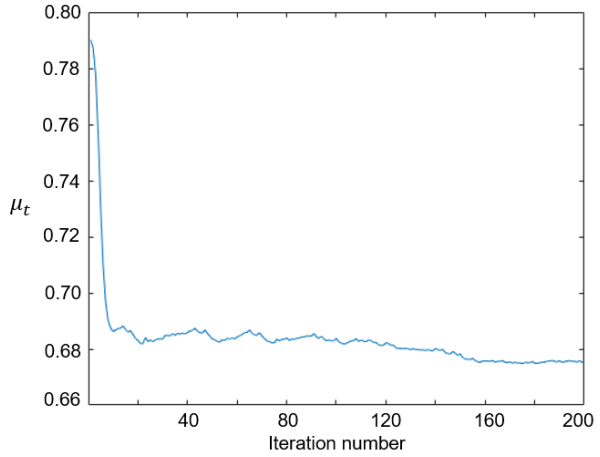

Figure 2: The optimized result of gradient-descent algorithm.

whole mask, but only to aim at a small block for dictionary training and mask optimization. Inspired by the compressed sensing methods[2], when the sample matrix meets a series of character, including mutual-coherence and finite isometric properties, the zero-norm problem can be solved through one-norm optimization. For better reconstruction, we need to make our projection matrix meet the mutual-coherence condition. In other words, we need to make the matrix more irrelative to reconstruct more efficiently.

By using the framework introduced by Elad et al.[5], we design a mask optimization algorithm for compressive hyperspectral imaging. The original method of projection matrix is an unconstrained optimization, for the application of our specific issues, the actual system places lots of strong constraints to the projection matrix. The real unknown variables are very limited for such a high-dimensional projection matrix. This kind constraints also brought some problems for our optimizing process, so we propose a new algorithm specifically designed for the compressive hyperspectral imaging.

First, in order to achieve a better reconstruction effect, we need to make the projection matrix satisfy mutual-coherence as much as possible.The stronger the mutual-coherence, the stronger the reconstruction ability. We use a common standard μ_t to measure the mutual-coherence of the matrix, and its detailed representation can be referred to [5].

For our problem, only the diagonal element of the projection matrix P is nonzero, which makes it different from the concrete optimization. Due to the Gram matrix is symmetric indefinite, we can not use the cholesky decomposition. Here we use the singular value decomposition (SVD) to get the approximate results. We applied gradient-descent algorithm to solve formula $|S_q - PD|_F^2$. The gradient of this formula can be calculated as:

$$\frac{df(x)}{dP^T} = -2DS_q + 2DD^T P^T \tag{5}$$

The solving process of the gradient-descent algorithm as follows: Step 0: To determine the initial point
Step 1: $\triangle x = - \bigtriangledown f(x)$
Step 2: linear searching, by using the backtracking line search method to determine the step length t.
Step 3: $x = x + \triangle x$.
Repeat step 1 to 3 until meet stop criterion.

The convergence curve of the algorithm is shown in Fig. 2. With the starting point of a random mask, the μ_t gradually converges to around 0.68. Then we did some numerical simulation to compare the results with random mask and optimized mask.

3 NUMERICAL SIMULATION

Based on the above analysis, we realized the algorithm and carried on some simulations to verify our method. We first read the corresponding matrix dictionary. Then the initial parameters are determined, including the size of required mask (according to the size of the dictionary), the parameters used in the process of linear search, etc. Finally, the above-mentioned procedures are undertaken to achieve several simulation results for performance analysis of our method.

The first simulation is to test the convergence speed of the optimization process using the 1-norm constraint. The μ_t convergence curve can converge to local optimal solution fast. In the process of linear search, all the values within the mask should be constrained between 0 to 1. Even a small change in each gradient process will lead the algorithm to converge to a local point and make it difficult for further optimization.

Different from the above optimization algorithm based on the 1-norm constraint, an improved method can be used for the fast convergence and a stable optimal value. The Convergence result of improved method as shown in Fig. 2,instead of using a gradient search in all dimensions after each contraction G, the improved gradient-descent algorithm just need to select one of the largest each time. Therefore, the μ_t curve can converge much faster.

Next, we compared the results of our optimized coding and the random coding. we only rebuild a certain part of a larger

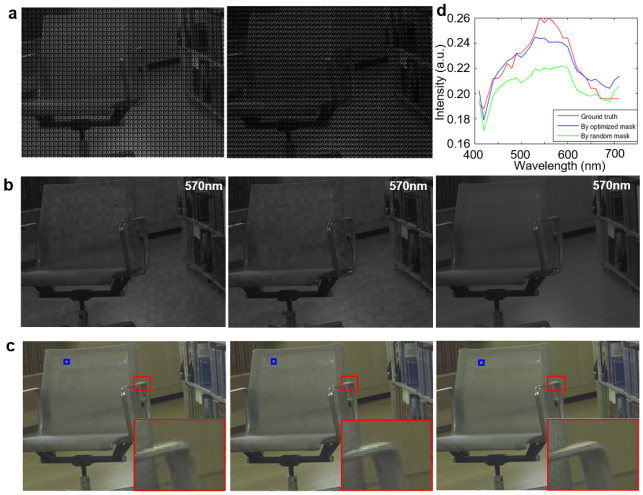

Figure 3: The comparison of the reconstruction by coded mask and the optimized mask on the door data. (a) are input spatial-spectral encoded sensor images, where the left is encoded by the random mask and the right is encoded by the optimized mask. (b) are reconstruction results, where the left and the middle images are using the random mask and the optimized mask, respectively. The right panel is the ground truth. (c) are the RGB synthesis images with their zoom-ins. The reconstruction with optimized mask (left panel), with random mask(middle panel) and the groundtruth(the right panel) are shown for comparision. Reconstruction results using our optimized mask provide better accuracy (PSNR=30.7dB) than using the random mask (PSNR=29.5dB) on the 'Door' data. The spectral profiles of a certain patch marked by a blue box are shown in (d). a.u., arbitary units.

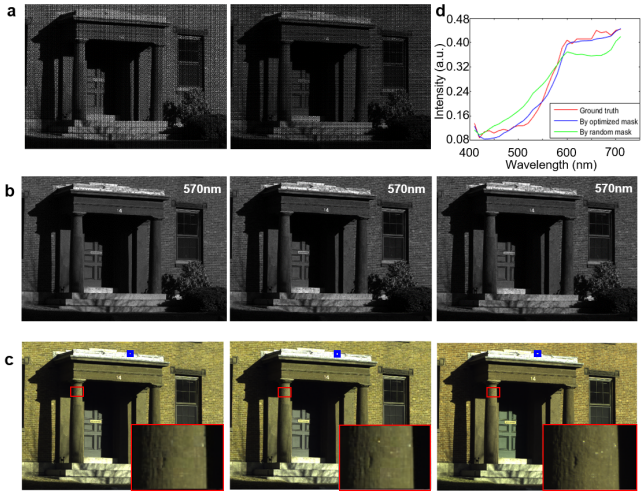

Figure 4: The comparison of the reconstruction by coded mask and the optimized mask on the chair data. (a) are input spatial-spectral encoded sensor images, where the left is encoded by the random mask and the right is by the optimized mask. (b) are reconstruction results, where the left and the middle images are respectively using the random mask and the optimized mask. The right of (b) is the ground truth. (c) are the RGB synthesis images by optimized mask (left panel), random mask (middle panel) and groundtruth (right panel). Reconstruction results using our optimized mask can apparently provide better accuracy (PSNR=37.5dB) than using the random mask (PSNR=36.2dB) on 'Chair' data. The spectral profiles of a certain patch using our optimized mask and the random mask on 'Chair' data are shown in (d). a.u., arbitary units.

image for a proof of concept. The reconstruction of a whole image still takes a lot time of computation. Reconstructed results are shown in Fig. 3 and Fig. 4. Results using the optimized mask are close to the optimized reconstruction result in terms of PSNR and reconstruction speed. Although the image contrast is low, the PSNR of the optimized coding can reach 30.7361 dB in the dataset of 'Door' and 37.5139 dB in the dataset of 'chair'. In comparison, the reconstruction of the random coding only reaches 29.5332 dB and 36.1903 dB respectively. If we have a good initial guess, the results are better than the least squares method and the total-least squares method. In table 1, we evaluated the reconstruction quality of the dataset 'Chair' and 'Door' using MAE, MSE and PSNR parameters.

4 CONCLUSION

In this paper, we propose a mask optimization algorithm for the application of compressive hyperspectral imaging. By the projection optimization framework and the gradient descent algorithm, we achieve a reduction and convergence of the mutual coherence coefficient. The reconstruction quality using the optimized mask has an obvious improvement from that using the random mask, in terms of PSNR and MAE. We believe that our idea of solving practical problems with constrained projection matrix optimization, can inspire more multimedia applications.

REFERENCES

[1] Robert W. Basedow. 1995. HYDICE system: implementation and performance. *Proceedings of SPIE - The International Society for Optical Engineering* 2480, 1 (1995), 258–267.
[2] Emmanuel J Candes and Terence Tao. 2005. Decoding by linear programming. *IEEE transactions on information theory* 51, 12 (2005), 4203–4215.
[3] A. Chakrabarti and T. Zickler. 2011. Statistics of real-world hyperspectral images. In *Computer Vision and Pattern Recognition*. 193–200.
[4] Hao Du, Xin Tong, Xun Cao, and Stephen Lin. 2010. A prism-based system for multispectral video acquisition. In *IEEE International Conference on Computer Vision*. 175–182.
[5] Michael Elad. 2007. Optimized projections for compressed sensing. *IEEE Transactions on Signal Processing* 55, 12 (2007), 5695–5702.

[6] Nahum Gat, Gordon Scriven, John Garman, Ming De Li, and Jingyi Zhang. 2006. Development of four-dimensional imaging spectrometers (4D-IS). In *SPIE Optics+ Photonics*. International Society for Optics and Photonics, 63020M–63020M.

[7] Ling Ge, Ran Ju, Tongwei Ren, and Gangshan Wu. 2015. Interactive RGB-D Image Segmentation Using Hierarchical Graph Cut and Geodesic Distance. In *Pacific Rim Conference on Multimedia*. 114–124.

[8] ME Gehm, R John, DJ Brady, RM Willett, and TJ Schulz. 2007. Single-shot compressive spectral imaging with a dual-disperser architecture. *Optics express* 15, 21 (2007), 14013–14027.

[9] William R Johnson, Daniel W Wilson, and Greg Bearman. 2006. Spatial-spectral modulating snapshot hyperspectral imager. *Applied optics* 45, 9 (2006), 1898–1908.

[10] Xing Lin, Yebin Liu, Jiamin Wu, and Qionghai Dai. 2014. Spatial-spectral encoded compressive hyperspectral imaging. *ACM Transactions on Graphics (TOG)* 33, 6 (2014), 233.

[11] X. Lin, G Wetzstein, Y. Liu, and Q. Dai. 2014. Dual-coded compressive hyperspectral imaging. *Optics Letters* 39, 7 (2014), 2044.

[12] Jing Liu, Tongwei Ren, Yuantian Wang, Sheng Hua Zhong, Jia Bei, and Shengchao Chen. 2016. Object proposal on RGB-D images via elastic edge boxes. *Neurocomputing* 236 (2016).

[13] D. Schonfeld and J. Goutsias. 1991. Optimal Morphological Pattern Restoration from Noisy Binary Images. *Historical Methods A Journal of Quantitative and Interdisciplinary History* 7, 3 (1991), 225–244.

[14] Y. Sun, Q. Liu, J. Tang, and D. Tao. 2014. Learning Discriminative Dictionary for Group Sparse Representation. *IEEE Trans Image Process* 23, 9 (2014), 3816–3828.

[15] J. Tang, Z. Li, M. Wang, and R. Zhao. 2015. Neighborhood Discriminant Hashing for Large-Scale Image Retrieval. *IEEE Transactions on Image Processing* 24, 9 (2015), 2827–2840.

[16] Jiamin Wu, Xiong Bo, Lin Xing, Jijun He, Jinli Suo, and Qionghai Dai. [n. d.]. Snapshot Hyperspectral Volumetric Microscopy. *Scientific Reports* 6 ([n. d.]), 24624.

[17] Xiangyang Xu, Yuncheng Li, Gangshan Wu, and Jiebo Luo. 2017. Multi-modal Deep Feature Learning for RGB-D Object Detection. *Pattern Recognition* 72 (2017).

[18] Masahiro Yamaguchi, Hideaki Haneishi, Hiroyuki Fukuda, Junko Kishimoto, and Hiroshi Kanazawa. 2008. High-fidelity video and still-image communication based on spectral information: natural vision system and its applications. In *Spectral Imaging: Eighth International Symposium on Multispectral Color Science*. 60620G–60620G–12.

[19] Daniel K. Zhou, Henry E. Revercomb, Allen M. Larar, Hung Lung Huang, and Bormin Huang. 2001. Hyperspectral remote sensing of atmospheric profiles from satellites and aircraft. *Proceedings of SPIE - The International Society for Optical Engineering* 4151 (2001), 94–102.