Compressive hyperspectral image reconstruction with variational autoencoders

Mark Nishimura

February 13, 2018

1 Motivation

Hyperspectral imaging is a type of imaging aimed at capturing both the spectral and spatial content of a scene. Instead of a 3-channel RGB image, the output is a 3D image volume with two spatial dimensions and a third spectral dimension along which the wavelength varies, making hyperspectral imaging useful for a variety of purposes in astronomy, surveillance, and microscopy.

Due to redundancy in the information content of different frequencies, it is possible to capture hyperspectral images in a compressive fashion, in which each pixel of the sensor captures light from many different frequencies, and compressive sensing is used to fill in the gaps. Clever architectures such as CASSI (Coded Aperture Snapshot Spectral Imager) allow this multiplexing to happen in a single shot, eliminating the need to scan slowly over the spatial or spectral domains.

Convolutional neural networks have been applied to great success at various imaging-related tasks including denoising and inpainting [6]. In particular, convolutional autoencoders are a type of deep neural network which learns latent representations of images in an unsupervised fashion by simultaneously training an encoder $\boldsymbol{E}()$ and a decoder $\boldsymbol{D}()$ such that $i = \boldsymbol{D}(\boldsymbol{E}(i))$, and such that $\alpha = \boldsymbol{E}(i)$ is a relatively compact representation. The α vector is the latent representation of the image i.

2 Related Work

Compressive sensing techniques previously applied to the hyperspectral image problem include Approximate Message Passing (AMP-3D-Wiener, a variant of a technique from graphical models) [5], Two-step iterative shrinkage/thresholding (TwIST) [1], and Gradient Projection for Sparse Reconstruction (a faster method for solving the basis pursuit problem) [4].

3 Project Overview

This project aims to replicate and extend the results from Choi et. al. 2017 [3]. In most compressive sensing setups, the main step is to identify a transform Φ such that the image $i = Ax = A\Phi s$ is sparse in the s-domain. The main idea in this paper is to train a convolutional autoencoder from existing hyperspectral image data to learn (nonlinear) Encoding and Decoding functions $\boldsymbol{E}()$ and $\boldsymbol{D}()$. such that $\alpha = \boldsymbol{E}(x)$ is compact (but not necessarily sparse). Once such functions have been learned, the reconstruction of a new image is performed by solving the following convex optimization problem. Given input data (e.g. captured using CASSI) i, we seek the latent vector α^* that is the solution of:

$$\min_{\alpha} \|i - \Phi \boldsymbol{D}(\alpha)\|_2^2 + \tau_1 \|\alpha - \boldsymbol{E}(\boldsymbol{D}(\alpha))\|_2^2 + \tau_2 \|\nabla_{xy} \boldsymbol{D}(\alpha)\|_1$$

This objective can be summarized as the sum of

1. A data fidelity term,
2. A prior on the quality of the learned autoencoder, and
3. A (spatial) TV prior on the reconstructed 3D block $\boldsymbol{D}(\alpha)$.

In the paper, the authors solve this problem with ADMM - we will likely do the same.

3.1 Notes on Data and Implementation

We will use the Harvard [2] and Columbia [7] hyperspectral image datasets, in addition to Choi et. al.'s new dataset, to train our autoencoder. For the autoencoder, itself, we will reconstruct the model using PyTorch (the authors provide a TensorFlow model). Fortunately, the Choi paper is relatively thorough in its description of its architecture and hyperparameters used to tune the model. We also have access to suitable compute (a 1080 Ti GPU) to train the model in a reasonable amount of time.

4 Milestones

The timeline for this project is roughly as follows:

- End of Week 6 - Obtain dataset, begin training the autoencoder.
- End of Week 7 - Finish training the autoencoder and begin the ADMM implementation.
- End of Week 8 - Finish the Choi et. al. reproduction, think about possible extensions.
- End of Week 9 - Try one or two extensions, begin the writeup and poster.
- End of Week 10 - Poster presentation.

References

[1] J. M. Bioucas-Dias and M. A. T. Figueiredo. A new twist: Two-step iterative shrinkage/thresholding algorithms for image restoration. *IEEE Transactions on Image Processing*, 16(12):2992–3004, Dec 2007.

[2] A. Chakrabarti and T. Zickler. Statistics of real-world hyperspectral images. In *CVPR 2011*, pages 193–200, June 2011.

[3] I. Choi, D. S. Jeon, G. Nam, D. Gutierrez, and M. H. Kim. High-quality hyperspectral reconstruction using a spectral prior. *ACM Transactions on Graphics (Proc. SIGGRAPH Asia 2017)*, 36(6):218:1–13, 2017.

[4] M. A. T. Figueiredo, R. D. Nowak, and S. J. Wright. Gradient projection for sparse reconstruction: Application to compressed sensing and other inverse problems. *IEEE Journal of Selected Topics in Signal Processing*, 1(4):586–597, Dec 2007.

[5] J. Tan, Y. Ma, H. Rueda, D. Baron, and G. R. Arce. Compressive hyperspectral imaging via approximate message passing. *IEEE Journal of Selected Topics in Signal Processing*, 10(2):389–401, March 2016.

[6] J. Xie, L. Xu, and E. Chen. Image denoising and inpainting with deep neural networks. In F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems 25*, pages 341–349. Curran Associates, Inc., 2012.

[7] F. Yasuma, T. Mitsunaga, D. Iso, and S. K. Nayar. Generalized assorted pixel camera: Postcapture control of resolution, dynamic range, and spectrum. *IEEE Transactions on Image Processing*, 19(9):2241–2253, Sept 2010.