PANSHARPENING VIA COUPLED TRIPLE FACTORIZATION DICTIONARY LEARNING

Erik Skau* *Brendt Wohlberg*† *Hamid Krim** *Liyi Dai* *

* North Carolina State University, Raleigh, NC 27695, USA
† Los Alamos National Laboratory, Los Alamos, NM 87544 , USA

ABSTRACT

Data fusion is the operation of integrating data from different modalities to construct a single consistent representation. This paper proposes variations of coupled dictionary learning through an additional factorization. One variation of this model is applicable to the pansharpening data fusion problem. Real world pansharpening data was applied to train and test our proposed formulation. The results demonstrate that the data fusion model can successfully be applied to the pansharpening problem.

Index Terms— data fusion, dictionary learning, pansharpening, hyperspectral

1. INTRODUCTION

Data fusion methods integrate different sources of information to construct a single consistent representation. Pansharpening is a data fusion problem which aims to combine high spatial and high spectral resolution data sets. Specifically, given a high resolution panchromatic image, the goal is to sharpen corresponding low resolution multispectral imagery to obtain a high resolution multispectral image.

One approach to solving the pansharpening problem is dictionary learning. After training dictionaries to represent the different bands of a multispectral image, the trained model can be used to sharpen the low resolution images. The proposed dictionary learning model for data fusion can be applied to the pansharpening problem. Training and validating the model on real world data demonstrates the data fusion formulation's viability for pansharpening.

2. BACKGROUND

Dictionary learning is a method that finds a set of vectors with the goal of sparsely representing a family of signals. Typically, one wishes to represent a set of signals S, with a dictionary D, whose columns are called atoms, and sparse assignment vectors X. A classic approach to this is the minimization problem

$$\begin{aligned} \underset{D,X}{\arg\min} \quad & \frac{1}{2}\|DX - S\|_F^2 + \lambda\|X\|_1, \\ \text{subject to} \quad & \|D\|_{2,\infty} \leq 1, \end{aligned}$$

where λ is referred to as the regularization parameter that controls sparsity, and the constraint on D keeps the dictionary bounded [1]. More precisely the constraint

$$\|D\|_{2,\infty} = \max_i \sqrt{\sum_j |D_{j,i}|^2} \leq 1,$$

keeps the euclidean norm of each atom bounded above by 1. After learning a dictionary, one can use that dictionary to sparsely represent signals since $DX \approx S$. Dictionary learning has had numerous uses, of which two notable applications are image denoising and inpainting.

When dictionary learning is applied to images, the signals are constructed by vectorizing overlapping patches or subimages of a fixed size. Several preprocessing techniques can be applied to the vectorized patches such as centering, normalizing, and whitening [2]. If one wishes to reconstruct the image after sparsely encoding the overlapping patches it is typical to resolve the multivalued pixels by averaging.

Coupled dictionary learning trains two different dictionaries to represent two different but related signals. A typical formulation for coupled dictionary learning is

$$\begin{aligned} \underset{D_i,X_i,W}{\arg\min} \quad & \frac{\sigma_1}{2}\|D_1X_1 - S_1\|_F^2 + \frac{\sigma_2}{2}\|D_2X_2 - S_2\|_F^2 + \\ & \frac{\gamma}{2}\|WX_1 - X_2\|_F^2 + \lambda\|X\|_1, \\ \text{subject to} \quad & \|D_i\|_{2,\infty} \leq 1 \text{ for } i = 1, 2. \end{aligned}$$

Coupled dictionary learning has had significant success in image superresolution as well as image analogies [3, 4].

3. RELATED WORKS AND OUR CONTRIBUTION

There are a variety of approaches to the pansharpening problem [5, 6]. Two of the more basic approaches to pansharpening that we will compare against are Gram-Schmidt (GS) [7] and principal component analysis (PCA) [8]. GS and PCA

Thanks to the U.S. Army Research Office, Grant W911NF-04-D-0003-0022 and Department of Energy Grant DE-NA0002576 for partial funding of this research.

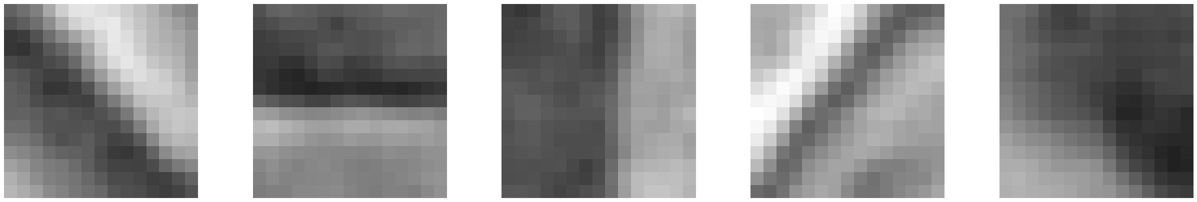


Fig. 1: Sample dictionary atoms transformed to the reference image space.

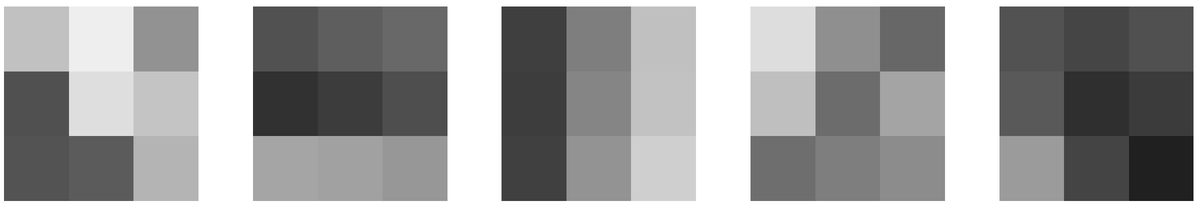


Fig. 2: Sample dictionary atoms transformed to the degraded input image space.

provide a good benchmark for validation of problem formulation. There have also been a few previous works using variations of dictionary learning or sparse coding for pansharpening [9, 10]. Our investigation entails adding an additional factorization to coupled dictionary learning to solve the pansharpening problem.

We propose a different and novel variation of coupled dictionary learning with the goal of demonstrating that an additional factorization provides some new flexibility that can be suitable for pansharpening. Variations of our formulation should be applicable to a variety of problems. The goal of this paper is to demonstrate that this formulation is applicable to the pansharpening problem.

4. PROBLEM FORMULATION

One may observe that in coupled dictionary learning there is often very similar information stored in the different dictionaries. This is especially true when dealing with multispectral images, as there is a lot of overlapping information in different spectral bands. Since there is a lot of redundant information in the different bands, it is natural to construct a model to exploit this. This is the motivation for our proposed method, to find one dictionary that contains information from all bands, as well as a set of linear transformations taking the cumulative information to a specific band. Specifically, our formulation for the pansharpening problem

$$\begin{aligned} &\underset{P_i,D,X}{\arg\min} \quad \sum_i \frac{\sigma_i}{2}\|P_iDX - S_i\|_F^2 + \sum_i \lambda_i \|P_i\|_1 + \lambda\|X\|_1, \\ &\text{subject to} \quad \|D\|_{2,\infty} \leq 1, \end{aligned}$$

trains a dictionary containing the information from every band D, as well as a set of sparse linear transformations P_i's from the general dictionary to the space of a particular band. The sparsity regularization on the linear transformations is appropriate when dealing with images because the value of any given pixel is most strongly correlated with its neighbors. One can imagine other applications where different regularization may be more appropriate.

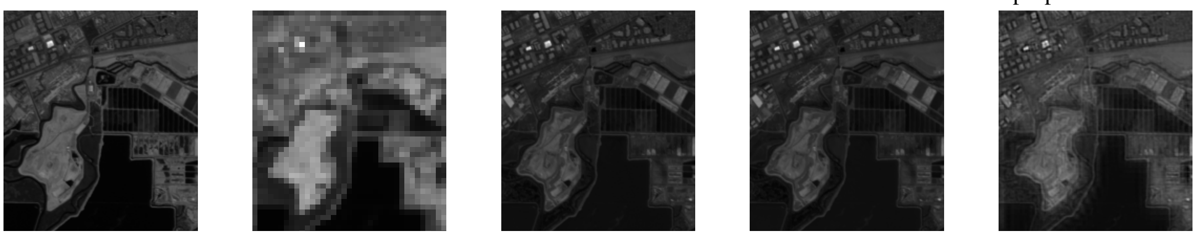
One potential advantage of this formulation can be seen by considering the dimensionality of our dictionary. In classic dictionary learning, each atom needs to be the dimensionality of the signal. In coupled dictionary learning there are multiple signals, each signal with its own dictionary of the corresponding dimensionality. In our proposed formulation there is a single dictionary whose dimensionality is adjustable, so it can be chosen appropriately to be able to hold the cumulative information from each signal. In addition to the single dictionary in our method, there are several sparse linear transformation matrices that map the generic dictionary to the space of a specific signal.

5. ALGORITHM

To perform the optimization, we used a classic alternating minimization algorithm consisting of fixing all but one variable at a time. For each P_i update and X update we utilized alternating direction method of multipliers (ADMM) [11], and a variation of method of optimal directions (MOD) was used for the D update [12].

$$\begin{aligned} X^* &= \underset{X}{\arg\min} \quad \sum_i \frac{\sigma_i}{2}\|P_iDX - S_i\|_F^2 + \lambda\|X\|_1 \\ D^* &= \underset{D}{\arg\min} \quad \sum_i \frac{\sigma_i}{2}\|P_iDX - S_i\|_F^2 \\ &\qquad \text{subject to} \quad \|D\|_{2,\infty} \leq 1 \\ P_i^* &= \underset{P_i}{\arg\min} \quad \frac{\sigma_i}{2}\|P_iDX - S_i\|_F^2 + \lambda_i\|P_i\|_1 \end{aligned}$$

(a) Reference image of a red band (b) Input image of a red band (c) Pansharpened image of a red band via GS (d) Pansharpened image of a red band via PCA (e) Pansharpened image of a red band via our proposed method



(f) Reference image of a green band (g) Input image of a green band (h) Pansharpened image of a green band via GS (i) Pansharpened image of a green band via PCA (j) Pansharpened image of a green band via our proposed method

Fig. 3: Some sample reference and degraded images to be sharpened from Moffett data set

6. TRAINING PROCEDURE

To train our model we require both reference images for each band as well as the degraded input images that are to be processed during validation. For each band in the data set, our model trains one linear transformation to the reference images, and one linear transformation to the input images. Along with the pairs of reference and input images, the single high resolution panchromatic image is also used during training. One half of the data was used for training the model, the other half was used as a validation set.

7. VALIDATION PROCEDURE

To test the trained model we applied it to the validation half of the data set. To compare our method to others we used three methods to measure the quality of the sharpened image. Cross correlation (CC) measures the spatial similarity of the sharpened image and the reference image. Spectral angle mapper (SAM) measures the spectral similarity between hyperspectral images. Root mean squared error (RMSE) measures the ℓ_2 distance between the reference and sharpened image. Numerous pansharpening methods involve interpolating a low resolution image to a high resolution and use the interpolated image as an approximation in which to fill in details. Our method is also suitable for this, but for clarity we skipped this to demonstrate the flexibility of our proposed technique. For a comparative study, we referred to two of the more basic pansharpening techniques, GS and PCA.

8. MOFFETT DATASET RESULTS

The publicly available Moffett data set is a hyperspectral image of 224 bands that contains both urban and vegetative regions taken by Airborne Visible Infrared Imaging Spectrometer (AVIRIS). Our method was compared using the framework developed in [6]. In summary, the reference image was degraded using Wald's protocol by a factor of 5 to construct the low resolution image to be sharpened [13]. The patches for the high resolution and low resolution images were chosen to be 15 by 15 and 3 by 3 respectively. The dictionary was chosen to be 225 by 300. The only image preprocessing steps taken were centering of the signal. Samples of the trained dictionary atoms transformed to some bands are shown in figure's 1 and 2. Applying GS, PCA, and our proposed method to this data set we obtain the numeric results shown in the table.

	GS	PCA	Proposed Method
CC	0.917	0.906	**0.935**
SAM	12.95	13.4	**10.6**
RMSE	420.5	445.1	**364.6**

Figure 3 depicts the reference images, input images, and results from the different methods. Visually one can see that the proposed method provides the best fidelity to the relative intensities of the different bands. Unfortunately our method also shows evidence of blocking artifacts, most notable around the high contrast areas at the bottom. These artifacts are a known disadvantage to dictionary learning techniques that originate from centering the data. The artifacts cause the detail in the GS and PCA methods to be visually more precise. Numerically from the three measures of image quality, our method is superior to both GS and PCA as shown in the table.

9. CONCLUSIONS

This work demonstrates that a triple factorization in coupled dictionary learning achieves better results than some of the more basic pansharpening techniques. Other variations of the proposed data fusion model are suitable to different problems and will be the subject of future work. One future avenue to explore is different methods of regularization for the linear transformations to each signal space.

10. REFERENCES

[1] M. Aharon, M. Elad, and A. Bruckstein, "k -svd: An algorithm for designing overcomplete dictionaries for sparse representation," *Signal Processing, IEEE Transactions on*, vol. 54, no. 11, pp. 4311–4322, Nov 2006.

[2] A. Hyvrinen, J. Hurri, and P. Hoyer, *Natural Image Statistics: A Probabilistic Approach to Early Computational Vision.*, Springer Publishing Company, Incorporated, 1st edition, 2009.

[3] T. Cao, V. Jojic, S. Modla, D. Powell, K. Czymmek, and M. Niethammer, "Robust multimodal dictionary learning," in *Medical Image Computing and Computer-Assisted Intervention MICCAI 2013*, Kensaku Mori, Ichiro Sakuma, Yoshinobu Sato, Christian Barillot, and Nassir Navab, Eds., vol. 8149 of *Lecture Notes in Computer Science*, pp. 259–266. Springer Berlin Heidelberg, 2013.

[4] S. Wang, L. Zhang, Y. Liang, and Q. Pan, "Semi-coupled dictionary learning with applications to image super-resolution and photo-sketch synthesis," in *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, June 2012, pp. 2216–2223.

[5] L. Alparone, L. Wald, J. Chanussot, C. Thomas, P. Gamba, and L. Bruce, "Comparison of pansharpening algorithms: Outcome of the 2006 grs-s data-fusion contest," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 45, no. 10, pp. 3012–3021, Oct 2007.

[6] L. Loncan, L. Almeida, J. Bioucas-Dias, X. Briottet, J. Chanussot, N. Dobigeon, S. Fabre, W. Liao, G. Licciardi, M. Simões, J-Y. Tourneret, M. Veganzones, G. Vivone, Q. Wei, and N. Yokoya, "Hyperspectral pansharpening: a review," *IEEE Geoscience and Remote Sensing Magazine*, p. Accepted, 2015.

[7] C. Laben and B. Brower, "Process for enhancing the spatial resolution of multispectral imagery using pansharpening," Jan. 4 2000, US Patent 6,011,875.

[8] P. Chavez and A. Kwarteng, "Extracting spectral contrast in landsat thematic mapper image data using selective principal component analysis," 1989.

[9] D. Liu and P. Boufounos, "Dictionary learning based pan-sharpening," in *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, March 2012, pp. 2397–2400.

[10] S. Chen, L. Xiao, Z. Wei, and W. Huang, "Pansharpening via coupled unitary dictionary learning," in *Image and Graphics*, Yu-Jin Zhang, Ed., vol. 9219 of *Lecture Notes in Computer Science*, pp. 1–10. Springer International Publishing, 2015.

[11] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, Jan. 2011.

[12] K. Engan, S. Aase, and J. Husøy, "Method of optimal directions for frame design," in *Acoustics, Speech, and Signal Processing, 1999. Proceedings., 1999 IEEE International Conference on*, 1999, vol. 5, pp. 2443–2446 vol.5.

[13] L. Wald, "Quality of high resolution synthesised images: Is there a simple criterion?," in *Third conference "Fusion of Earth data: merging point measurements, raster maps and remotely sensed images"*, Thierry Ranchin and Lucien Wald, Eds., Sophia Antipolis, France, Jan. 2000, pp. 99–103, SEE/URISCA.