Bayesian Subpixel Mapping Neural Network for Hyperspectral images

Yuan Fang, University of Waterloo
Linlin Xu, University of Waterloo
Yuxian Wang, China Univerisity of Geosciences
David A. Clausi, University of Waterloo
Email: y227fang@uwaterloo.ca

Abstract

Subpixel mapping (SPM) of a hyperspectral image (HSI) allocates land cover fractions or discrete abundances in original pixels, so that the resolution of the HSI label map becomes finer by dividing the mixed pixel to subpixels. Most of existing SPM approaches have the limitation of unlearnable spatial prior in HSIs and nonintegrated frameworks. In this paper, we present an unsupervised Bayesian subpixel mapping network for hyperspectral images. An end-to-end unified SPM network with an encoder-decoder architecture is designed to incorporate the fully convolutional neural network (FCNN) with the deep image prior and the forward models to effectively estimate the subpixel labels. The proposed approach is tested on a benchmark real HSI dataset, in comparison with several other SPM methods. The results demonstrate that the proposed method is more effective for SPM of HSIs with higher numerical accuracies and more accurate visual maps of subpixel labels.

1 Introduction

Hyperspectral imaging is a rapidly growing remote sensing technique and has been used widely used for applications, such as ground target classification [1–3], agricultural management [4] and environmental monitoring [5, 6]. Hyperspectral images (HSIs) have hundreds of narrow and contiguous spectral bands, recording the electromagnetic radiation from the earth surface and the atmosphere [7]. However, due to the trade-off between the spectral resolution and spatial resolution in HSIs, pixels in HSIs usually contain spectral contributions from multiple materials. Spectral unmixing (SU) decomposes these mixed pixels into both the spectral signatures of constituent components (i.e., endmembers) and their corresponding fractional proportions (i.e., abundances) from the mixed pixel in HSI [8]. However, SU processes HSIs in a manner of soft classification at the original observation scale, without knowing the spatial distribution of the endmembers inside the mixed pixel [9].

Subpixel mapping (SPM) assigns abundances to HSI pixels to create a finer label map by dividing the mixed pixel to subpixels [10]. Efficient SPM approaches are crucial due to the requirements of wide-range applications, such as environmental monitoring [11–13], target detection [14], and rural land cover objects [15]. An unsupervised Bayesian SPM network (BSMN) for HSI which integrates the forward downsampling model with a fully convolutional neural network (FCNN) in a Bayesian framework is designed and implemented.

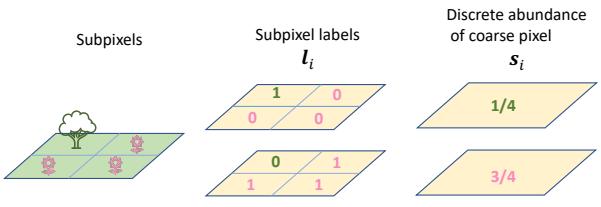

Fig. 1: Illustration of the relationship between the subpixel labels and discrete abundances coarse pixel.

2 Problem Formulation

We assume that an observed HSI data cube $\boldsymbol{X}$ has P spectral bands, N pixels containing m rows and n columns, and the term I represents the set of coarse pixel sites in HSI, we denote the observed reflectance of the pixel at site i by $\boldsymbol{x}_i$, which is a $P \times 1$ vector. Then the HSI can be expressed as $\boldsymbol{X} = \{\boldsymbol{x}_i | i = 1, 2, ..., m \times n\}$. The term J represents the set of subpixel positions within each coarse pixel, which contains a total of c^2 positions. For example, when $c = 2$, a coarse pixel is divided into 2×2 subpixels, and the coarse HSI containing $m \times n$ pixels corresponds to a fine HSI with the size of $2m \times 2n$ subpixels. Assuming that the HSI covers K classes, the SPM aims to infer subpixel labels $\boldsymbol{L} = \{\boldsymbol{l}_{i,j} | i \in I, j \in J\}$ in the HSI, where $\boldsymbol{l}_{i,j}$ is a one-hot vector $K \times 1$ vector. Then, a coarse pixel $\boldsymbol{x}_i$ can be formulated a linear combination of K endmembers $\boldsymbol{A} = \{\boldsymbol{a}_k | k = 1, 2, ..., K\}$ weighted by the discrete abundances plus noise $\boldsymbol{n}_i$:

$$\boldsymbol{x}_i = \sum_{k=1}^{K} \boldsymbol{a}_k \frac{\sum_{j=1}^{c^2} \delta\left(\boldsymbol{l}_{i,j}, \boldsymbol{k}\right)}{c^2} + \boldsymbol{n}_i, \tag{1}$$

The noise distribution is assumed to satisfy a Gaussian model. The discrete abundance $\boldsymbol{s}_i$ of a coarse pixel is determined by the proportions of subpixels labels. "$\boldsymbol{k}$" is the one-hot representation of "k". $\delta(\boldsymbol{u}, \boldsymbol{v})$ is the Kronecker delta function where $\delta(\boldsymbol{u}, \boldsymbol{v}) = 1$ for $\boldsymbol{u} = \boldsymbol{v}$ and $\delta(\boldsymbol{u}, \boldsymbol{v}) = 0$ otherwise. For example, when $c = 2$, if the subpixel labels indicate one subpixel in the class "tree" and three subpixels in the class "flower" among the 2×2 subpixels, the discrete abundance $\boldsymbol{s}_i$ corresponding to the coarse pixel is a 2×1 vector written as $[1/4; 3/4]$, as illustrated in Figure 1.

MAP estimation. The SPM problem can be solved by the MAP approach by maximizing the posterior distribution of $\boldsymbol{L}$ given the observed HSI $\boldsymbol{X}$ and the model parameters (i.e., endmembers $\boldsymbol{A}$). Maximizing $p(\boldsymbol{L} \mid \boldsymbol{X}, \boldsymbol{A})$ is equivalent to minimizing its negative logarithm likelihood. The objective function can be reformulated as follows [16],

$$\arg\min_{\mathbf{L}} \sum_{i=1}^{N} \left\{ \left\| \boldsymbol{x}_i - \left(\sum_{k=1}^{K} \boldsymbol{a}_k \frac{\sum_{j=1}^{c^2} \delta\left(E(\boldsymbol{l}_{i,j}), \boldsymbol{k}\right)}{c^2} \right) \right\|^2 \right\} \tag{2}$$

This objective function has following characteristics: (i) the EM algorithm estimates parameters by treating $\{\boldsymbol{l}_{i,j}\}$ as missing observations and $\{\boldsymbol{a}_k\}$ as model parameters, and iteratively updates the estimation of $\{\boldsymbol{l}_{i,j}\}$ and $\{\boldsymbol{a}_k\}$. (ii) We use $E(\boldsymbol{l}_{i,j})$ as the estimation of $\boldsymbol{l}_{i,j}$. we use $\{\boldsymbol{x}_i\}$ as input to the FCNN and optimize network parameters. Once the FCNN is trained, we obtain $\hat{\boldsymbol{l}}_{i,j} = f(\boldsymbol{l}_{i,j})$. (iii) When estimating parameters in the FCNN for obtaining $E(\boldsymbol{l}_{i,j})$, we use a reconstruction loss based on $||\boldsymbol{x}_i - \hat{\boldsymbol{x}}_i||^2$, which incorporates a forward model to constrain meaningful $\boldsymbol{l}_{i,j}$ estimation.

EM Iteration. The main steps in EM algorithm to estimate $\{\boldsymbol{l}_{i,j}\}$ and $\{\boldsymbol{a}_k\}$ are summarized as follows.

Initialization: Set the initial value for $\{\boldsymbol{a}_k\}$. The endmember of each class is manually selected from the coarse HSI.

E-step: Given endmembers $\{\boldsymbol{a}_k\}$, estimate subpixel labels $\{\boldsymbol{l}_{i,j}\}$ by optimizing a FCNN.

M-step: Given $\{\boldsymbol{l}_{i,j}\}$, estimate endmembers $\{\boldsymbol{a}_k\}$. Endmembers $\{\boldsymbol{a}_k\}$ are estimated with the purified means approach [17].

3 Experiment

We use the coarse HSI as the input of the FCNN, and the output is the soft subpixel labels. The forward model is implemented with an average pooling downsampler to map soft subpixel labels to discrete abundance of coarse pixels, and a linear combination to reconstruct the HSI using discrete abundances and endmembers. The FCNN is implemented with a U-Net type "hourglass" architecture with skip-connection [16] to model a mapping from the input to soft labels of subpixels.

Figure 3 shows the SPM results for the Jasper HSI scene obtained by different SPM methods. The subpixel label map obtained by the BSMN methods shows the most detailed spatial textural information and achieves the highest numerical accuracy. Table 1 shows the individual classification accuracy achieved by different methods on the HSI. The proposed BSMN performs the best on the class "water", "tree", and "soil". Although the "road" class accuracy is relatively low, the outline of the road is still clearly visible in Figure 3. We attribute the false negative pixels for road class to the smoothness property of convolutions.

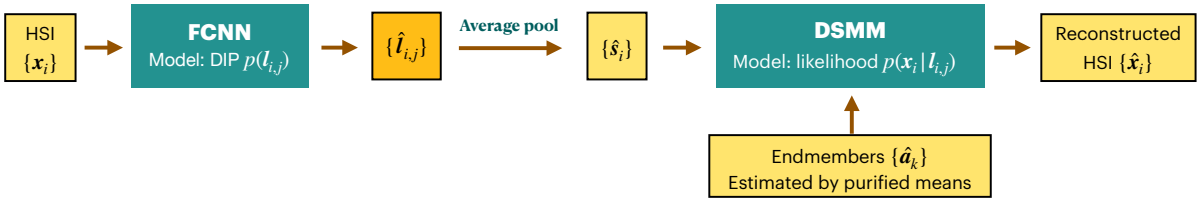

Fig. 2: Subpixel mapping framework. The encoder is a skip-connection FCNN designed for the estimation of soft labels of subpixels $\{l_{i,j}\}$, where DIP is used to model the spatial correlation of the label field. The decoder has two parts. One part is the forward downsampling model which maps soft labels of subpixels to the class proportions s_i. The other part of the encoder reconstructs the HSI $\hat{X}$ with s_i and endmembers A extracted from the HSI X.

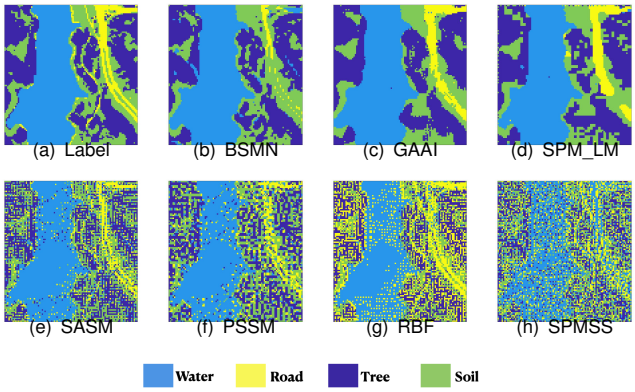

Fig. 3: Subpixel mapping results of different methods on Japser dataset.

Table 1: Individual and overall class accuracies (unit: %).

Methods	Water	Road	Tree	Soil	OA
SPMSS[18]	64.94	41.43	45.78	39.13	50.21
RBF[19]	88.45	**71.71**	0.5560	48.72	66.22
PSSM[20]	91.28	53.25	58.69	54.04	67.99
SASM[21]	92.09	59.36	60.38	57.87	70.24
SPM_LM[22]	99.91	65.74	84.88	73.15	85.59
GAAI[23]	99.97	64.28	87.12	71.21	85.81
BSMN	**100.00**	23.24	**90.58**	**83.03**	**86.81**

4 Conclusion

We presented an unsupervised Bayesian subpixel mapping network for HSIs. An encoder-decoder architecture was designed to incorporate the FCNN with DIP prior and the forward models to effectively estimate the subpixel labels. BSMN adopted FCNN rather than fully connected layers to better exploit the spatial correlation effect in HSI. An efficient purified means approach was adopted to the SPM framework for the endmember estimation. The resulting Bayesian MAP framework is solved by the proposed EM approach. The proposed approach was tested on both real and simulated HSIs, in comparison with several other SPM methods. The results demonstrated that the proposed BSMN method was more effective for SPM of HSIs with higher numerical accuracies and more accurate visual maps of subpixel labels.

References

[1] M. Fauvel, Y. Tarabalka, J. A. Benediktsson, J. Chanussot, and J. C. Tilton, "Advances in spectral-spatial classification of hyperspectral images," *Proceedings of the IEEE*, vol. 101, no. 3, pp. 652–675, 2012.

[2] X. Sun, Q. Qu, N. M. Nasrabadi, and T. D. Tran, "Structured priors for sparse-representation-based hyperspectral image classification," *IEEE geoscience and remote sensing letters*, vol. 11, no. 7, pp. 1235–1239, 2013.

[3] Y. Fang, L. Xu, J. Peng, H. Yang, A. Wong, and D. A. Clausi, "Unsupervised bayesian classification of a hyperspectral image based on the spectral mixture model and markov random field," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 9, pp. 3325–3337, 2018.

[4] T. Adão, J. Hruška, L. Pádua, J. Bessa, E. Peres, R. Morais, and J. Sousa, "Hyperspectral imaging: A review on uav-based sensors, data processing and applications for agriculture and forestry," *Remote Sensing*, vol. 9, no. 11, p. 1110, 2017.

[5] A. Plaza, Q. Du, J. M. Bioucas-Dias, X. Jia, and F. A. Kruse, "Foreword to the special issue on spectral unmixing of remotely sensed data," *IEEE transactions on geoscience and remote sensing*, vol. 49, no. 11, pp. 4103–4110, 2011.

[6] R. Jackisch, S. Lorenz, R. Zimmermann, R. Möckel, and R. Gloaguen, "Drone-borne hyperspectral monitoring of acid mine drainage: An example from the sokolov lignite district," *Remote Sensing*, vol. 10, no. 3, p. 385, 2018.

[7] B. Palsson, J. Sigurdsson, J. R. Sveinsson, and M. O. Ulfarsson, "Hyperspectral unmixing using a neural network autoencoder," *IEEE Access*, vol. 6, pp. 25 646–25 656, 2018.

[8] J. M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE journal of selected topics in applied earth observations and remote sensing*, vol. 5, no. 2, pp. 354–379, 2012.

[9] D. He, Y. Zhong, X. Wang, and L. Zhang, "Deep convolutional neural network framework for subpixel mapping," *IEEE Transactions on Geoscience and Remote Sensing*, 2020.

[10] Mertens, KC, De, Baets, B, Verbeke, LPC, De, Wulf, and RR, "A sub-pixel mapping algorithm based on sub-pixel/pixel spatial attraction models," *Int J Remote Sens*, 2006.

[11] Dagrun, Vikhamar, , , Rune, and Solberg, "Subpixel mapping of snow cover in forests by optical remote sensing," *Remote Sensing of Environment*, 2003.

[12] Y. Lv, W. Gao, C. Yang, and Z. Fang, "A novel spatial–spectral extraction method for subpixel surface water," *International Journal of Remote Sensing*.

[13] F. Ling, Y. Du, Y. Zhang, X. Li, and F. Xiao, "Burned-area mapping at the subpixel scale with modis images," *IEEE Geoscience and Remote Sensing Letters*, vol. 12, no. 9, pp. 1963–1967, 2015.

[14] L. Zhang, K. Wu, Y. Zhong, and P. Li, "A new sub-pixel mapping algorithm based on a bp neural network with an observation model," *Neurocomputing*, vol. 71, no. 10-12, pp. 2046–2054, 2008.

[15] M. W. Thornton, P. M. Atkinson, and D. A. Holland, "A linearised pixel-swapping method for mapping rural linear land cover features from fine spatial resolution remotely sensed imagery," *Computers geosciences*, vol. 33, no. 10, pp. 1261–1272, 2007.

[16] D. Ulyanov, A. Vedaldi, and V. Lempitsky, "Deep image prior," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 9446–9454.

[17] L. Xu, J. Li, A. Wong, and J. Peng, “Kp-means: A clustering algorithm of k “purified” means for hyperspectral endmember estimation,” *IEEE Geoscience and Remote Sensing Letters*, vol. 11, no. 10, pp. 1787–1791, 2014.

[18] P. Wang, L. Wang, and J. Chanussot, “Soft-then-hard subpixel land cover mapping based on spatial-spectral interpolation,” *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 12, pp. 1851–1854, 2016.

[19] Q. Wang, W. Shi, and P. M. Atkinson, “Sub-pixel mapping of remote sensing images based on radial basis function interpolation,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 92, pp. 1–15, 2014.

[20] P. M. Atkinson, “Sub-pixel target mapping from soft-classified, remotely sensed imagery,” *Photogrammetric Engineering & Remote Sensing*, vol. 71, no. 7, pp. 839–846, 2005.

[21] K. C. Mertens, B. De Baets, L. P. Verbeke, and R. R. De Wulf, “A sub-pixel mapping algorithm based on sub-pixel/pixel spatial attraction models,” *International Journal of Remote Sensing*, vol. 27, no. 15, pp. 3293–3310, 2006.

[22] F. Ling, Y. Du, F. Xiao, and X. Li, “Subpixel land cover mapping by integrating spectral and spatial information of remotely sensed imagery,” *IEEE Geoscience and Remote Sensing Letters*, vol. 9, no. 3, pp. 408–412, 2011.

[23] X. Tong, X. Xu, A. Plaza, H. Xie, H. Pan, W. Cao, and D. Lv, “A new genetic method for subpixel mapping using hyperspectral images,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 9, no. 9, pp. 4480–4491, 2016.