This item is the archived peer-reviewed author-version of:

MRF-based decision fusion for hyperspectral image classification

Reference:

Andrejchenko Vera, Heylen Rob, Liao Wenzhi, Philips Wilfried, Scheunders Paul.- MRF-based decision fusion for hyperspectral image classification
IEEE International Geoscience and Remote Sensing Symposium - ISSN 2153-7003 - (2017), p. 8066-8069
Full text (Publisher's DOI): <https://doi.org/10.1109/IGARSS.2018.8519182>

MRF-BASED DECISION FUSION FOR HYPERSPECTRAL IMAGE CLASSIFICATION.

Vera Andrejchenko ^a, Rob Heylen ^a, Wenzhi Liao ^b, Wilfried Philips ^b, Paul Scheunders ^a

Visionlab, University of Antwerp ^a

Image Processing and Interpretation, University of Gent ^b

ABSTRACT

The high dimensionality of hyperspectral images, the limited availability of ground-truth data as well as the limited spatial resolution, causing pixels to contain mixtures of materials, hinder hyperspectral image classification. In this work we propose to combine the outcome of spectral unmixing with the outcome of a supervised classifier to improve classification. In particular, we consider fractional abundances obtained from a Sparse Unmixing method along with posterior probabilities obtained from a Multinomial Logistic Regression classifier. Both sources of information are fused using a Markov Random Field framework. We conducted experiments on publicly available real hyperspectral images: Indian Pines and University of Pavia using a very limited number of training samples. Our results indicate that the proposed decision fusion approach largely improves the classification result over using the individual sources and outperforms the state of the art methods.

Index Terms— unmixing, supervised classification, MRF decision fusion

1. INTRODUCTION

Classification of hyperspectral images is challenging due to the high dimensionality of the data, the limited availability of labeled training data, and the limited spatial resolution. The latter causes hyperspectral pixels to contain mixtures of different materials on the ground. In order to address these issues, supervised classification methods that consider pixelwise information based on e.g. support vector machines (SVM) [1] or multinomial logistic regression (MLR) [2] can be used. On the other hand, spectral unmixing methods [3] try to determine the pure materials and infer the fractional abundances of these materials within the pixel. Recently, unmixing methods were applied for classification [4].

However, the performance of these approaches relies solely on pixelwise information. This has fostered the need

to develop spectral-spatial classification techniques [5]. Another strategy is to combine multiple classification outcomes using decision fusion methods. Several multi-source and multi-feature fusion techniques based on Markov Random Fields were proposed [6], [7], [8], [9], [10] [11] that account for spatial neighborhood information.

A limited amount of studies has considered the combination of classification and unmixing methods. Recently, in [12], a framework was proposed to fuse spectral unmixing and classification results based on: a) a certainty degree of the class probabilities and fractional abundances, b) confidence scores extracted from the unmixing (as a classifier) and a SVM. These certainty degrees and confidence scores together defined joint class probabilities which are further refined with a MRF to carry out the final classification.

In this paper we propose a new decision fusion approach: MRF with cross links (MRFL), combining abundance and probability information modeled as a single MRF graph. The fractional abundances are obtained from a sparse unmixing method [3], and the posterior probabilities from a supervised MLR classifier [13]. Both are then combined using an interaction link in the MRF graph. With this strategy, we exploit the complementarity of both decision sources to improve the classification accuracies.

2. METHOD

In this section, we propose an approach for fusion of fractional abundances produced by a spectral unmixing method and probability outputs from a supervised classifier. The ultimate goal is to improve the overall accuracy by including complementary information from both decision sources in the case of low number of training pixels. Let $\mathbf{x} = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ be an image containing n pixels as reflectances, where $\mathbf{x}_i \in \mathbb{R}^d$, d being the number of spectral bands. $D = \{(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_m, y_m)\}$ is a training set containing $j = 1, \dots, m$ labeled samples $\mathbf{x}_j$ and their associated labels $y_j \in \{1, \dots, C\}$ where C is the number of classes. The aim is to assign labels y_i to each image pixel $\mathbf{x}_i$.

As a first step, we obtain class probabilities for pixel $\mathbf{x}_i$

using a Multinomial Logistic Regression classifier (MLR) [13]: $\mathbf{p}(\mathbf{x}_i) = (p_1(\mathbf{x}_i), \dots, p_C(\mathbf{x}_i))$, with:

$$p_c(\mathbf{x}_i) = p(y_i = c | \mathbf{x}_i) = \frac{\exp(\beta_c^T \mathbf{x}_i)}{\sum_{c=1}^C \exp(\beta_c^T \mathbf{x}_i)} \quad (1)$$

where $\beta_c \in R^d$, ($c = 1, \dots, C$) are the regression coefficients, estimated from the training data.

Furthermore, we compute the fractional abundances of pixel $\mathbf{x}_i$ with the SunSAL unmixing method [3], with the training data being used as a dictionary of endmembers, $\mathbf{E} = [\mathbf{x}_1, \dots, \mathbf{x}_m]$:

$$\begin{aligned} \boldsymbol{\alpha}^* &= (\alpha_1^*, \dots, \alpha_m^*) \\ &= \arg \min_{\boldsymbol{\alpha}} \frac{1}{2} \|\mathbf{E}\boldsymbol{\alpha} - \mathbf{x}_i\|_2^2 + \lambda \|\boldsymbol{\alpha}\|_1, \\ &\text{s.t. } \boldsymbol{\alpha} \geq 0. \end{aligned} \quad (2)$$

Then, the obtained abundances for all endmembers belonging to the same class c ($\mathbf{x}_j$ with $y_j = c$) are summed up to obtain one fractional abundance $\alpha_c(\mathbf{x}_i)$ for each class c , and the abundance vector: $\boldsymbol{\alpha}(\mathbf{x}_i) = (\alpha_1(\mathbf{x}_i), \dots, \alpha_C(\mathbf{x}_i))$.

Once the individual soft decisions $\boldsymbol{\alpha}$ and $\mathbf{p}$ are obtained from the unmixing and the MLR classifier, the fusion of these modalities is described in terms of graphs using a MRF. In the classical MRF approach, a graph is defined over a set of observed pixels $\mathbf{x} = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ and their corresponding class labels $\mathbf{y} = \{y_1, \dots, y_n\}$, being the nodes in the graph, such that the graph edges model the spatial neighborhood dependencies between the pixels. While the pixel values are known, the labels are the variables which have to be estimated. In order to accomplish that, the probability distribution of the labels, given the observed data: $P(\mathbf{y}|\mathbf{x})$ is then maximized. In terms of energies, the optimal labels are inferred by minimizing the following energy function:

$$E(\mathbf{y}) = \sum_{i=1}^n \psi_i(y_i) + \beta \sum_{i=1, j \in N_i} \psi_{i,j}(y_i, y_j) \quad (3)$$

where $\psi_i(y_i)$ is the unary potential, defined on the single observation $\mathbf{x}_i$ and its label y_i [14] (e.g. given by $-\log p(\mathbf{x}_i | y_i)$), whereas $\psi_{i,j} = (1 - \delta(y_i, y_j))$ are the pairwise potentials which are only label dependent and impose smoothness on the labels within the spatial neighborhood N_i of pixel i . In the above, $\delta(y_i, y_j)$ denotes the indicator function ($\delta(a, b) = 1$ for $a = b$ and $\delta(a, b) = 0$ otherwise).

2.1. MRF with cross links for fusion

In our approach, we consider a slightly different configuration in which the 2 sets of observations: the entire set of derived fractional abundances obtained from the unmixing method: $\{\boldsymbol{\alpha}(\mathbf{x}_1), \dots, \boldsymbol{\alpha}(\mathbf{x}_n)\}$ and the derived classification probabilities obtained by the MLR method: $\{\mathbf{p}(\mathbf{x}_1), \dots, \mathbf{p}(\mathbf{x}_n)\}$, along with their corresponding labels: $\mathbf{y}^\alpha = \{y_1^\alpha, \dots, y_n^\alpha\}$

and $\mathbf{y}^p = \{y_1^p, \dots, y_n^p\}$ are accounted for in the graph. For each pixel i , two nodes are defined: on node containing the fractional abundance vector, and its corresponding label: $(\boldsymbol{\alpha}(\mathbf{x}_i), y_i^\alpha)$, and the other node containing the classification probability vector and its corresponding label: $(\mathbf{p}_i, y_i^p)$. Along with the edges connecting neighboring pixels, an interaction link is defined, joining label y_i^α with the corresponding label y_i^p [6], [7], see Fig. 1.

The goal is now to optimize the joint distribution over all the labels from both sources given their features: $P(\mathbf{y}^\alpha, \mathbf{y}^p | \boldsymbol{\alpha}, \mathbf{p})$. For this, the following energy function is minimized:

$$\begin{aligned} E(\mathbf{y}^\alpha, \mathbf{y}^p) &= \sum_{i=1}^n \psi_i^\alpha(y_i^\alpha) + \sum_{i=1}^n \psi_i^p(y_i^p) \\ &+ \beta \left[\sum_{i=1, j \in N_i} \psi_{i,j}^\alpha(y_i^\alpha, y_j^\alpha) + \sum_{i=1, j \in N_i} \psi_{i,j}^p(y_i^p, y_j^p) \right] \\ &+ \gamma \sum_{i=1}^n \psi_{i,i}^{\alpha p}(y_i^\alpha, y_i^p) \end{aligned} \quad (4)$$

where the unary potentials are defined as:

$\psi_i^\alpha(y_i) = -\log(\alpha_c(\mathbf{x}_i))$, and $\psi_i^p(y_i) = -\log(p_c(\mathbf{x}_i))$ with $y_i = c$. The pairwise potentials from the individual sources: $\psi_{i,j}^\alpha = (1 - \delta(y_i^\alpha, y_j^\alpha))$ and $\psi_{i,j}^p = (1 - \delta(y_i^p, y_j^p))$ impose smoothness on the labeling of the spatial neighborhood of pixel i , obtained from the fractional abundances and the classification probabilities respectively. The last pairwise term $\psi_{i,i}^{\alpha p} = (1 - \delta(y_i^\alpha, y_i^p))$ penalizes dissimilar labeling ($y_i^\alpha \neq y_i^p$) of pixel i . The optimal labeling is computed with the graph-cut α -expansion algorithm [15], [16]. Since the last term $\psi_{i,i}^{\alpha p}$ encourages cross-source label consistency, a vast majority of pixels is expected to have equivalent labels $y_i^\alpha = y_i^p$. For this reason, any of the two may be used as the final labeling result.

Fig. 1: Part of the defined MRF graph. The blue circles are the label nodes from the abundance source and the green squares are the label nodes from the probability source. Dashed lines denote edges between neighboring nodes. The red link is the interaction potential between the central blue and central green label node.

3. EXPERIMENTS AND RESULTS

For our experiments we used the "AVIRIS Indian Pines" and the "ROSIS-03 University of Pavia" hyperspectral images.

3.1. Indian Pines

This scene consist of 145 x 145 pixels with a spatial resolution of 20m and 220 spectral bands, ranging from 0.2 to 2.3 μm . Prior to using the dataset, we manually discarded the noisy bands and the water absorption bands, leaving us with 164 bands. A ground reference map is available, consisting of 16 classes, from which we removed the smallest classes and used the rest for extraction of training and test pixels. The training data was composed of only 10 pixels per class.

The value for the regularization parameter λ in the sparse unmixing problem (2) was empirically selected to be $\lambda = 4 \cdot 10^{-4}$. Afterwards the abundances were normalized. In the proposed MRFL approach, the β parameter which weights the influence of the spatial neighbours (4-neighborhood) was set to $\beta = 0.5$, following the literature [12]. γ which weights the influence of the connection term, promoting equality of corresponding label pairs from the abundance and probability sources, was set to a value: $\gamma = 0.5$.

The performance of our methods was compared with the following methods:

- U, the unmixing method used as a classifier, acting on the abundances as a single source : $\hat{y}_i^\alpha = \arg \max_c \alpha_c(\mathbf{x}_i)$,
- the MLR classifier acting on the probabilities as a single source: $\hat{y}_i^p = \arg \max_c p_c(\mathbf{x}_i)$,
- the linear weighting approach - a simple decision fusion approach using a linear combination of the abundance and probability sources,
- and the MRFG fusion approach [12]. This method is applied with 2 sources of information as input: 1) the probabilities of the MLR classifier and 2) classification probabilities, obtained from applying a SVM classifier on the abundances obtained from sparse unmixing. In the original paper, a third source given by spectral-spatial features was included.

Table 1 shows that the proposed method MRFL outperforms the first two methods U and MLR by $\approx 16\%$ and MLR by $\approx 14\%$ respectively, which classify the pixels without performing any fusion and base their decisions solely on the individual sources. MRFL also outperforms the linear combination of abundances and probabilities by more than $\approx 10\%$ (the results using the original abundances are mentioned in this table, other transformations of these abundances produced similar results) and the MRFG method by more than $\approx 6\%$.

Table 1: Average accuracies [%] (and standard deviations) on the Indian Pines image

U	MLR	LC	MRFG	MRFL
64.86 $\pm(3.4)$	66.87 $\pm(2.7)$	71.01 $\pm(2.28)$	74.3 $\pm(5.2)$	81.19 $\pm(4.2)$

Table 2: Average accuracies [%] (and standard deviations) on the University of Pavia image

U	MLR	LC	MRFG	MRFL
63.04 $\pm(4.06)$	65.72 $\pm(4.1)$	68.6 $\pm(3)$	73.6 $\pm(6)$	80.76 $\pm(3.9)$

3.2. University of Pavia

We performed the same experiments on the University of Pavia hyperspectral image of 610 x 340 pixels, with 115 bands, with a spectral range from 0.43 to 0.86 μm , with a spatial resolution of 1.3 meter per pixel. Twelve noisy bands have been removed, and the remaining 103 spectral channels are used. The ground truth map consists of nine classes: Trees, asphalt, bitumen, gravel, metal sheets, shadow, self-blocking bricks, meadows and bare soil. As in the previous image, only 10 pixels per class were used as training data. The unmixing regularization parameter λ was set, similarly as in the previous experiment to: $\lambda = 4 \cdot 10^{-4}$, the abundances were normalized, the β parameter in the proposed method was set, following the literature [12] to $\beta = 0.5$, weighting the 4-neighborhood and γ was set to $\gamma = 0.5$ for the same reasons as in the Indian Pines image.

Similar results as in the first experiment can be observed in Table 2. Fig. 2 visually illustrates the improvement of our method (e) over the other methods, where it can be seen that the abundance fractions and the probabilities complement each other and combining them produces a classification map much closer to the ground-truth map (f).

3.3. Conclusion

In this paper we proposed a novel framework for decision fusion of obtained fractional abundances and classification probabilities on hyperspectral images in the case of low number of labeled samples. The fusion method combines both sources of information in a single MRF graph. Our experiments showed that the method improves the classification accuracy over state of the art methods.

4. REFERENCES

- [1] G. Camps-Valls and L. Bruzzone, "Kernel-based methods for hyperspectral image classification," *IEEE Trans-*

Fig. 2: Classification maps

actions on Geoscience and Remote Sensing, vol. 43, no. 6, pp. 1351–1362, June 2005.

- [2] J. Li, J. M. Bioucas-Dias, and A. Plaza, “Hyperspectral image segmentation using a new bayesian approach with active learning,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 10, pp. 3947–3960, Oct 2011.
- [3] J.M. Bioucas-Dias and M.A.T. Figueiredo, “Alternating direction algorithms for constrained sparse regression: Application to hyperspectral unmixing,” in *Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, 2010 2nd Workshop on, June 2010, pp. 1–4.
- [4] Scheunders P. Philips W Liao W. Andrejchenko V., Heylen R., “Classification of hyperspectral images with very small training size using sparse unmixing,” in *International Geoscience and Remote Sensing Symposium - IGARSS 2016*, 2016.
- [5] Jocelyn Chanussot Johannes R. Sveinsson Mathieu Fauvel, Jon Atli Benediktsson, “Spectral and spatial classification of hyperspectral data using svms and morphological profiles,” *IEEE Transactions on Geoscience and Remote Sensing, Institute of Electrical and Electronics Engineers*, vol. 46, pp. 3804–3814, 2008.
- [6] S. T. Namin, M. Najafi, M. Salzmann, and L. Petersson, “A multi-modal graphical model for scene analysis,” in *2015 IEEE Winter Conference on Applications of Computer Vision*, Jan 2015, pp. 1006–1013.
- [7] Mikkel Kragh and James Underwood, “Multi-modal obstacle detection in unstructured environments with conditional random fields,” *CoRR*, vol. abs/1706.02908, 2017.
- [8] A. H. S. Solberg, T. Taxt, and A. K. Jain, “A markov random field model for classification of multisource satellite imagery,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 34, no. 1, pp. 100–113, Jan 1996.
- [9] L. Gmez-Chova, D. Tuia, G. Moser, and G. Camps-Valls, “Multimodal classification of remote sensing images: A review and future directions,” *Proceedings of the IEEE*, vol. 103, no. 9, pp. 1560–1584, Sept 2015.
- [10] M. Xu, H. Chen, and P. K. Varshney, “An image fusion approach based on markov random fields,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 12, pp. 5116–5127, Dec 2011.
- [11] Q. Lu, X. Huang, J. Li, and L. Zhang, “A novel mrf-based multifeature fusion for classification of remote sensing images,” *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 4, pp. 515–519, April 2016.
- [12] T. Lu, S. Li, L. Fang, X. Jia, and J. A. Benediktsson, “From subpixel to superpixel: A novel fusion framework for hyperspectral image classification,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 8, pp. 4398–4411, Aug 2017.
- [13] Friedman J. Hastie T., Tibshirani R., *The Elements of Statistical Learning*, Springer, 2008.
- [14] Christopher M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)*, Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2006.
- [15] Shai Bagon, “Matlab wrapper for robust higher order potentials,” January 2009.
- [16] M. Schmidt, “Ugm: A matlab toolbox for probabilistic undirected graphical models,” 2007.