Functions of multiple instances for sub-pixel target characterization in hyperspectral imagery

Alina Zare, Changzhe Jiao

Alina Zare, Changzhe Jiao, "Functions of multiple instances for sub-pixel target characterization in hyperspectral imagery," Proc. SPIE 9472, Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XXI, 947212 (21 May 2015); doi: 10.1117/12.2176889

SPIE.

Event: SPIE Defense + Security, 2015, Baltimore, Maryland, United States

Functions of multiple instances for sub-pixel target characterization in hyperspectral imagery

Alina Zare and Changzhe Jiao

Department of Electrical and Computer Engineering, University of Missouri, Columbia, MO, USA;

ABSTRACT

In this paper, the Multi-target Extended Function of Multiple Instances (Multi-target eFUMI) method is developed and described. The method is capable of learning multiple target spectral signatures from weakly- and inaccurately-labeled hyperspectral imagery. Multi-target eFUMI is a generalization of the Function of Multiple Instances approach (FUMI). The FUMI approach differs significantly from standard Multiple Instance Learning (MIL) approach in that it assumes each data is a function of target and non-target "concepts." In this paper, data points which are convex combinations of multiple target and several non-target "concepts" are considered. Moreover, it allows both "proportion-level" and "bag-level" uncertainties in training data. Training data needs only binary labels indicating whether some spatial area contains or does not contain some proportion of target; the specific target proportions for the training data are not needed. Multi-target eFUMI learns the target and non-target concepts, the number of non-target concepts, and the proportions of all the concepts for each data point. After learning the target concepts using the binary "bag-level" labeled training data, target detection can be performed on test data. Results for sub-pixel target detection on simulated and real airborne hyperspectral data are shown.

Keywords: sub-pixel, Multi-target, target detection, unmixing, endmember, hyperspectral, multiple instance learning

1. INTRODUCTION

Multiple Instance Learning is a variation of supervised learning in which training data is paired with only incomplete training labels. In particular, instead of having instance-level labeled training data, it assumes only bag-level labeled training data is available. Each bag is a collection of many data points or feature vectors that have been labeled into one of two categories: "positive" or "negative." A bag is defined to be positive if at least one of the instances in the bag falls to the target category. A negatively labeled bag must be composed of all non-target instances. The goal of Multi-target eFUMI is to learn all target concepts based on bag-level labeled training data. These concepts can then be used to predict the unknown labels of test instances or bags. MIL was originally proposed to address the application of predicting the drug molecule activities[1]. Since then, many effective methods have been proposed and developed [2–6].

The FUMI approaches [7–10] are a generalization of MIL. FUMI assumes each instance has a latent class label indicating whether each individual instance is or is not a target point. These latent class labels also contribute to the positive or negative bag-level labels. In FUMI, each data point is regarded as a function of positive and negative concepts. Here, concepts refer to the generalized class prototypes in the feature space. In the case of hyperspectral imagery, a FUMI concept is equivalent to the spectral signature of an endmember. Suppose there is a given data set $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_N\}$ and some assumed functional form $f(\cdot)$, then FUMI methods assume that each data point can be represented as $\mathbf{x}_i = f(\mathbf{E}_i, \mathbf{p}_i)$ where $\mathbf{E}_i$ is the "bag" of concepts that defines the latent label of $\mathbf{x}_i$ and $\mathbf{p}_i$ are the set of functional parameters which relate $\mathbf{x}_i$ to $\mathbf{E}_i$. Each training point $\mathbf{x}_i$ is given a binary label $l(\mathbf{x}_i)$ where $l(\mathbf{x}_i) = 1$ if $\mathbf{e}_T \in \mathbf{E}_i$ and $l(\mathbf{x}_i) = 0$ if $\mathbf{e}_T \notin \mathbf{E}_i$ where $\mathbf{e}_T$ is the single target concept.

The cFUMI (Convex FUMI) algorithm [7] was the first implementation of a FUMI method. cFUMI assumes each data point is a convex combination of target and non-target prototypes and is able to address the unspecificity

Further author information:

A. Z.: zarea@missouri.edu, C. J.: cjr25@mail.missouri.edu. This material is based upon work supported by the National Science Foundation under Grant No. IIS-1350078 - CAREER: Supervised Learning for Incomplete and Uncertain Data and a scholarship from China Scholarship Council (No. 201206960005).

Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XXI, edited by Miguel Velez-Reyes, Fred A. Kruse, Proc. of SPIE Vol. 9472, 947212
 · doi: 10.1117/12.2176889

associated with unknown target and non-target proportion values. cFUMI requires specific instance-level labels to denote whether there is some non-zero proportion of the target concept in each data point. Thus, cFUMI fails to address bag-level uncertainty. For example, in many cases, even binary labeled training with unspecificity in target proportion values is not available, e.g., location uncertainty in target detection problems where given GPS coordinates have an accuracy that ranges across several pixels. The eFUMI [9, 10] algorithm was proposed to address both unspecificity in proportion values and uncertainty in location in training data. eFUMI and was shown to be effective in learning target and target concepts, number of non-target concepts, and the presence or absence of these concepts in each data point. After learning a target concept, target detection can then be performed using a signature-based target detector [11–13].

In this paper the Multi-target Extended Function of Multiple Instances (Multi-target eFUMI) method is proposed. This method extends the eFUMI algorithm to estimate multiple target signatures. This algorithm address the case when a target class cannot be effectively represented by only one target concept. Results from multi-target sub-pixel target detection on simulated and real airborne hyperspectral data are shown.

2. THE MULTI-TARGET eFUMI ALGORITHM

Multi-target eFUMI extends previous work to learn the multiple target concepts given bag-level labeled training data. Specifically, assume the input training data, $\{\mathbf{x}_i\}_{i=1}^N \in \mathbb{R}^D$ have been partitioned into K "bags," $\mathbf{B} = \{B_1, \ldots, B_K\}$, and has been labeled with bag-level labels, $L = \{L_1, \ldots, L_K\}$. Each training point is assumed to be a function of target and non-target concepts $\mathbf{E}$, $\mathbf{x}_i = f(\mathbf{E}_i, \mathbf{p}_i)$, where $\mathbf{E}_i$ is a subset of $\mathbf{E}$ and $\mathbf{p}_i$ is the functional parameter that defines the relationship between $\mathbf{x}_i$ and $\mathbf{E}_i$; $\mathbf{E} = \cup_i \mathbf{E}_i$. The goal of Multi-target eFUMI is to estimate target and background concepts, $\mathbf{E}$, the number of needed non-target concepts, M, and the function parameters, $\mathbf{p}_i$. In this paper, we consider the functional form $f(\cdot)$ to be a convex combination, $\mathbf{x}_i = \sum_{t=1}^T p_{it}\mathbf{e}_t + \sum_{k=1}^M p_{ik}\mathbf{e}_k$ subject to the constraints shown in (1).

$$\sum_{t=1}^{T} p_{it} + \sum_{k=1}^{M} p_{ik} = 1, p_{it} \geq 0, p_{ik} \geq 0 \tag{1}$$

The bag-level labels for the training data are assigned as shown in (2). If B_j is a positive bag (i.e., $L_j = 1$), there exists least one data point in B_j with a non-zero p_{it} indicating some presence of target. However, the exact number of data points in a positive bag with a target contribution and how significant this contribution is (i.e., value of p_{it}), are unknown. Furthermore, if B_j is a negative bag (i.e., $L_j = 0$), then this indicates that none of the data in B_j contains any target concept.

$$\begin{array}{l} \text{if } L_j = 1, \exists \mathbf{x}_i \in B_j \text{ s.t. } \mathbf{x}_i = \sum_{t=1}^T p_{it}\mathbf{e}_t + \sum_{k=1}^M p_{ik}\mathbf{e}_k + \boldsymbol{\varepsilon}_i; \sum_{t=1}^T p_{it} > 0 \\ \text{if } L_j = 0, \forall \mathbf{x}_i \in B_j, \mathbf{x}_i = \sum_{k=1}^M p_{ik}\mathbf{e}_k + \boldsymbol{\varepsilon}_i \end{array} \tag{2}$$

Given training data in this assumed form, the *complete* data log-likelihood is proportional to (3), where u is a regularization parameter controlling the relative importance of the first, second and third terms. The first term in (3) computes the squared error between the input data and its approximation using currently estimated concepts and corresponding proportions. The weights are set to $w_{l(\mathbf{x}_i)} = \frac{\alpha N_n}{N_t}$ for points in positively labeled bags and $w_{l(\mathbf{x}_i)} = 1$ for points in negatively labeled bags in order to balance the influence from positive bags and negative bags on the objective function, as described in [8]. The second and third terms of the objective encourages a compact simplex defined by the concepts by minimizing the squared difference between each concept and the global data mean, $\boldsymbol{\mu}_0$. These terms were motivated by the volume-related term in the SPICE [14] algorithm. The fourth term is a sparsity promoting term used to determine M, where $\gamma_k = \frac{\Gamma}{\sum_{n=1}^N p_{nk}^{(j-1)}}$, Γ is a parameter constant that controls the degree sparsity is promoted and $p_{nk}^{(j-1)}$ are the proportion values estimated in the previous iteration. During each iteration, this term drives proportion values with respect to unnecessary endmembers to zero and prunes the unnecessary endmembers as in [14].

Multi-target eFUMI estimates the desired parameter values (i.e., $\mathbf{e}_t, \quad \forall t = 1, \ldots T$, $\mathbf{E}$, $\mathbf{P}$ and M) by minimizing the objective function shown in (3). To address the fact that the instance level label z_i for each data point is unknown, Multi-target eFUMI adopts an EM approach to estimate the probability of instance label value $P(z_i|\boldsymbol{x}_i, \boldsymbol{\theta}^{(j-1)})$ in

$$F = \frac{(1-u)}{2}\sum_{i=1}^{N} w_i \left\| (\mathbf{x}_i - z_i \sum_{t=1}^{T} p_{it}\mathbf{e}_t - \sum_{k=1}^{M} p_{ik}\mathbf{e}_k) \right\|_2^2 + \frac{u}{2}\sum_{t=1}^{T} \left\| (\mathbf{e}_t - \boldsymbol{\mu}_0) \right\|_2^2 + \frac{u}{2}\sum_{k=1}^{M} \left\| (\mathbf{e}_k - \boldsymbol{\mu}_0) \right\|_2^2 + \sum_{k=1}^{M}\gamma_k \sum_{i=1}^{N} p_{ik} \tag{3}$$

$$\begin{aligned} E[F] \quad = \quad & \sum_{z_i \in \{0,1\}} \left[\frac{(1-u)}{2}\sum_{i=1}^{N} w_i P(z_i|\mathbf{x}_i, \boldsymbol{\theta}^{(j-1)}) \left\| \mathbf{x}_i - z_i \sum_{t=1}^{T} p_{it}\mathbf{e}_t - \sum_{k=1}^{M} p_{ik}\mathbf{e}_k \right\|_2^2 \right] + \frac{u}{2}\sum_{t=1}^{T} \|\mathbf{e}_t - \boldsymbol{\mu}_0\|_2^2 \\ & + \frac{u}{2}\sum_{k=1}^{M} \|\mathbf{e}_k - \boldsymbol{\mu}_0\|_2^2 + \sum_{k=1}^{M}\gamma_k \sum_{i=1}^{N} p_{ik} \end{aligned} \tag{4}$$

the *E-step* using (5) where $\boldsymbol{\theta}^{j-1}$ is the set of parameters estimated at iteration $j-1$, $r_b = \left\| \mathbf{x}_i - \sum_{k=1}^{M} p_{ik}\mathbf{e}_k \right\|_2^2$ is the approximation residual between $\mathbf{x}_i$ and its representation using only background endmembers, and β is a scaling parameter for r_b.

$$P(z_i|\mathbf{x}_i, \boldsymbol{\theta}^{(j-1)}) = \begin{cases} e^{-\beta \left\| \mathbf{x}_i - \sum_{k=1}^{M} p_{ik}\mathbf{e}_k \right\|_2^2} & \text{if } z_i = 0, L_i = 1 \\ 1 - e^{-\beta \left\| \mathbf{x}_i - \sum_{k=1}^{M} p_{ik}\mathbf{e}_k \right\|_2^2} & \text{if } z_i = 1, L_i = 1 \\ 0 & \text{if } z_i = 1, L_i = 0 \\ 1 & \text{if } z_i = 0, L_i = 0 \end{cases} \tag{5}$$

$P(z_i|\mathbf{x}_i, \boldsymbol{\theta}^{(j-1)})$ in (5) can be interpreted as follows: if a point $\mathbf{x}_i$ is a negative point it should be well represented by only the background endmembers with very small approximation residual, r_b, thus $P(z_i = 0|\mathbf{x}_i, \boldsymbol{\theta}^{(j-1)}) = e^{-\beta r_b} \rightarrow 1$. Otherwise, if $\mathbf{x}_i$ is a target point, it may not be fully represented by only the background endmembers, so the approximation residual r_b will be large and $P(z_i = 1|\mathbf{x}_i, \boldsymbol{\theta}^{(j-1)}) = 1 - e^{-\beta r_b} \rightarrow 1$. Note, z_i is unknown only for points from positive bags; for all points from negative bags z_i are fixed to zero. This constitutes the *E-step* of the EM algorithm. The *M-step* is performed by optimizing the expectation of the objective function shown in (4) for each of the desired parameters. The method is summarized in Alg. 1.

Algorithm 1 Multi-target eFUMI EM algorithm

1: Initialize $\boldsymbol{\theta}^0 = \{\mathbf{e}_t, t = 1, \ldots T, \mathbf{E}, \mathbf{P}\}$, normalize input data, $j = 1$
2: **repeat**
3: ***E-step***: Compute $P(z_i|\mathbf{x}_i, \boldsymbol{\theta}^{(j-1)})$ given $\boldsymbol{\theta}^{j-1}$
4: ***M-step***:
5: Update $\mathbf{e}_t$ and $\mathbf{E}$ by maximizing (4) wrt. $\mathbf{e}_t, t = 1, \ldots, T, \mathbf{E}$
6: Normalize $\mathbf{e}_t, t = 1, \ldots T, \mathbf{E}$
7: Update $\mathbf{P}$ by maximizing (4) wrt. $\mathbf{P}$ s.t. (1)
8: Prune each $\mathbf{e}_k, k = 1, \ldots, M$ if $\max_n(p_{nk}) \leq \tau$ where τ is a fixed threshold (e.g. $\tau = 10^{-6}$)
9: $j \leftarrow j + 1$
10: **until** Convergence
11: **return** $\mathbf{e}_t, t = 1, \ldots T, \mathbf{E}, \mathbf{P}$

3. EXPERIMENTAL RESULTS

The proposed Multi-target eFUMI was tested both on synthetic data and real hyperspectral data.

3.1 Multi-target eFUMI on Synthetic Data

Multi-target eFUMI was applied to simulated data generated from four spectra selected from the ASTER spectral library [15]. Specifically, the Red Slate, Verde Antique, Phyllite and Pyroxenite spectra from the rock class with 211

bands and wavelengths ranging from 0.4μm to 2.5μm (as shown in Fig. 1) were used as endmembers to generate hyperspectral data. Red Slate and Phyllite were labeled as the target endmembers. Five bags were generated with 500 points in each bag. The first two bags were positive bags and each had 100 points with non-zero target proportion values corresponding to at least one of the target endmembers. The constituent background endmembers for each data point were drawn randomly. The corresponding proportion values were generated by drawing from a Dirichlet distribution with mean $[0.2, 0.2, 0.3, 0.3]$. The synthetic data was added by Gaussian white noise and has signal-to-noise ratio at $30dB$.

In this synthetic experiment, the normalized mean square error (NMSE) and mean spectral angle distance (MSAD) between estimated target spectra and ground truth spectra (Red Slate and Phyllite) were calculated to evaluate the results, where $NMSE = \mathbf{E}\left[\frac{\|\mathbf{e}_{tru}-\mathbf{e}_{est}\|}{\|\mathbf{e}_{tru}\|}\right]$; $MSAD = \mathbf{E}\left[\cos^{-1}\left(\frac{\mathbf{e}_{tru}^T\cdot\mathbf{e}_{est}}{\|\mathbf{e}_{tru}\|\cdot\|\mathbf{e}_{est}\|}\right)\right]$; $\mathbf{e}_{tru}$ and $\mathbf{e}_{est}$ are the true and estimated target spectra, respectively. The parameter settings for Multi-target eFUMI used here were $u = 0.05, M = 4, \Gamma = 15, \beta = 25$. The results are shown in Table 1, Fig. 2(a) and 2(b) are example plots of the estimated target signatures. Experimental results show that Multi-target eFUMI algorithm is capable of estimating Multi-target sub-pixel target spectra and is robust to noise.

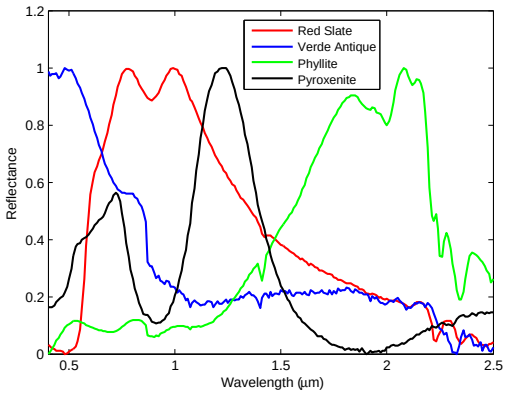
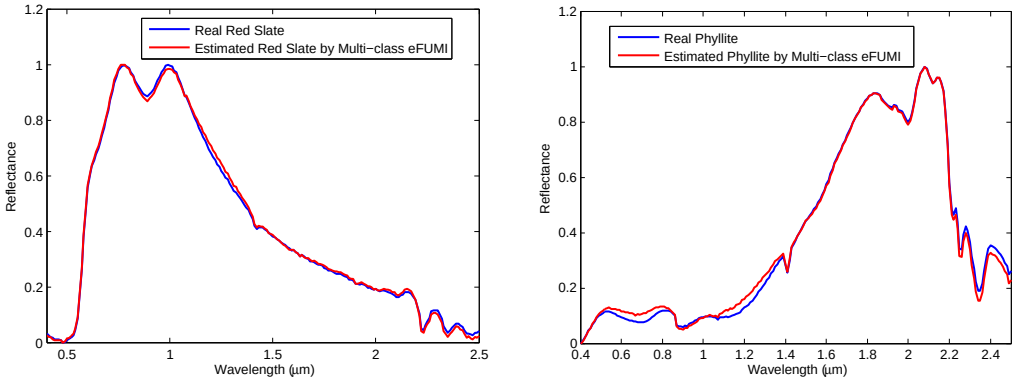


Figure 1. Signatures from ASTER library used to generate simulated data



(a) Red Slate (b) Phyllite

Figure 2. Estimated target spectra

3.2 Multi-target eFUMI on RIT SHARE 2012 Data

Then Multi-target eFUMI was also tested on the RIT SHARE 2012 data set collected near Rochester, NY [16]. A subset of the AVON hyperspectral imagery, AVON PM collected in the afternoon of September 20, 2012 containing 305×320 pixels with 360 bands corresponding to wavelengths from $400.3\ nm$ to $2452.8\ nm$ were used (following [16], the noisy bands 256-279 were pruned). The spatial resolution was $1m^2$. There were two target types with 12

Table 1. Error and standard deviation (over 10 runs)

Target	NMSE($\times 10^{-2}$)	MSAD($\times 10^{-5}$)
Red Slate	2.94 ± 0.44	2.79 ± 0.39
Phyllite	4.64 ± 1.07	4.33 ± 1.08

Figure 3. Labeled bags and 24 scattered target points

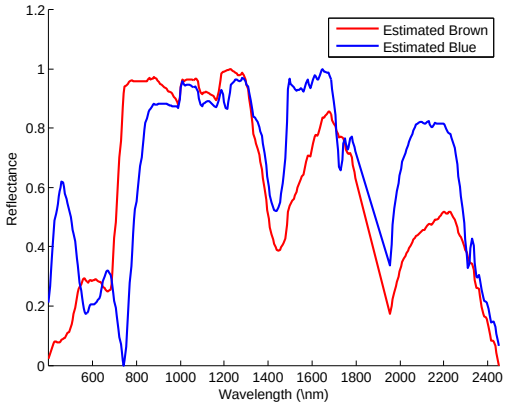


Figure 4. Estimated Brown and Blue spectra

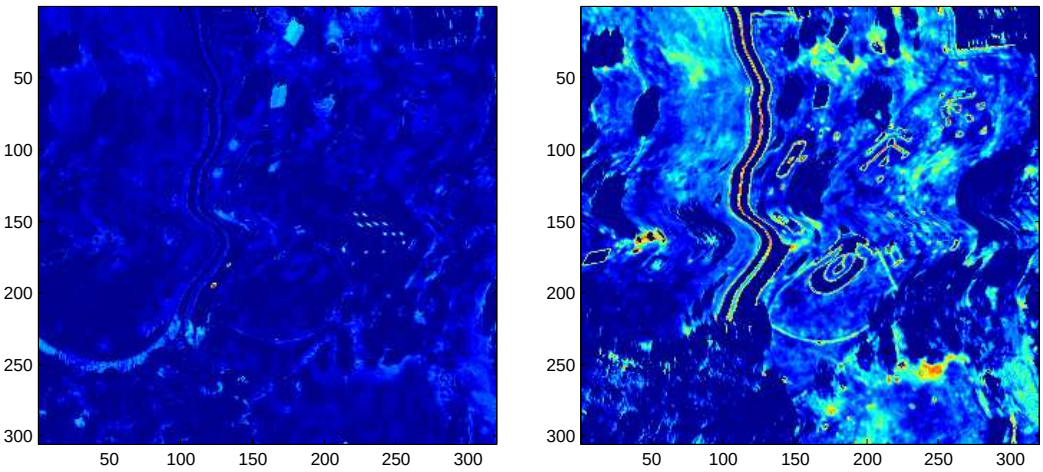
examples each type: man-made blue and brown tarps. An illustration shows how the positive and negative bags were formed are shown in Fig. 3, where a rectangular region (indicated with a red rectangle) was grouped together to form one large positive bag. Several large background regions (marked with the yellow rectangles) were labeled as negative bags. Fig. 6(b) is shows a zoomed image of the target region. The parameter setting for Multi-target eFUMI was $u = 0.05, M = 4, \Gamma = 1$ and $\beta = 120$.

After running Multi-target eFUMI, the estimated Blue and Brown target signatures are shown in Fig. 4; Fig. 5(a)-Fig. 5(e) display the corresponding proportion maps for estimated target endmembers Blue and Brown, and estimated background endmembers, vegetation, asphalt and cars, respectively. Fig. 5(a) clearly indicates the presence of Blue targets. However, it is difficult to differentiate the brown targets in Fig. 5(b) since the brown material is likely to be confused with the brown soil in the edges of roads or mixed in with grass. Fig. 5(c)-Fig. 5(e) also clearly indicate the presence of tree and grass, asphalt road and parking lot, and cars, respectively.

To more effectively verify the quality of estimated target spectra, the ACE detector [11–13] was applied for target detection for both target types. Fig. 6 shows the detection map for both Blue and Brown targets. Fig. 6(a) shows the entire ACE detection map based on estimated Brown signature. When comparing this map with the proportion map of the brown target signature in Fig. 5(b), it can be seen that the target signature is effective for target detection of the brown targets (and effective at distinguishing between borwn target and brown soil). Also, when comparing Fig. 6(b) that displays a zoomed target region of the RGB image to the zoomed ACE detection maps for Brown and Blue targets shown in Fig. 6(d) and 6(c), it can be seen that the targets are accurately detected. This illustrates that Multi-target eFUMI is capable of detecting quality Multi-target sub-pixel target spectra for real hyperspectral detection task given training data with spatial uncertainty.

4. DISCUSSION AND FUTURE WORK

In this paper, a Multi-target sub-pixel target signature characterization algorithm that is based on the functions of multiple instances framework is presented. The proposed method estimates target and non-target concepts from binary labeled bags of training data. Target detection experiments on synthetic data and real hyperspectral data show that Multi-target eFUMI is able to learn Multi-target target concepts from bags of data following the definition of MIL problem. Future work will include investigating the performance of some functional form $f(\cdot)$ other than linear mixing model. Current work also includes generalizing Multi-target eFUMI to a more general semi-supervised dictionary learning approach that can be applied to a wide range of applications and data types.



(a) Blue

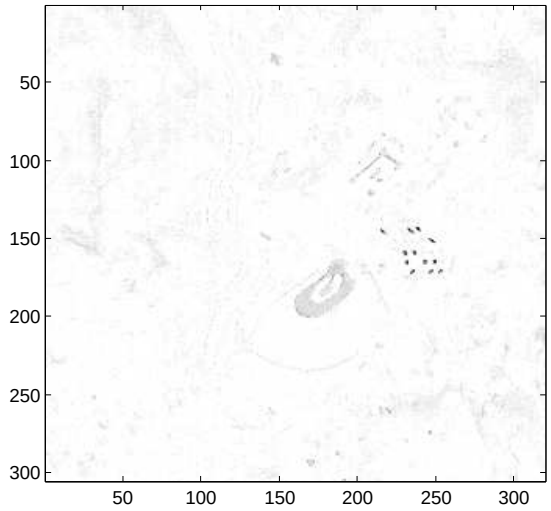
(b) Brown

(c) Vegetation

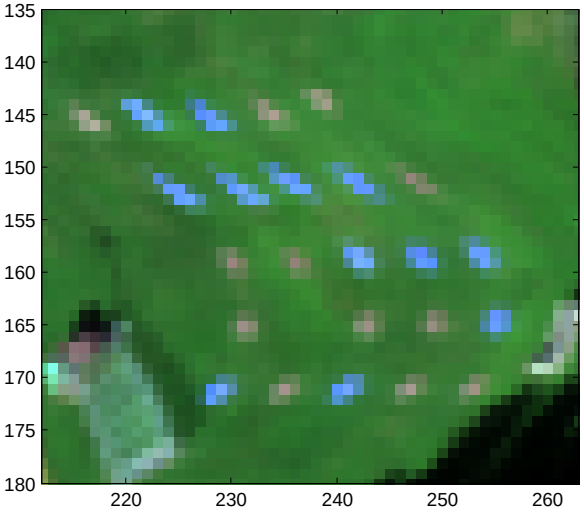
(d) Asphalt

(e) Cars

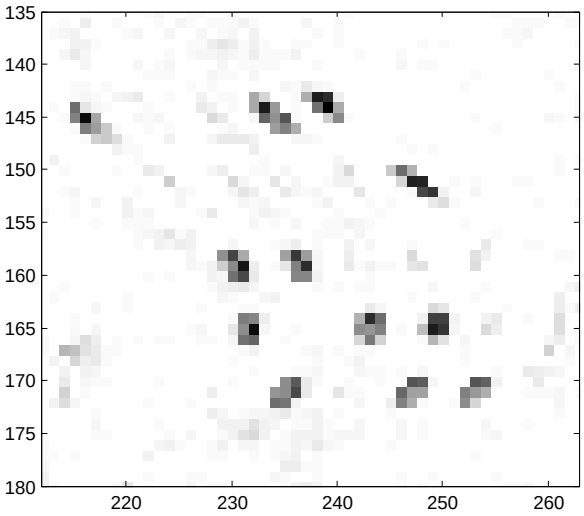
Figure 5. RIT SHARE 2012 AVON PM Proportion Map



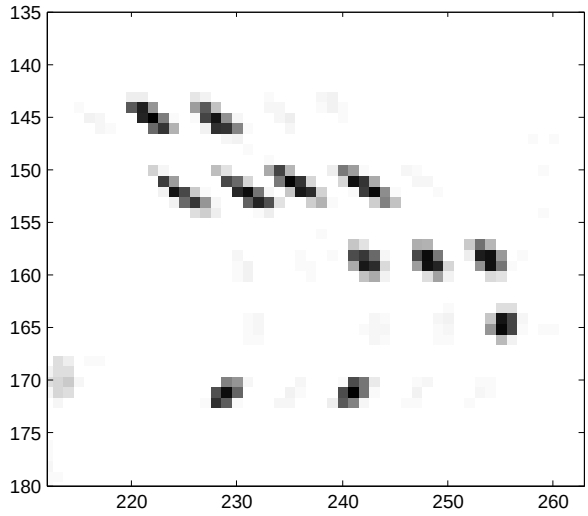
(a) ACE output map for brown targets



(b) Zoomed target region of Fig. 3



(c) ACE output map zoomed for brown targets



(d) ACE output map zoomed for blue targets

Figure 6. ACE output with estimated Blue and Brown spectra

References

[1] Dietterich, T. G., Lathrop, R. H., and Lozano-Pérez, T., "Solving the multiple instance problem with axis-parallel rectangles," *Artificial intelligence* **89**, 31–71 (1997).

[2] Maron, O. and Lozano-Pérez, T., "A framework for multiple-instance learning," in [*Advances in neural information processing systems*], **10**, 570–576 (1998).

[3] Zhang, Q. and Goldman, S., "EM-DD: An improved multiple-instance learning technique," in [*Advances in Neural Information Processing Systems*], **2**, 1073–1080, MIT; 1998 (2002).

[4] Raykar, V., Krishnapuram, B., Bi, J., Dundar, M., and Rao, R., "Bayesian multiple instance learning: automatic feature selection and inductive transfer," in [*Proceedings of the 25th international conference on Machine learning*], 808–815, ACM New York, NY, USA (2008).

[5] Bolton, J. and Gader, P., "Multiple instance learning for hyperspectral image analysis," in [*Geoscience and Remote Sensing Symposium (IGARSS), 2010 IEEE International*], 4232–4235 (2010).

[6] Bolton, J., Gader, P., Frigui, H., and Torrione, P., "Random set framework for multiple instance learning," *Information Sciences* **181**(11), 2061–2070 (2011).

[7] Zare, A. and Gader, P., "Pattern recognition using functions of multiple instances," in [*Proceedings of the International Conference on Pattern Recognition*], 1092–1095 (Aug. 2010).

[8] Zare, A. and Gader, P., "Multiclass subpixel target detection using functions of multiple instances," in [*SPIE Defense, Security, and Sens.*], 804811, International Society for Optics and Photonics (2011).

[9] Zare, A. and Jiao, C., "Extended functions of multiple instances for target characterization," in [*Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), 2014 6th Workshop on*], 1–4 (2014).

[10] Jiao, C. and Zare, A., "Functions of multiple instances for learning target signatures," *IEEE Trans. Geosci. Remote Sens.* . (in press).

[11] Kraut, S. and Scharf, L., "The CFAR adaptive subspace detector is a scale-invariant GLRT," *IEEE Trans. Signal Process.* **47**, 2538 –2541 (Sept. 1999).

[12] Kraut, S., Scharf, L., and McWhorter, L., "Adaptive subspace detectors," *IEEE Trans. Signal Process.* **49**(1), 1–16 (2001).

[13] Basener, W. F., "Clutter and anomaly removal for enhanced target detection," in [*SPIE Defense, Security, and Sens.*], **7695**, 769525 (2010).

[14] Zare, A. and Gader, P., "Sparsity promoting iterated constrained endmember detection for hyperspectral imagery," *IEEE Geoscience and Remote Sensing Letters* **4**, 446–450 (July 2007).

[15] Baldridge, A., Hook, S., Grove, C., and Rivera, G., "The ASTER spectral library version 2.0," *Remote Sensing of Environment* **113**(4), 711–715 (2009).

[16] Giannandrea, A. et al., "The SHARE 2012 data campaign," in [*SPIE Defense, Security, and Sensing*], 87430F (2013).