Band Selection in Hyperspectral Imagery Using Sparse Support Vector Machines

Sofya Chepushtanova[1], Christopher Gittins[2], and Michael Kirby[1]

[1] *Department of Mathematics, Colorado State University, Fort Collins, CO*
[2] *UTC Aerospace Systems, Westford, MA*

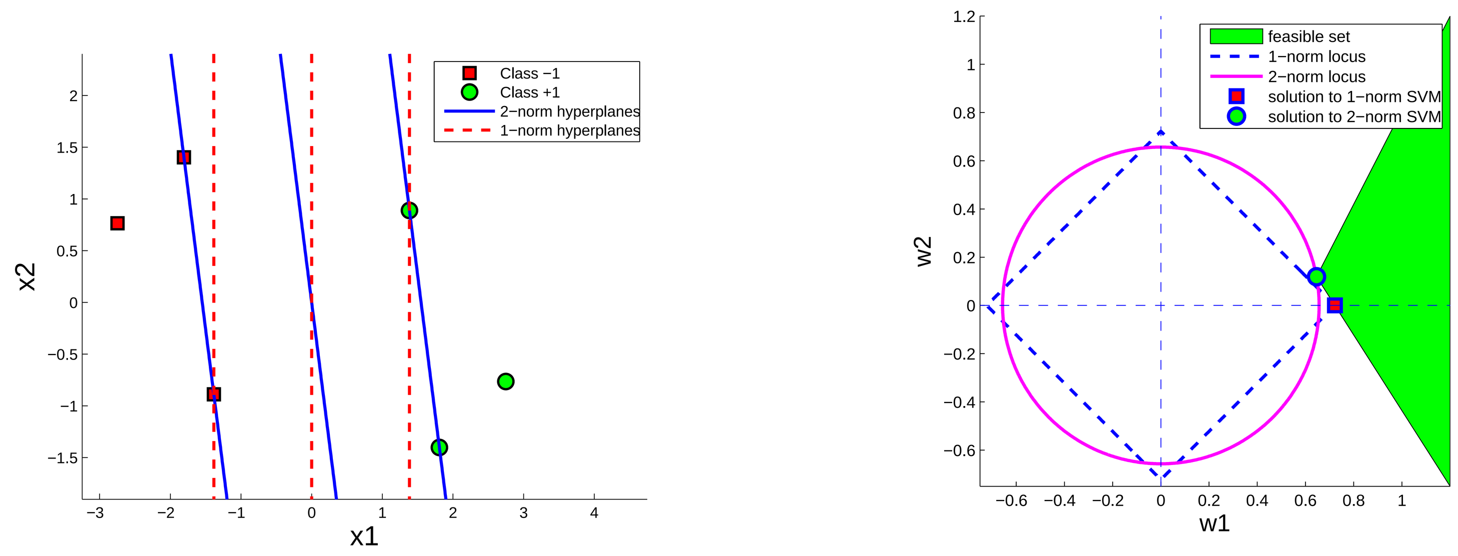
Introduction

- **Goal:** to identify a subset of spectral bands that contain the most discriminatory information during classification process.
- A **sparse support vector machine (SSVM)** is a binary classifier with ℓ_1-norm regularization that sets many feature weights to zero $\Rightarrow$ embedded feature selection tool.
- We propose a **band selection framework** based on sparsity and effectiveness of SSVMs combined with **bagging** (**b**ootstrap **agg**regat**ing**) to enhance the robustness of the classifier.
- We extend the binary band selection to a **multiclass case** and illustrate the performance of the method on two data sets.

Sparse SVM: why does ℓ_1-norm induce sparsity?

$$\min_{\mathbf{w},b,\xi} \ \|\mathbf{w}\|_1 + C e^T \xi$$
$$\text{s. t. } D(X\mathbf{w} + b\mathbf{e}) \geq \mathbf{e} - \xi,$$
$$\xi \geq 0$$

$$\|\mathbf{w}\|_1 = |w_1| + |w_2|$$
$$\|\mathbf{w}\|_2 = \sqrt{w_1^2 + w_2^2}$$



Loci of points in the band space and SVM solutions corresponding to ℓ_1-norm and ℓ_2-norm regularization.

Band selection

Two-class band selection algorithm:

Input: Training data X, set of kept bands $S = [1, 2, \dots, n]$

Variability reduction:

- Sample with replacement from $\mathbf{X}$ to obtain $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N$
- Train N SSVM models $f_j(\mathbf{x}) = (\mathbf{w}^j)^T\mathbf{x} + b^j \rightarrow$ weight vectors $\mathbf{w}^j$
- For $k = 1 : n$, remove k-th band if $\#\{|w_k^j| <$ tolerance, $j = 1, \dots, N\} \geq 0.95 * N$, $\rightarrow$ update S.
- Restrict $\mathbf{X}$ to selected bands: $\mathbf{X}_{new} = \mathbf{X}(:, S)$

Final selection:

- Train an SSVM model f on $\mathbf{X}_{new} \rightarrow \mathbf{w}$
- Rank $\mathbf{w}$ values by magnitude $\rightarrow \mathbf{w}^r$, keep ranked band indices in R
- Compare: if $|w_{i_k}^r| / |w_{i_{k+1}}^r| = O(10^L)$ and $L \geq 1$ for some $k = k^*$, remove bands from R starting from index $i_{k*+1} \rightarrow$ update $S = S \setminus S(R)$
- Restrict $\mathbf{X}_{new}$ to selected bands: $\mathbf{X}_{new} = \mathbf{X}_{new}(:, S)$

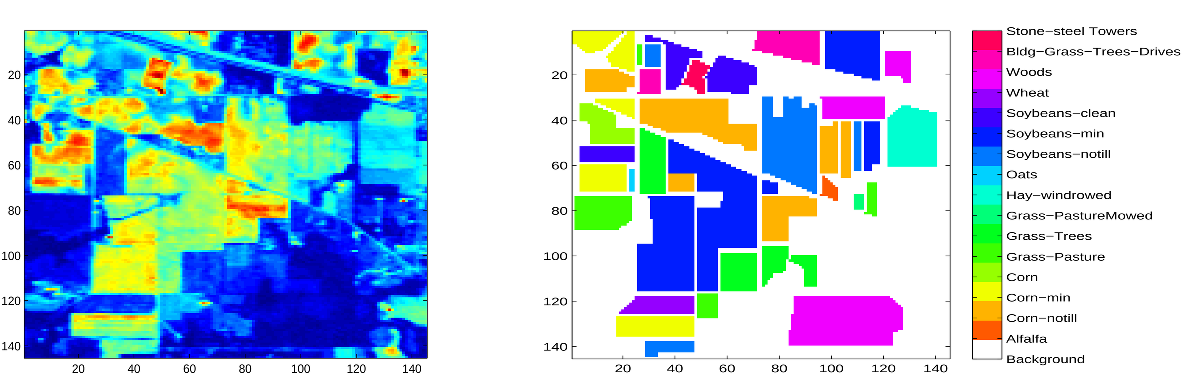
Output: Selected bands S, SSVM model f

Multiclass band selection options:

- **Method I:** Rank bands by their occurence number in all the two-class subsets and select K bands with the highest frequency values.
- **Method II:** Rank bands in each two-class subset by magnitude and take the superset of M top bands in each subset (we take $M = 1$).
- **Method III:** Use Ward's Linkage Strategy Using Mutual Information (WaLuMI) method [3] as a pre-selection step followed by an application of the SSVM.

***Note**: To compare our multiclass results with other methods, we adopted one-against-one (OAO) SSVM classification approach.*

Data set 1: AVIRIS Indian Pines data set



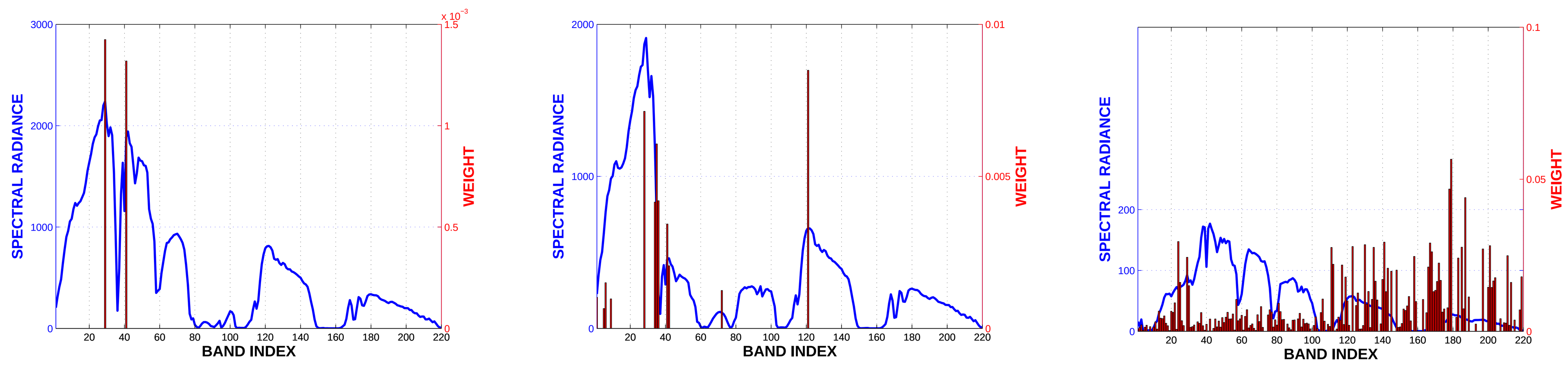
- **0.4** to **2.4** μm.
- **145** $\times$ **145** images, **220** spectral bands, **16** classes (from **20** to **2468** pixels).
- Two-thirds agriculture and one-third other natural vegetation.
- Water absorption bands: **104 − 108, 150 − 163**, and **220**.

Data set 1: binary band selection experiments

Classes	All Bands Kept	SSVM		WaLuMI [3] + SSVM		Lasso Logistic Regression	
		# Bands Kept	Accuracy %	# Bands Kept	Accuracy %	# Bands Kept	Accuracy %
Corn-min and Woods	100.00	2	100.00	2	99.9	12	100.00
Corn-notill and Grass/Trees	99.73	12	99.73	12	100	19	98.9
Soybeans-notill and Soybeans-min	89.58	179	89.23	-	-	127	89.52

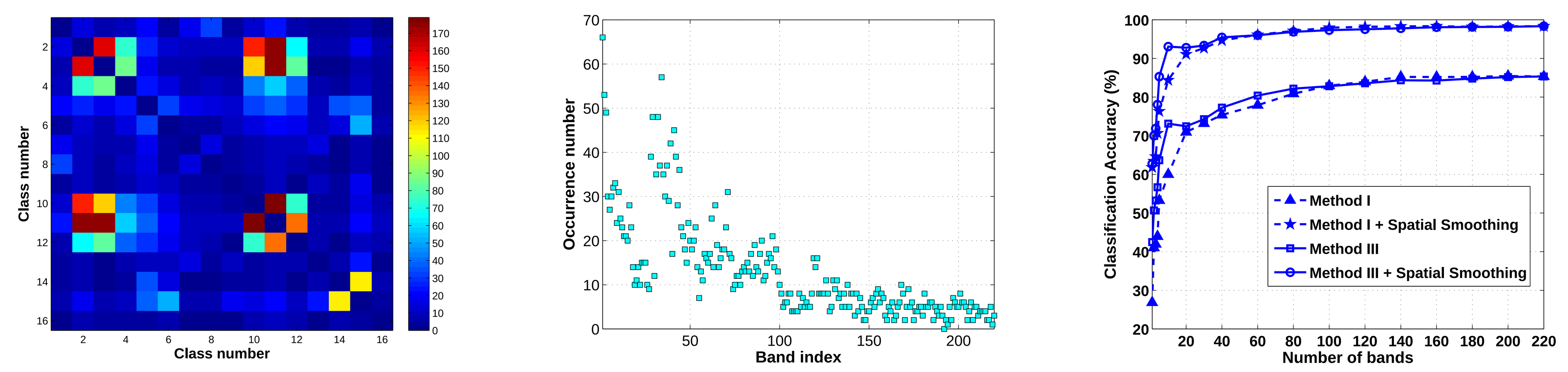
Difference in spectral signatures and the selected band weights for:

Corn-min & Woods | Corn-notill & Grass/Trees | Soybeans-notill & Soybeans-min



Data set 1: multiclass band selection experiments

- **Spacial smoothing:** after OAO SSVM is performed, we sum over labels in the **3** $\times$ **3** neighborhood of a testing pixel and assign the most frequently occuring class label to the pixel.



Comparison of accuracy rates (%):

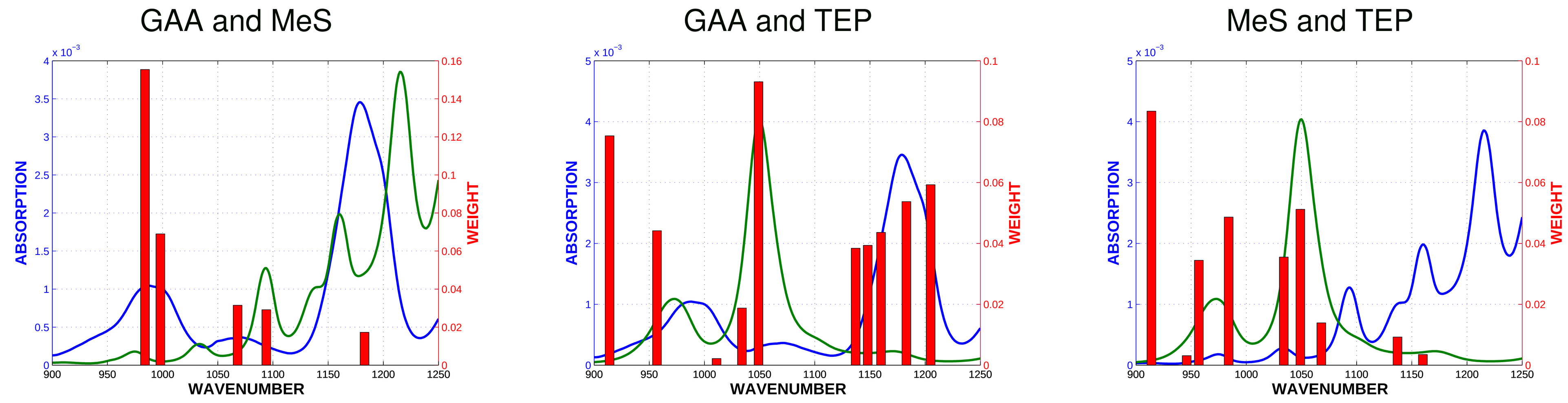
# Bands Kept	Method I	Method II	Method III (WaLuMI + SSVM)	B-SPICE + RVM [4]	WaLuMI + NN [3]
220	98.36	-	98.36	93.9	-
57	95.66	97.3	96.22	-	-
34	93.15	-	93.03	86.4	80
18	88.59	-	92.78	78.3	82
10	84.37	-	93.07	-	81
5	76.32	-	85.29	-	71

Data set 2: LWIR (long-wavelength infrared) data set

- **8** to **11** μm.
- **256** $\times$ **256** images, **20** spectral bands.
- Three chemicals: Glacial Acetic Acid (GAA), Methyl Salicylate (MeS), and Triethyl Phosphate (TEP).

Binary band selection results:

GAA and MeS | GAA and TEP | MeS and TEP



Band	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Wavenumber (cm^{-1})	1011	1034	1049	1068	1094	1107	1137	1148	1160	1183	1205	1216	1237	914	936	946	957	971	984	998

Conclusions

- The SSVM algorithm is a supervised embedded technique that simultaneously performs band selection and classification.
- Bagging and spatial smoothing steps are employed for variability reduction and improving the accuracy rates, respectively.
- The SSVM algorithm in both binary and multiclass cases demonstrated a good performance compared to other methods which makes it an effective framework in identifying the relevant spectral wavelengths.

References

[1] O. L. Mangasarian, *Exact 1-norm support vector machines via unconstrained convex differentiable minimization*, Journal of Machine Learning Research, 7, pp. 1517-1530, 2006.
[2] J. Bi, K. P. Bennett, M. Embrechts, C. M. Breneman, and M. Song, *Dimensionality reduction via sparse support vector machines*, Journal of Machine Learning Research, 3, pp. 1229-1243, 2003.
[3] A. Martinez-Uso, F. Pla, P. Garcia-Sevilla, and J. M. Sotoca, *Automatic band selection in multispectral images using mutual information-based clustering*, Proc. 11th Iberoamerican Congr.Pattern Recog., Cancun, Mexico, pp. 644-654, 2006.
[4] A. Zare and P. Gader, *Hyperspectral band selection and endmember detection using sparsity promoting priors*, IEEE Geoscience and Remote Sensing Letters, 2, vol.5, pp. 256-260, 2008.
[5] S. Chepushtanova, C. Gittins, and M. Kirby, *Band Selection in Hyperspectral Imagery Using Sparse Support Vector Machines*, submitted to SPIE DSS Proceedings 2014.