High-Dimensional Signal Processing Group

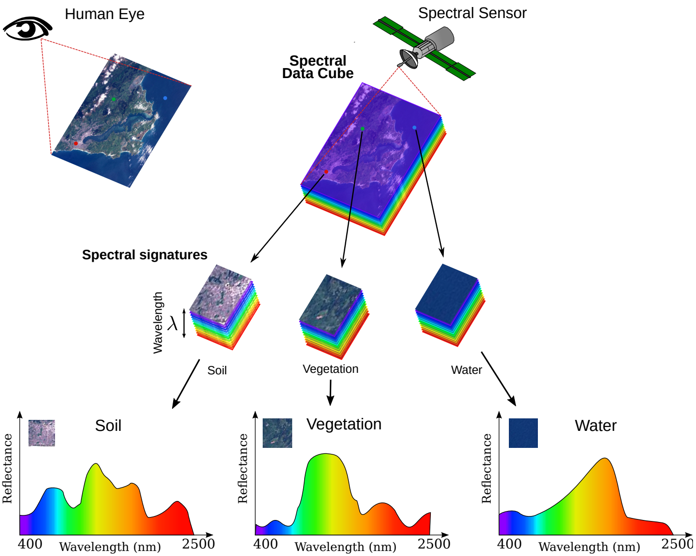
Spectral-spatial Classification From Multi-sensor Compressive Measurements Using Superpixels

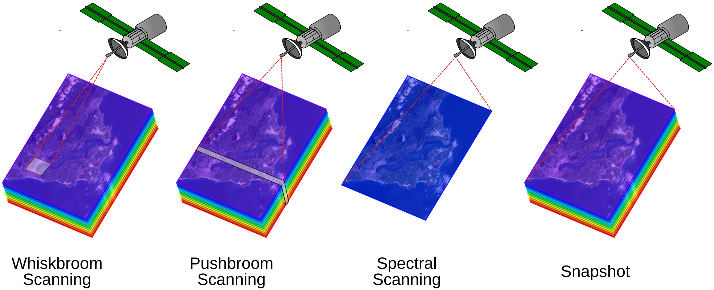
Carlos Hinojosa, Juan Marcos Ramirez, Henry Arguello

Universidad Industrial de Santander,
Department of Computer Science,
Bucaramanga, 680002, Colombia
carlos.hinojosa@saber.uis.edu.co

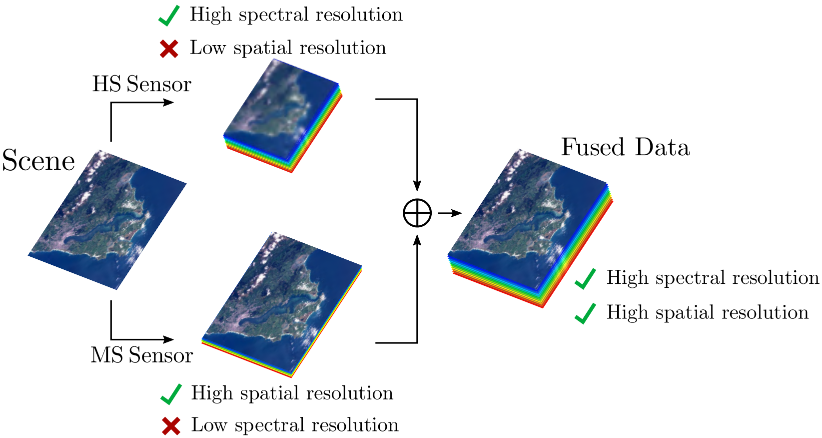
Taipei, Taiwan
September 25, 2019

Spectral Imaging

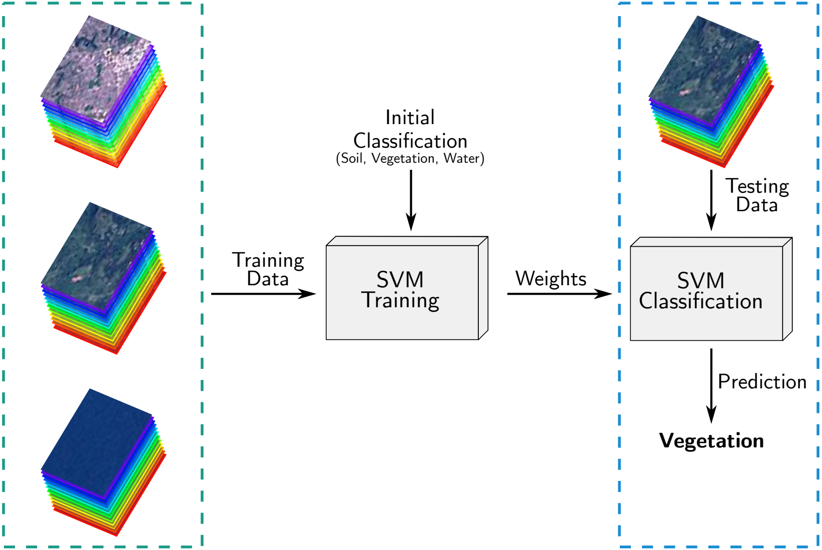


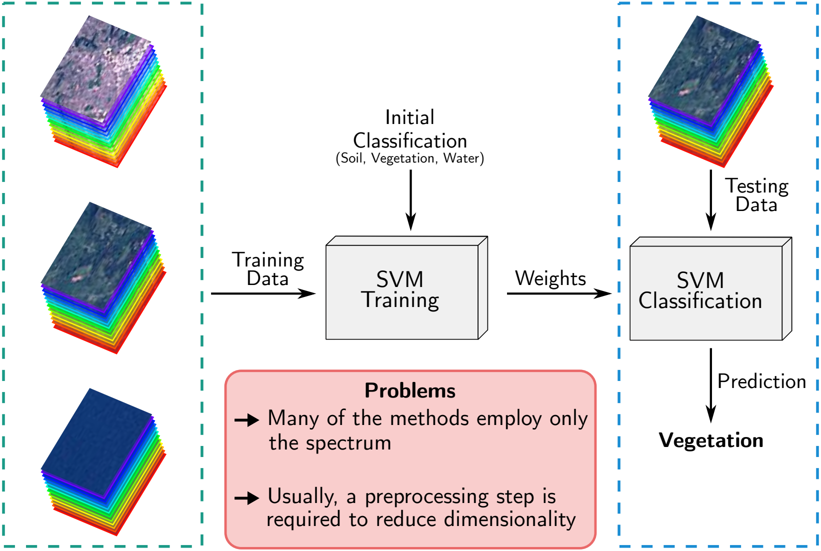


- Traditional spectral imaging techniques relies on Nyquist–Shannon sampling theorem.
- Require a fixed sampling rate along the three dimensions, leading to a large amount of captured data and large acquisition times.


High spectral resolution
Low spatial resolution
HS Sensor
Scene
MS Sensor
High spatial resolution
Low spectral resolution
Fused Data
High spectral resolution
High spatial resolution

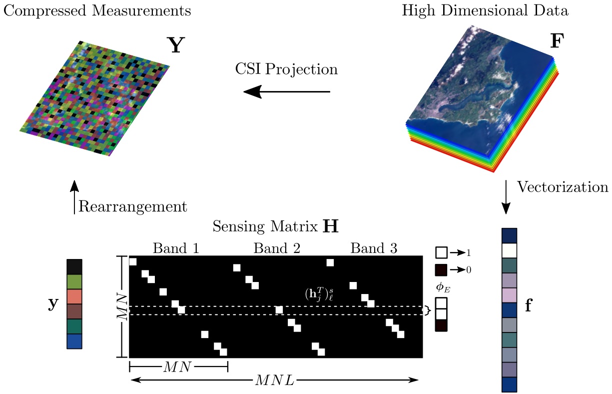
Classification From Fused Data





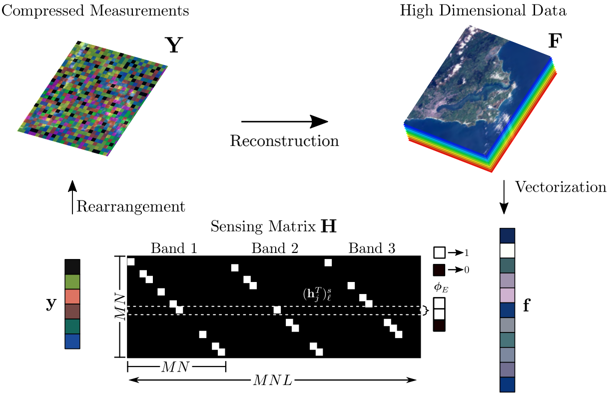
Problems

- Many of the methods employ only the spectrum
- Usually, a preprocessing step is required to reduce dimensionality



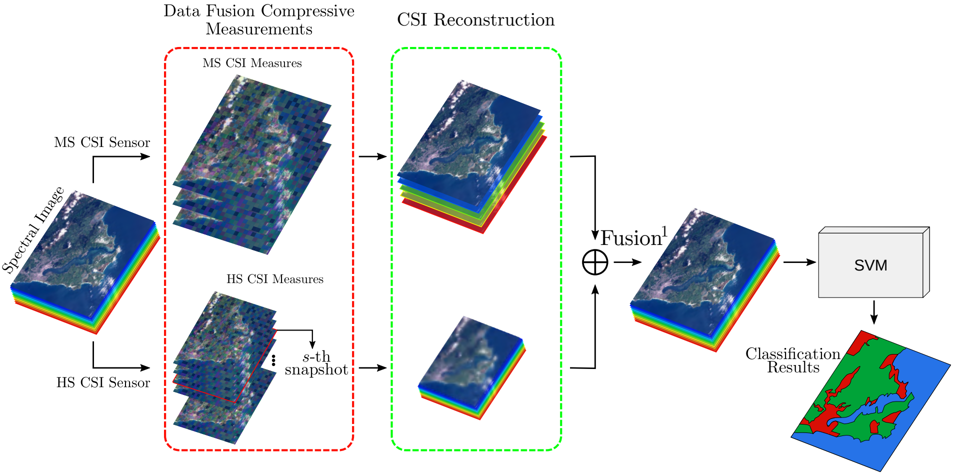
- Senses and simultaneously reduces the data dimension without any further processing step by capturing less samples.
- Assumes that $\mathbf{f}$ can be represented as a sparse vector $\boldsymbol{\theta}$ in some basis $\boldsymbol{\Psi}$, i.e., $\mathbf{f} = \boldsymbol{\Psi}\boldsymbol{\theta}$.
- CSI projections can be written in matrix notation as

$$\mathbf{y} = \mathbf{H}\mathbf{f} = \mathbf{H}\boldsymbol{\Psi}\boldsymbol{\theta}$$



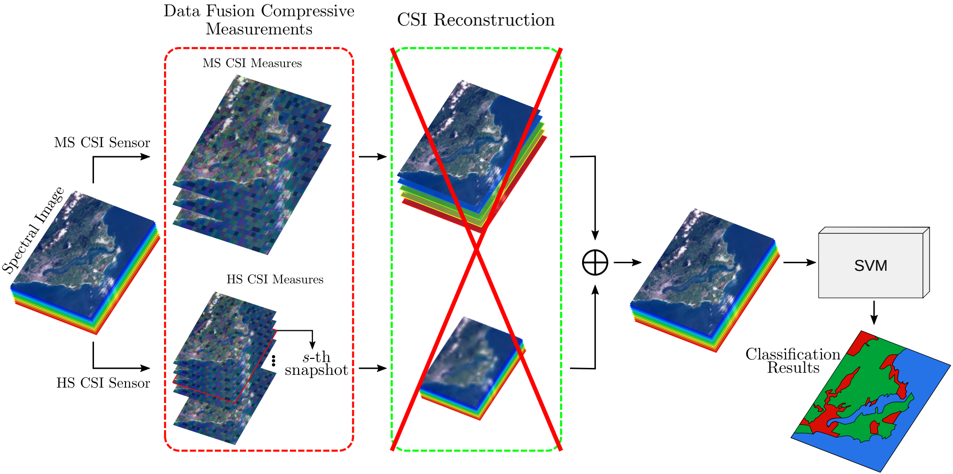
- CSI recovery consists of finding a sparse approximation $\hat{\boldsymbol{\theta}}$ by solving
$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \|\mathbf{y} - \mathbf{H}\boldsymbol{\Psi}\boldsymbol{\theta}\|_2^2 + \tau\|\boldsymbol{\theta}\|_1$$
- Computationally expensive optimization problem.
- GPSR[1] algorithm computes $O(kM^4N^4L)$ operations.

[1] M. Figueiredo, R. Nowak, & S. Wright. Gradient Projection for Sparse Reconstruction: Application to compressed sensing and other inverse problems. *IEEE Journal of Selected Topics in Signal Processing*.

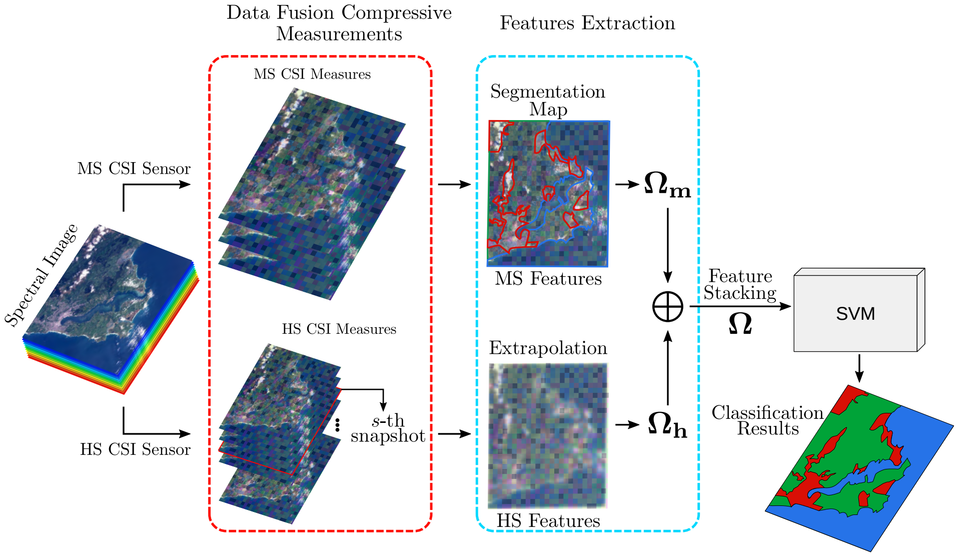


[1] Q. Wei, J. Bioucas-Dias, N. Dobigeon, & J. Tourneret. Hyperspectral and Multispectral Image Fusion Based on a Sparse Representation. IEEE Transactions on Geoscience and Remote Sensing, 2015.

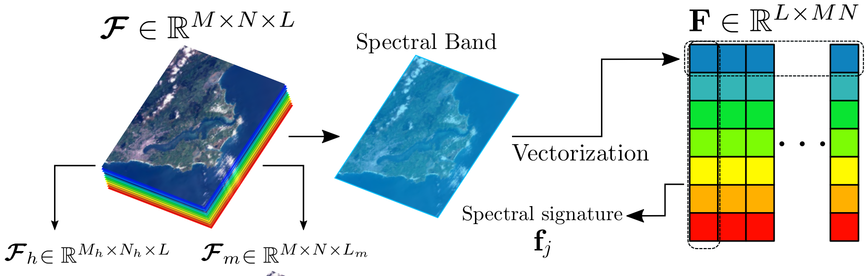
An Intuitive approach to CSI Classification

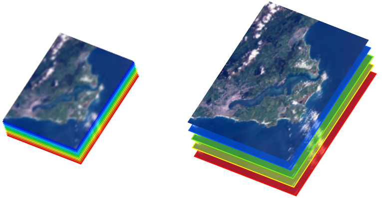


Proposed Method for CSI Classification



Multi-sensor Model



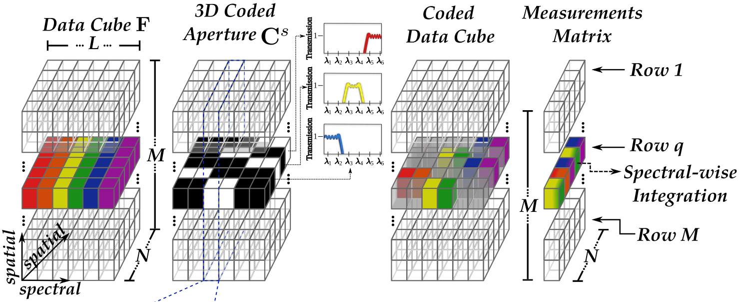


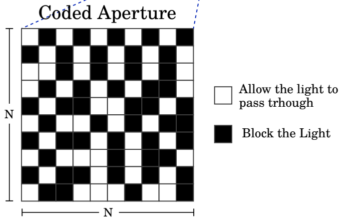
Matrix form

$$\mathbf{F_m} = \mathbf{D_m F}$$
$$\mathbf{F_h} = \mathbf{F D_h}$$

$\mathbf{D_m} \in \mathbb{R}^{L_m \times L}$ → Spectral downsampling matrix

$\mathbf{D_h} \in \mathbb{R}^{MN \times M_h N_h}$ → Spatial downsampling matrix



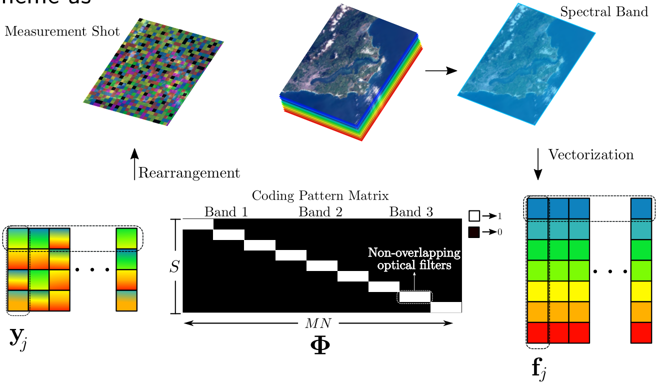


$$\hat{Y}^s_{m,n} = \sum_{k=0}^{L-1} \mathcal{F}_{m,n,k} C^s_{m,n,k}$$

- In this work, we use non-overlapping optical filters

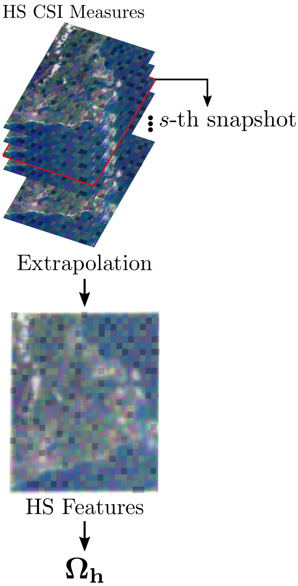
[1] X. Cao, T. Yue, X. Lin, S. Lin, X. Yuan, Q. Dai, & D. J. Brady. *Computational snapshot multispectral cameras: Toward dynamic capture of the spectral world. IEEE Signal Processing Magazine, 33(5), 95-108.*

- The number of measurement shots is assumed to be equal to the number of coding patterns.
- The CSI sensing scheme can be rewritten in a random projection scheme as



$$\mathbf{Y} = \boldsymbol{\Phi}\mathbf{F}$$

HS feature extraction



- CSI hyperspectral measurements

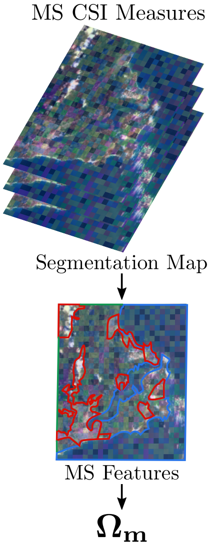
$$\mathbf{Y}_{\mathbf{h}} = \boldsymbol{\Phi}_{\mathbf{h}} \mathbf{F}_{\mathbf{h}},$$

  where $\boldsymbol{\Phi}_{\mathbf{h}} \in \mathbb{R}^{S_h \times L}$ is the coding pattern matrix and S_h is the number of measurement shots acquired with the HS CSI sensor.

- Obtain the feature matrix $\boldsymbol{\Omega}_{\mathbf{h}}$ from the extrapolation process expressed as

$$\boldsymbol{\omega}_{\mathbf{h}}^{j} = \mathbf{y}_{\mathbf{h}}^{\left(\lfloor \frac{j'}{p} \rfloor + \frac{M}{p}\left[\lfloor \frac{j'}{M} \rfloor + \lfloor \frac{j'}{Mp} \rfloor\right]\right)},$$

  where $\boldsymbol{\omega}_{\mathbf{h}}^{j}$ is the j-th column of $\boldsymbol{\Omega}_{\mathbf{h}}$; j and j' are the column indexes of $\boldsymbol{\Omega}_{\mathbf{h}}$ and $\mathbf{Y}_{\mathbf{h}}$, respectively.



- CSI multispectral measurements

$$\mathbf{Y_m} = \mathbf{\Phi_m} \mathbf{F_m},$$

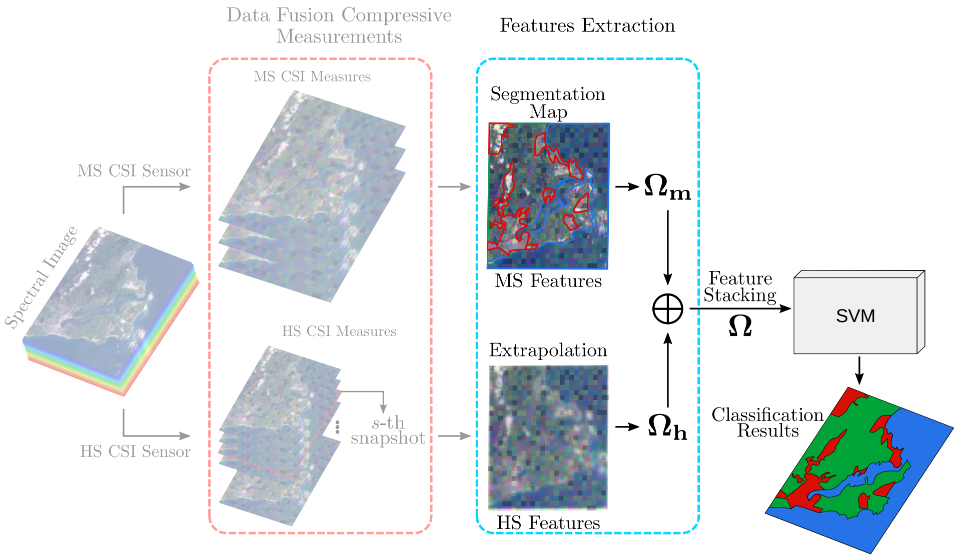
where $\mathbf{\Phi_m} \in \mathbb{R}^{S_m \times L}$ is the coding pattern matrix and S_m is the number of measurement shots acquired with the MS CSI sensor.

- Obtain the segmentation map by applying the SLIC[2]algorithm on the MS compressed measurements.
- Using the segmentation map, the feature matrix $\mathbf{\Omega_m}$ is obtained as follows

$$\boldsymbol{\omega}_{\mathbf{m}}^{\mathbf{p}^e} = \frac{\sum_{l=0}^{n_e-1} \mathbf{Y_m}^{(\mathbf{p}^e)_l}}{n_e}.$$

[1] R Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua & S. Süsstrunk. *SLIC superpixels compared to state-of-the-art superpixel methods. IEEE transactions on pattern analysis and machine intelligence, 2012.*

Feature Stacking

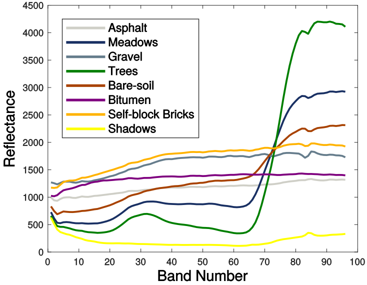


Simulations and Results

Pavia University Dataset

False Color

Ground truth



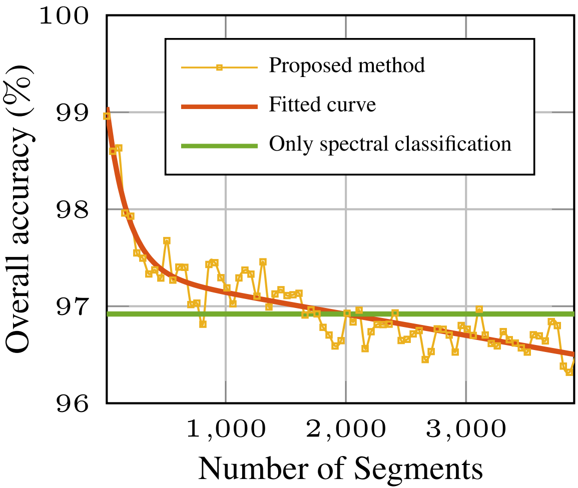
Spectral curves

- The presented results are the average of 10 trials.
- Training rate: select 10% of the pixels from each class.

- CSI compression ratio:

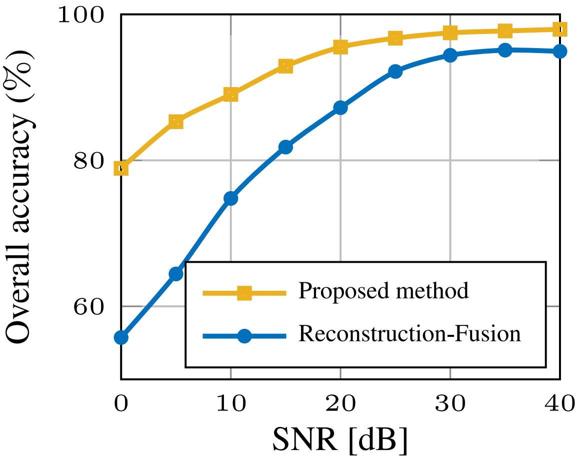
$$\rho = \frac{S_h + S_m}{L},$$

in the experiments we set $\rho = 25\%$.



The number of segments is fixed to $N_{seg} = 10$ for the subsequent experiments on the Pavia University dataset.

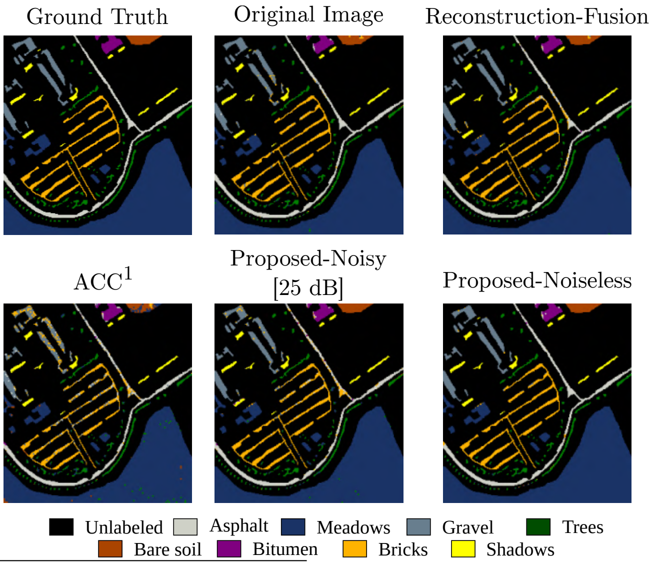
Pavia University



$$\mathbf{Y} = \mathbf{\Phi}\mathbf{F} + \mathbf{\Sigma},$$

where $\mathbf{\Sigma} \sim N(0, \sigma^2)$ represents the noise of the system.

Pavia University



[1] J. Hahn, S. Rosenkranz, & A. M. Zoubir. Adaptive Compressed Classification for Hyperspectral Imagery. *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). pp. 1020-1024. 2014.*

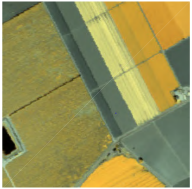
Table 1: Performance of the various classification approaches on the Pavia University dataset.

Class	Original image	Recontruction-Fusion	ACC Framework	Proposed-Noisy	Proposed-Noiseless
Asphalt	86.80 ± 2.03	84.62 ± 1.05	91.20 ± 1.21	95.05 ± 4.62	**98.63 ± 0.70**
Meadows	99.07 ± 0.02	99.23 ± 0.22	95.78 ± 0.14	98.95 ± 0.17	**99.77 ± 0.02**
Gravel	82.39 ± 1.35	80.03 ± 6.67	79.62 ± 0.31	78.04 ± 4.58	**99.67 ± 0.16**
Trees	88.61 ± 2.41	91.55 ± 2.62	92.06 ± 0.27	86.86 ± 3.13	**93.35 ± 0.07**
Bare-Soil	61.96 ± 5.89	72.45 ± 5.89	85.57 ± 0.98	88.98 ± 6.46	**98.25 ± 2.47**
Bitumen	**93.29 ± 0.97**	90.82 ± 2.51	77.11 ± 0.16	93.70 ± 3.10	92.19 ± 0.97
Self-Block Bricks	90.40 ± 0.20	85.14 ± 3.19	83.16 ± 0.24	83.05 ± 1.19	**97.58 ± 1.03**
Shadows	**100.00 ± 0.00**	99.89 ± 0.15	98.47 ± 0.66	98.42 ± 0.74	98.74 ± 0.00
OA (%)	94.51 ± 0.35	94.05 ± 0.72	90.88 ± 0.43	94.55 ± 0.60	**98.90 ± 0.03**
AA (%)	87.81 ± 1.26	87.97 ± 0.01	87.87 ± 1.05	90.38 ± 0.86	**97.27 ± 0.40**
κ	0.91 ± 0.0062	0.90 ± 0.0119	0.88 ± 0.0147	0.91 ± 0.0105	**0.98 ± 0.0005**
Time (s)	1.17 ± 0.007	87.43 ± 1.77	24.97 ± 2.35	**0.66 ± 0.050**	0.74 ± 0.037

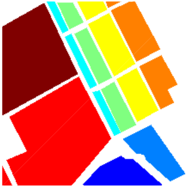
Salinas Valley

Salinas Valley Dataset

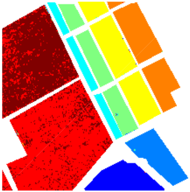
False-Color Image



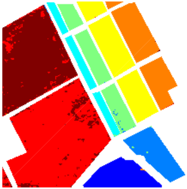
Ground Truth



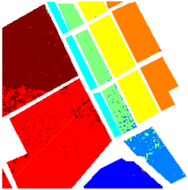
Full Image



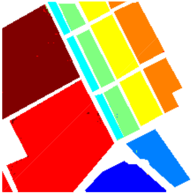
Reconstruction Fusion



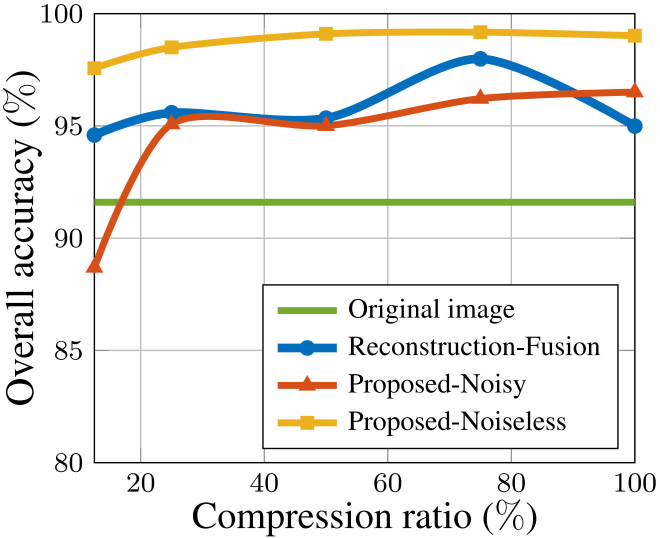
Proposed-Noisy [25 dB]



Proposed Noiseless

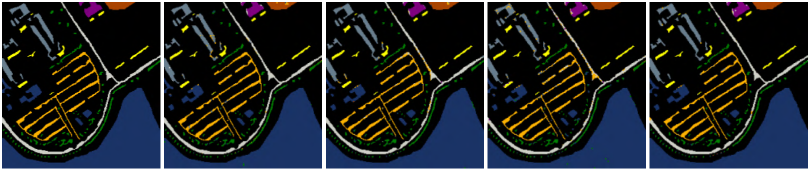


Salinas Valley



Conclusions

- The proposed method incorporates spatial neighboring information by using the superpixel technique.
- Features are extracted from the HS CSI measures using an extrapolation procedure.
- In general, the results show that performing the classification directly with the compressive measurements provides similar accuracy results.
- A maximum difference of just 3% in terms of OA was observed when comparing the classification results obtained by the full 3D data with those achieved using the CSI data fusion measurements.



Thank you! 謝謝！

ICIP 2019

IEEE International Conference on Image Processing

22-25 September 2019

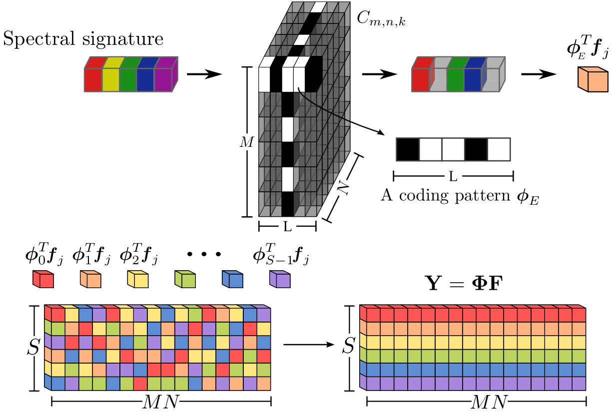
Taipei . Taiwan

Universidad Industrial de Santander

High-Dimensional Signal Processing Research Group

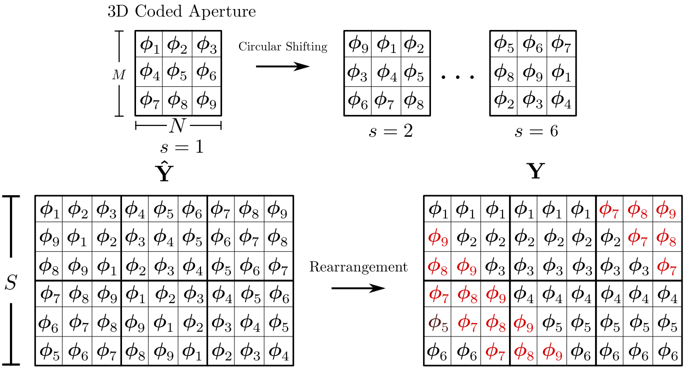
www.hdspgroup.com

Questions?



Rearrangement of the matrix $\mathbf{Y}$ such that the s-th row contains the compressed measurements acquired with the s-th coding pattern ϕ_s. In this figure, colors represent a specific codification, e.g, red pixels denotes the compressed measurements acquired with the ϕ_0 coding pattern.

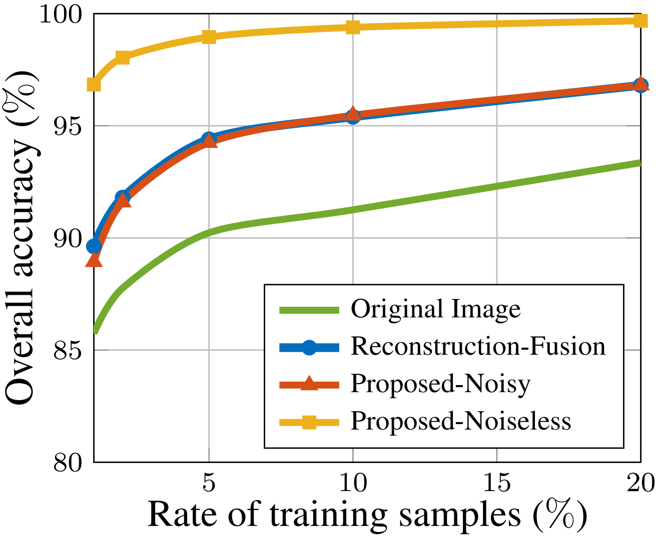
$S = 6$
$P = 9$
$N = M = 3$



Ideas

- Replace the missplaced measurement with the most correlated one.
- Put 0 and use matrix completion.

Salinas Valley



SLIC algorithm

- Algorithm works in the 5-D $[labxy]$ space, where $[lab]$ is the pixel color vector in CIELAB color space, and xy is the pixel position.
- Given K desired equally-sized superpixels. The approximate size of each superpixel is therefore N/K pixels. Then, there would be a superpixel center at every grid interval $S = \sqrt{N/K}$.
- The algorithm choose K superpixel cluster centers $C_k = [l_k, a_k, b_k, x_k, y_k]^T$ with $k = [1, K]$ at regular grid intervals S. The search are would be $2S \times 2S$.
- Distance measure D_s defined as

$$
\begin{aligned}
d_{lab} &= \sqrt{(l_k - l_i)^2 + (a_k - a_i)^2 + (b_k - b_i)^2} \\
d_{xy} &= \sqrt{(x_k - x_i)^2 + (y_k - y_i)^2} \\
D_s &= d_{lab} + \frac{m}{S} d_{xy}
\end{aligned}
$$

m controls the compactness of a superpixel. It is usually chosen as $m = 10$.

- Image gradients are computed as:

$$G(x,y) = \|\mathbf{I}(x+1,y) - \mathbf{I}(x-1,y)\|^2 + \|\mathbf{I}(x,y+1) - \mathbf{I}(x,y-1)\|^2,$$

where $\mathbf{I}(x,y)$ is the lab vector corresponding to the pixel at position (x,y), and $\|\cdot\|$ is the L_2 norm.

Algorithm 1 Efficient superpixel segmentation

```
Initialize cluster centers C_k = [l_k, a_k, b_k, x_k, y_k]^T by sampling pixels at regular grid
  steps S.
Perturb cluster centers in an n × n neighborhood, to the lowest gradient position.
repeat
    for each cluster center C_k do
        Assign the best matching pixels from a 2S × 2S square neighborhood around
          the cluster center according to the distance measure (Eq. 1).
    end for
    Compute new cluster centers and residual error E {L1 distance between previous
      centers and recomputed centers}
until E ≤ threshold
Enforce connectivity.
```

SLIC complexity

The time complexity for the classical *k-means* algorithm is $O(NKI)$ where N is the number of data points (pixels in the image), K is the number of clusters, and I is the number of iterations required for convergence.

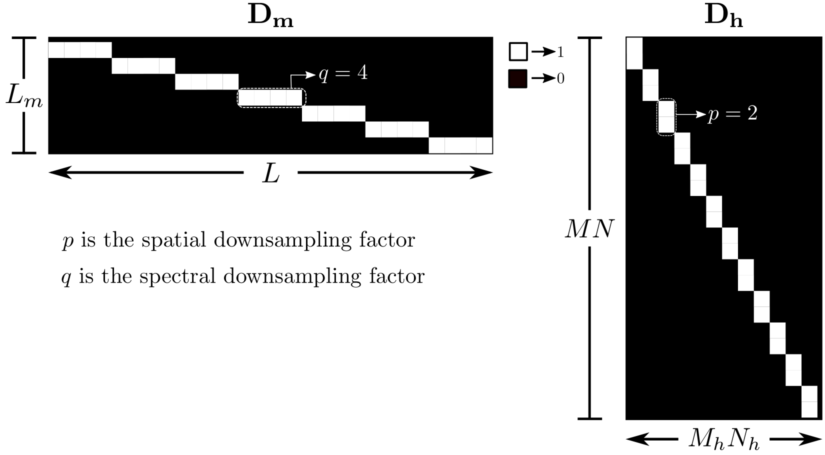
The complexity of SLIC algorithm is $O(N)$, where N is the total number of pixels, since it needs to compute distances from any point to no more than eight cluster centers and the number of iterations is constant.

Superpixels to extract MS compressed features

- Create the segmentation map using SLIC algorithm.
- Obtain all the vectors $\mathbf{p}^e$ containing the indexes of all pixels belonging to the superpixel e.
- The columns of the MS feature matrix $\mathbf{\Omega_m}$ are created as

$$\boldsymbol{\omega}_{\mathbf{m}}^{\mathbf{p}^e} = \frac{\sum_{l=0}^{n_e-1} \mathbf{Y_m}^{(\mathbf{p}^e)_l}}{n_e},$$

  where N_{seg} is the number of segments generated by the superpixel algorithm, $(\mathbf{p}^e)_l$ denotes the l-th entry of the $\mathbf{p}^e$ vector and $\boldsymbol{\omega}_{\mathbf{m}}^{\mathbf{p}^e}$ represents the columns in $\mathbf{\Omega_m}$ indexed by the vector $\mathbf{p}^e$.
- Note that the above equation simply replace all vectors in a segment e by its mean spectral pixel. This procedure incorporates the spatial neighboring information of the superpixel in the classification method.



p is the spatial downsampling factor

q is the spectral downsampling factor

Accuracy Metrics

		Reference Data			
		Water	Forest	Urban	Total
Classified Data	Water	21	6	0	**27**
	Forest	5	31	1	**37**
	Urban	7	2	22	**31**
	Total	**33**	**39**	**23**	***95***

- **Overall Accuracy** $=$ #corrected classified site/Total number of reference site $= 21 + 31 + 22/95 = 77.9\%$.
- **Average Accuracy** $=$ is the average of each accuracy per class. sum of accuracy for each class predicted/# of classes
- **Producer's Accuracy** $=$ correctly classified reference sites/total # of reference site. For instance (Water): $21/33 = 64\%$.
- **Kappa Coefficient**: essentially evaluate how well the classification performed as compared to just randomly assigning values, i.e. did the classification do better than random.