Feature Mining for Effective Subspace Detection of Hyperspectral Image Classification

Presented By
Md. Tanvir Ahmed
Roll: 123086

Supervised by
Dr. Md. Ali Hossain
Assistant Professor
Department of CSE

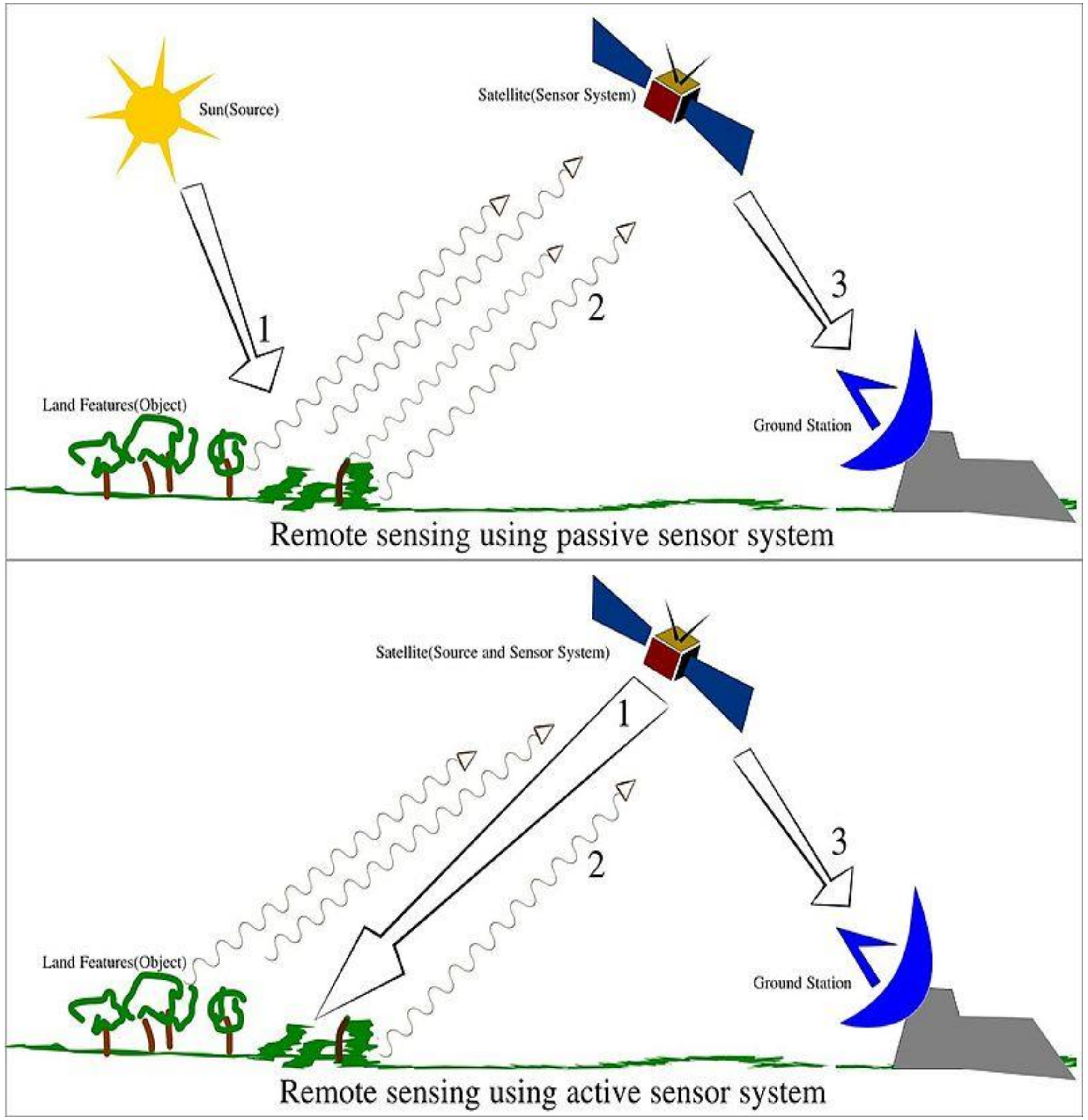
August 26, 2017

Outlines

- Remote Sensing
- Hyperspectral Images
- Objectives
- Proposed Method
- Feature Extraction
- Feature Selection
- Classification
- Conclusion
- Future Work
- References

- ❖ What is remote sensing?
  - ❖ Acquisition of information about an object without making physical contact with the object
  - ❖ Used in geography, land surveying, military, intelligence and planning

- ❑ Types:
  - ❑ Passive Remote Sensing
  - ❑ Active Remote Sensing



Remote Sensing(Con't)

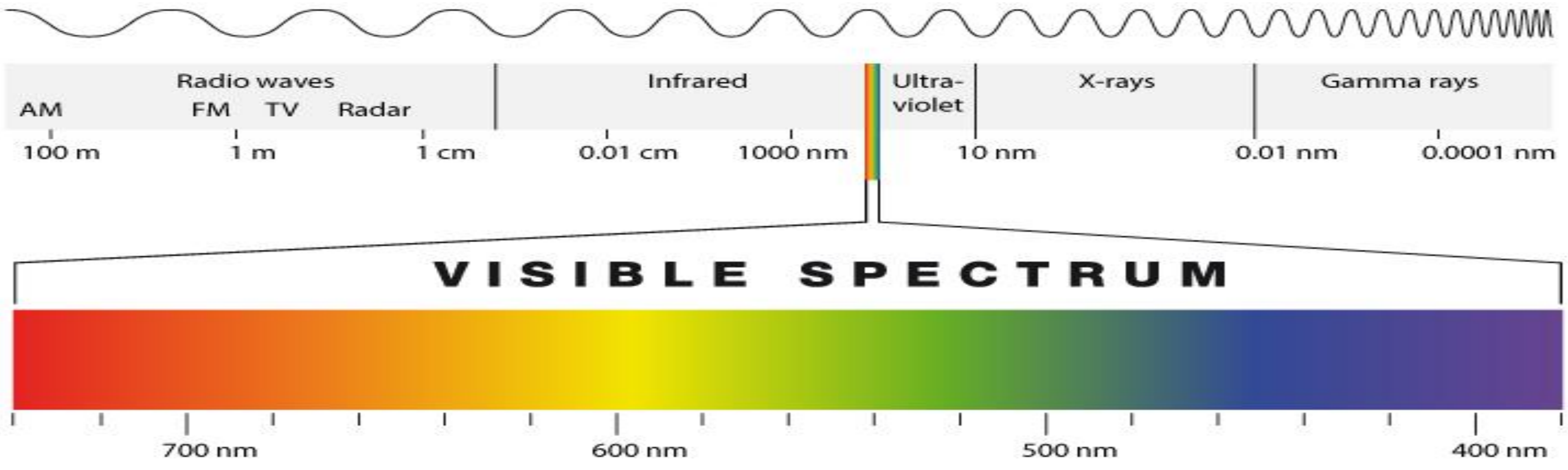


Figure: Electromagnetic Spectrum

Figure: Surveillance aircraft

- ❖ Visible and infrared range of wavelengths represents only part of the story in remote sensing
- ❖ Can also image the earth in microwave range

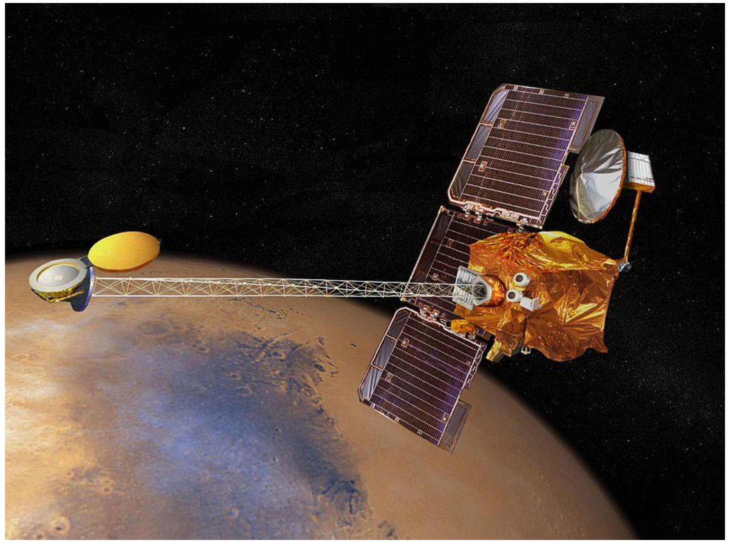


Figure: Remote sensing satellite

Hyperspectral Images

Figure: Two dimensional projection of Hyperspectral cube [5]

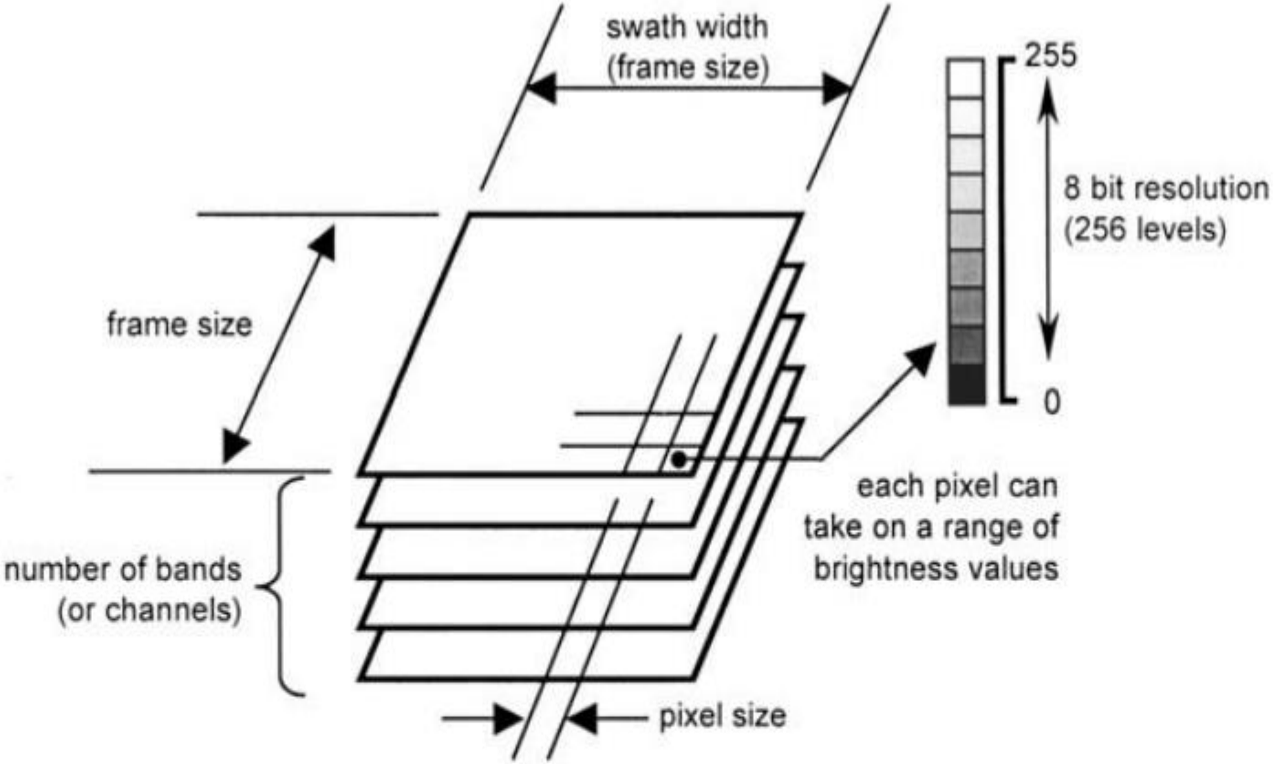


Figure: Technical characteristics of digital image data [2]

Challenges with Hyperspectral Images

- ❖ Curse of dimensionality
- ❖ High correlation among data
- ❖ All the feature are not equally important

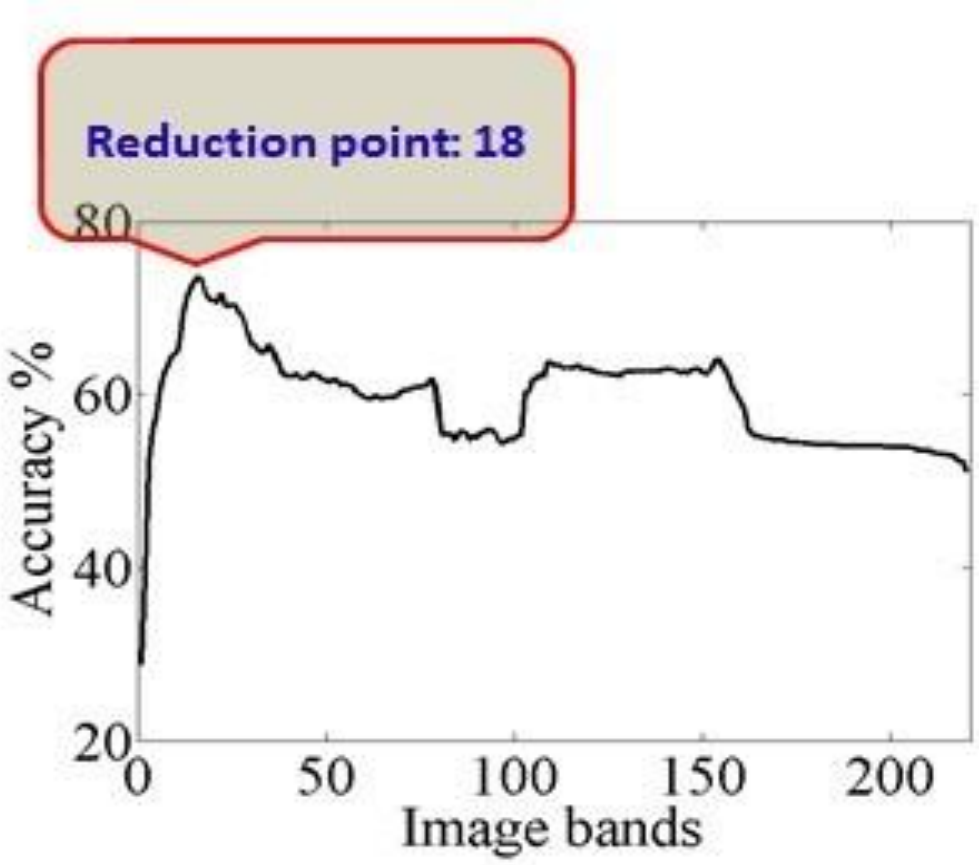


Figure: Curse of dimensionality [4]

Motivations and Objectives

- Extract only relevant features
- Address curse of dimensionality
- Improve classification accuracy

Proposed Method

Collaboration of feature mining through feature extraction and feature selection

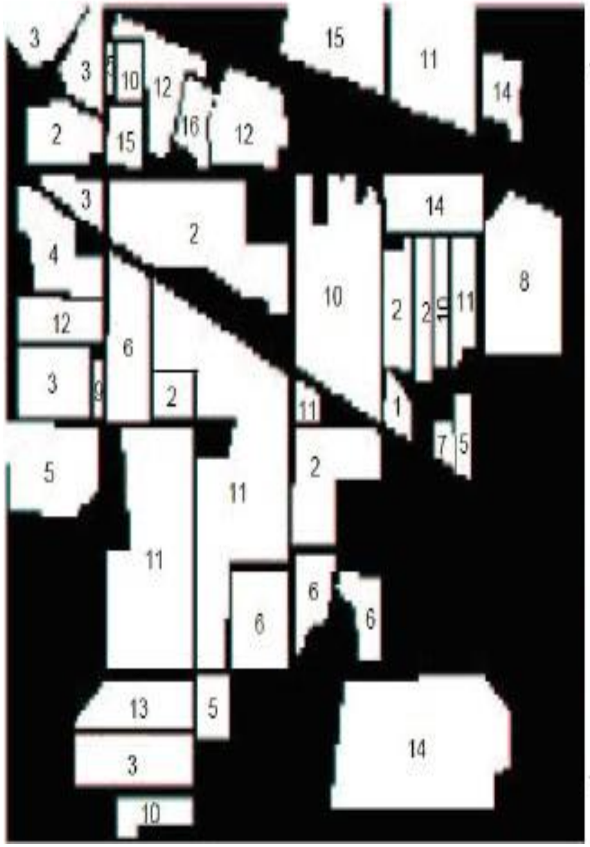
- **Feature Extraction:**
  Principal Component Analysis(PCA)
- **Feature Selection:**
  Normalized Mutual Information(nMI)
- **Classification:**
  Kernel Support Vector Machine(KSVM)

Dataset Description

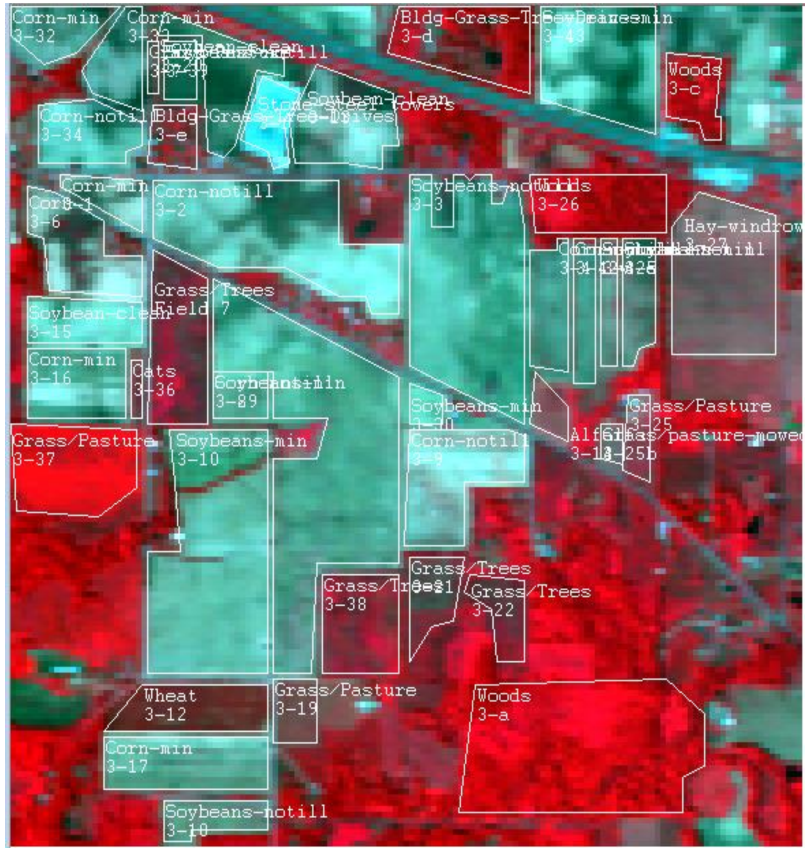
- **Dataset:** AVIRIS 92AV3C
- **Spectral Resolution:**
  0.4μ m-2.5 μ m
- **Spatial Resolution:** 30m
- **Band:** 220
- **Classes:**16

Data Set:
Captured Date:
June 12, 1992 Indian Pine
Test Site 3
Version 1.0

220 Band AVIRIS
Hyperspectral Image
Each band has 145 x 145
=21025 pixel



1. Alfalfa
2. Corn-notill
3. Corn-min
4. Corn
5. Grass/Pasture
6. Grass/Trees
7. Grass/pasture-mowed
8. Hay-windrowed
9. Oats
10. Soybeans-notill
11. Soybeans-min
12. Soybean-clean
13. Wheat
14. Woods
15. Bldg-Grass-Tree-Drives
16. Stone-steel towers



Feature Extraction

- Feature means unique measurable characteristics of an image/observed object (e.g., each band)
- Feature extraction means creating new features which have the essence of original features.
- Feature extraction helps to obtain relevant information from large dataset.

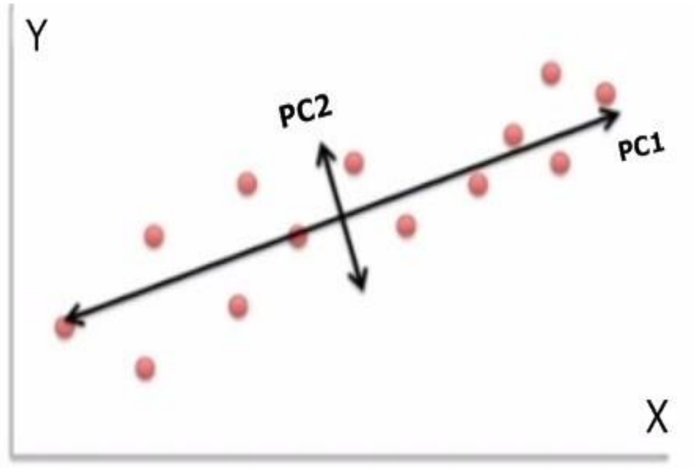
Principal Component Analysis(PCA)

- Perform orthogonal transformation
- Data are ordered based on high variance
- Dimensionality reduced by taking only first few components
- Classification accuracy Increased from the original set

Feature extraction using Principal Component Analysis(PCA)

- Principal Component Analysis (PCA) is a process of reduction input dimensionality. PC2 can be removed

- Transformed data is uncorrelated



Feature extraction using Principal Component Analysis(PCA)(Con't)

All the steps we need to perform Principal Component Analysis[1]:

- **Step 1:** Input the image
- **Step 2:** Subtract the mean
- **Step 3:** Calculate the covariance matrix
- **Step 4:** Calculate the eigenvectors and eigenvalues of the covariance matrix
- **Step 5:** Choose components and forming a feature vector

  FeatureVector = (eig1 eig2 eig3…eign)
- **Step 6:** Derive the new data set

  FinalData = RowFeatureVector x RowDataAdjust

Original Band

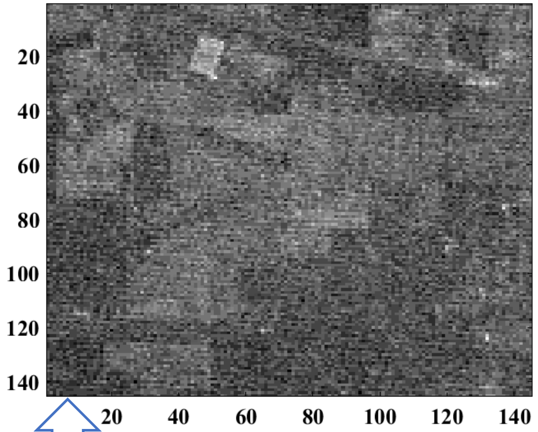
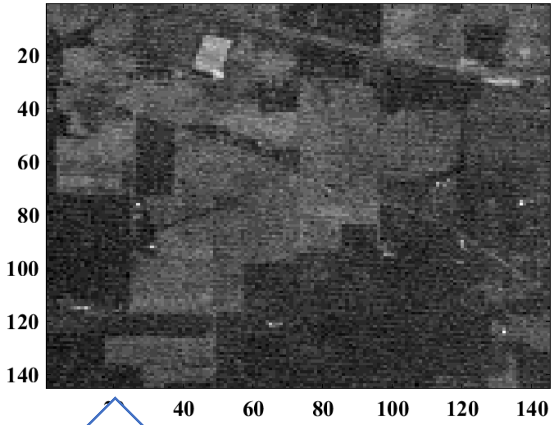
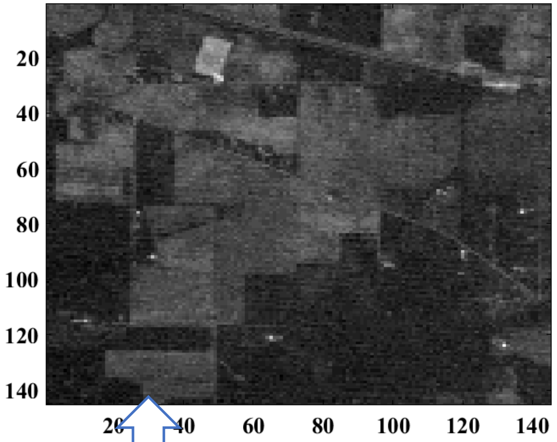


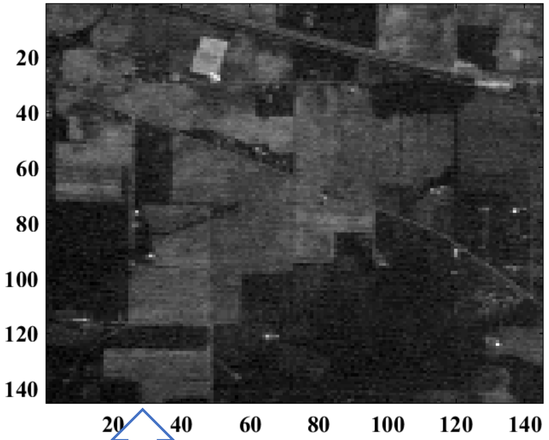
Image Band 1



Band 2

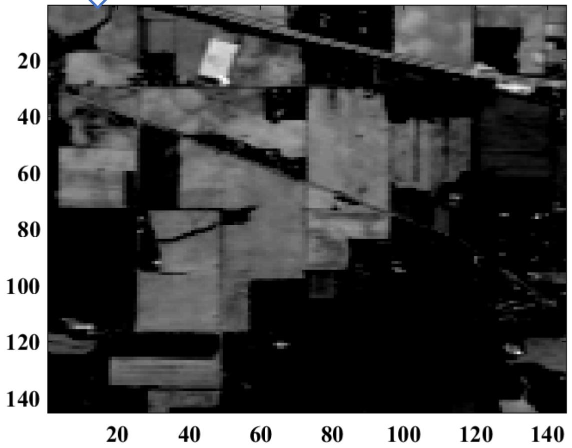


Band 3

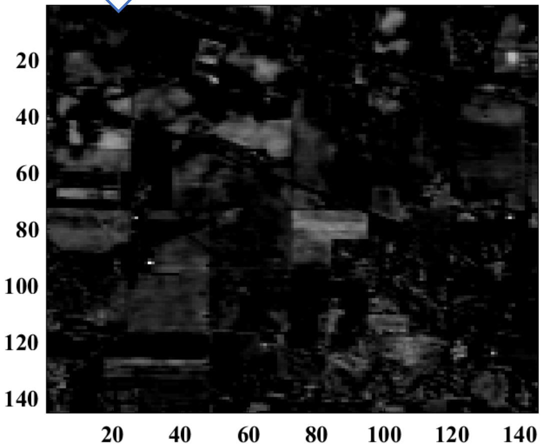


Band 4

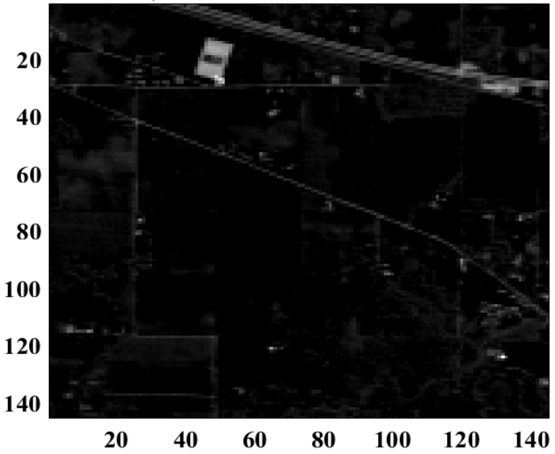
PCA Band



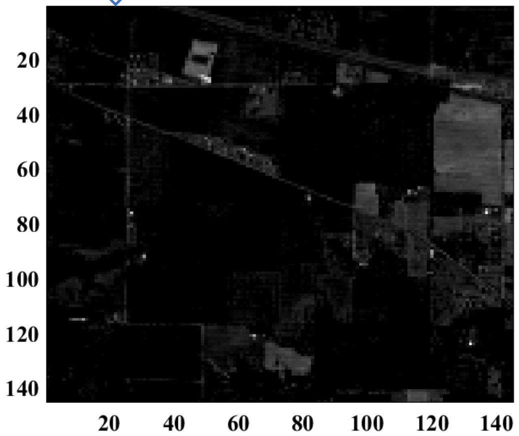
PC 1



PC 2



PC 3



PC 4

Cumulative variance of Transformed data

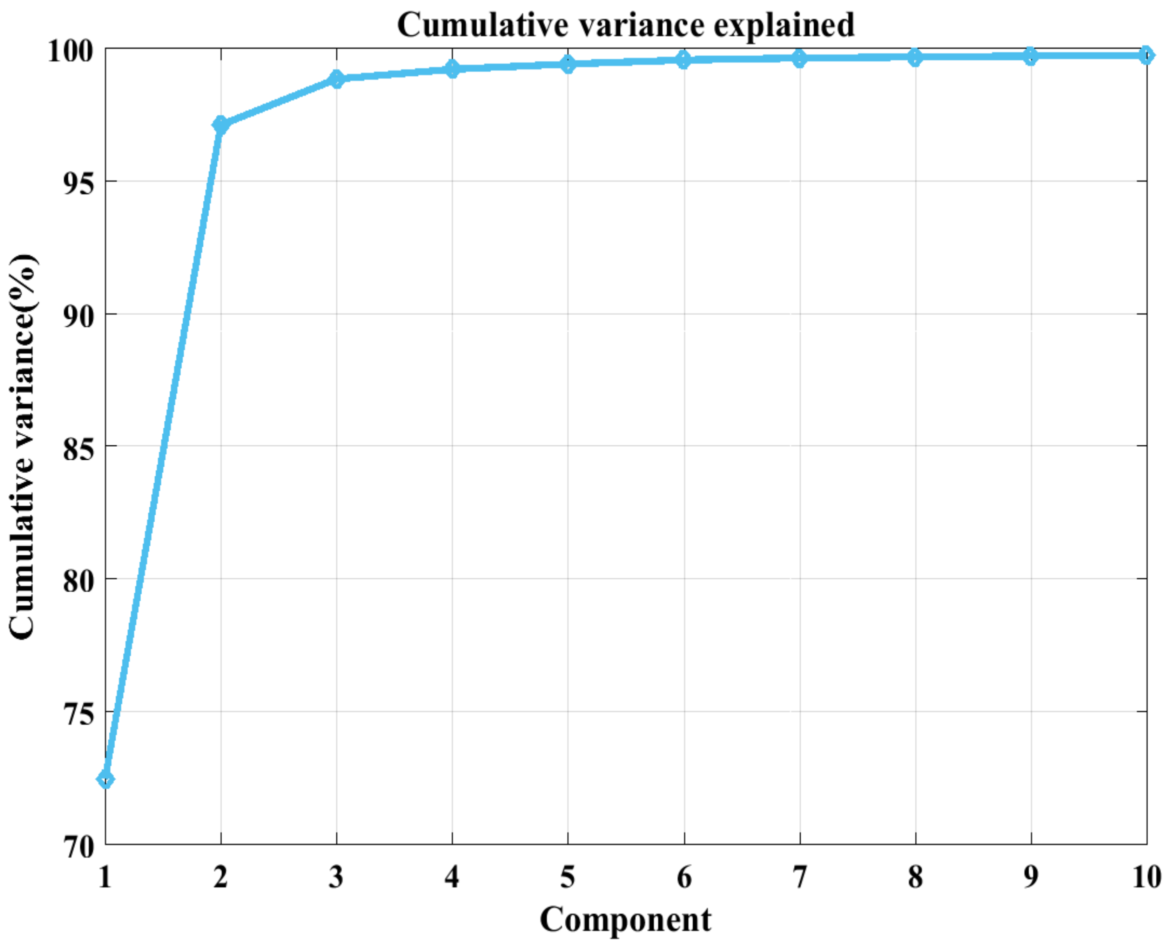
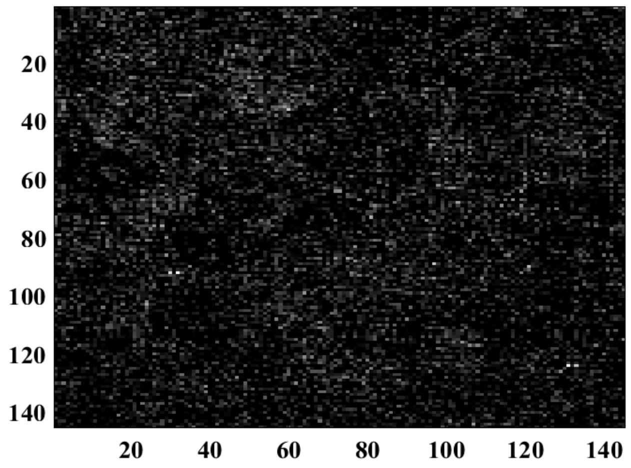


Figure: Cumulative variance

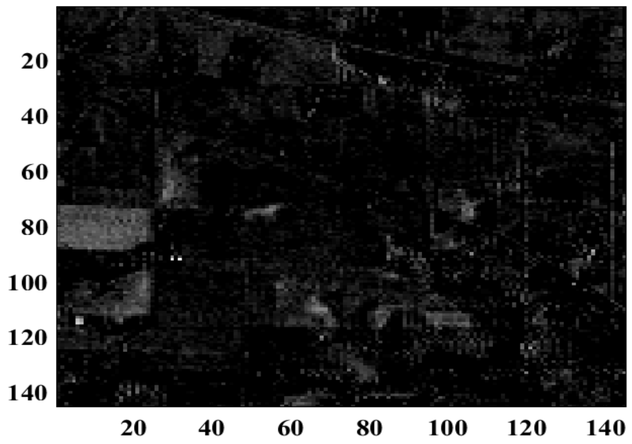
After principal component analysis first **10** components have **99.75%** variance of the total

Limitations of PCA

- Depends on overall variance. Assumes high variance means high information
- Here though PC 8 has high variance but it contains less information of image contents.
- On the other hand PC 9 has low variance but it contains more information of image contents.



PC 8



PC 9

Mutual Information

$$I(\mathbf{X},\mathbf{Y}) = \sum_{y=1}^{L_1}\sum_{x=1}^{L_2} p(x,y)\log\frac{p(x,y)}{p(x)p(y)}$$ [7]

$$I(\mathbf{X}_i,\mathbf{C}) < I(\mathbf{X}_j,\mathbf{C}) \text{ for all } j \neq i, \text{and } i = 1,2,\ldots,N.$$

$$G(\mathbf{X}_i,k)=I(\mathbf{X}_i,\mathbf{C})-\beta \sum_{\mathbf{X}_{sk}\in\mathbf{S}_k} I(\mathbf{X}_i,\mathbf{X}_{sk}), \quad \mathbf{X}_i \notin \mathbf{S}_k.$$

Where $\mathbf{S}_k = \{\mathbf{X}_{s1},\mathbf{X}_{s2},\ldots,\mathbf{X}_{sk}\}$

$$G(\mathbf{X}_i,k)=I(\mathbf{X}_i,\mathbf{C})-\frac{1}{k} \sum_{\mathbf{X}_{sk}\in\mathbf{S}_k} I(\mathbf{X}_i,\mathbf{X}_{sk}), \quad \mathbf{X}_i \notin \mathbf{S}_k.$$

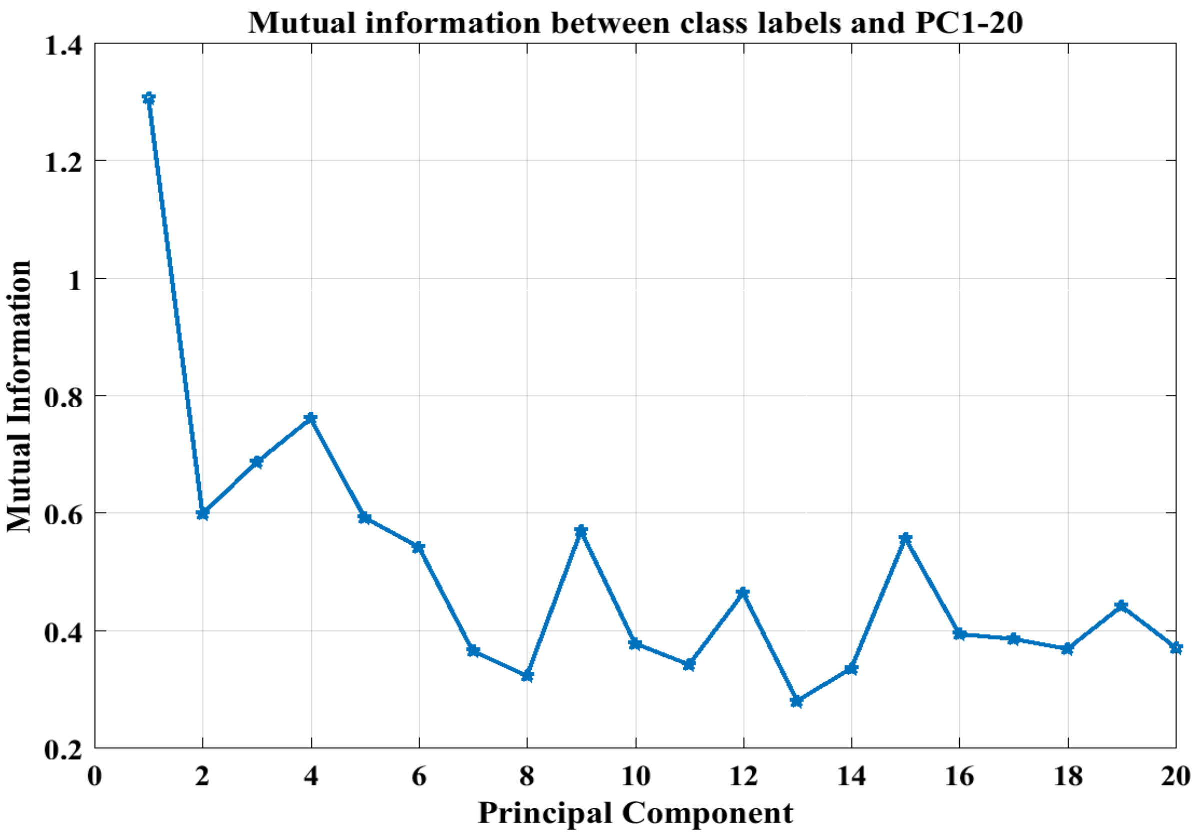


Figure: Mutual information between input classes and features

Normalized Mutual Information

$$\hat{I}(\mathbf{X},\mathbf{C}) = \frac{I(\mathbf{X},\mathbf{C})}{\sqrt{I(\mathbf{X},\mathbf{X})}\sqrt{I(\mathbf{C},\mathbf{C})}}$$ [6]

$$\widehat{G}(\mathbf{X}_i,k)=\widehat{I}(\mathbf{X}_i,\mathbf{C})-\frac{1}{k}\sum_{\mathbf{X}_{sk}\in\mathbf{S}_k}\widehat{I}(\mathbf{X}_i,\mathbf{X}_{sk}),\quad \mathbf{X}_i\notin\mathbf{S}_k.$$

Two Constraints:

$$\text{if } \widehat{I}(\mathbf{X}_i,\mathbf{C}) < T \text{ remove } \mathbf{X}_i$$

$$\widehat{G}(\mathbf{X}_i,k) > 0.$$

PCA	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
PCA-nMI	PC1	PC4	PC3	PC9	PC2	PC6	PC5	PC15	PC19	PC12

Table: Selected Component for PCA and for PCA-nMI

Classification

Kernel support vector machine (SVM)[3] classifier to classify 14 classes

Steps of classification:

Step 1: Get train and test data
Step 2: Scale data set into 0 to 1
Step 3: Calculate best c and g (kernel width)for RBF kernel of SVM

- 10 fold cross validation was used

Step 4: Assess the test sample accuracy

- Radial basis function (RBF) kernel was used

Class Name	Train	Test
Alfalfa	12	12
Corn-notill	42	36
Corn-min	49	42
Corn	30	30
Grass/Pasture	49	42
Grass/Trees	48	48
Hay-windrowed	77	88
Soybeans-notill	117	117
Soybeans-min	120	84
Soybean-clean	54	54
Wheat	54	60
Woods	120	96
Bldg-Grass-Tree-Drives	63	48
Stone-steel towers	25	24

Table: Data set

Conclusion

- 14 class is selected
- Selecting first 9 features after applying PCA-nMI gives 86.04% accuracy

Data set	C	g
Original Data	10	2.85
PCA Data	10	0.83
PCA-nMI Data	10	0.28

Data set	Classification Result (First 10 features only)
Original Data	43.91%
PCA Data	80.66%
PCA-nMI Data	85.53%

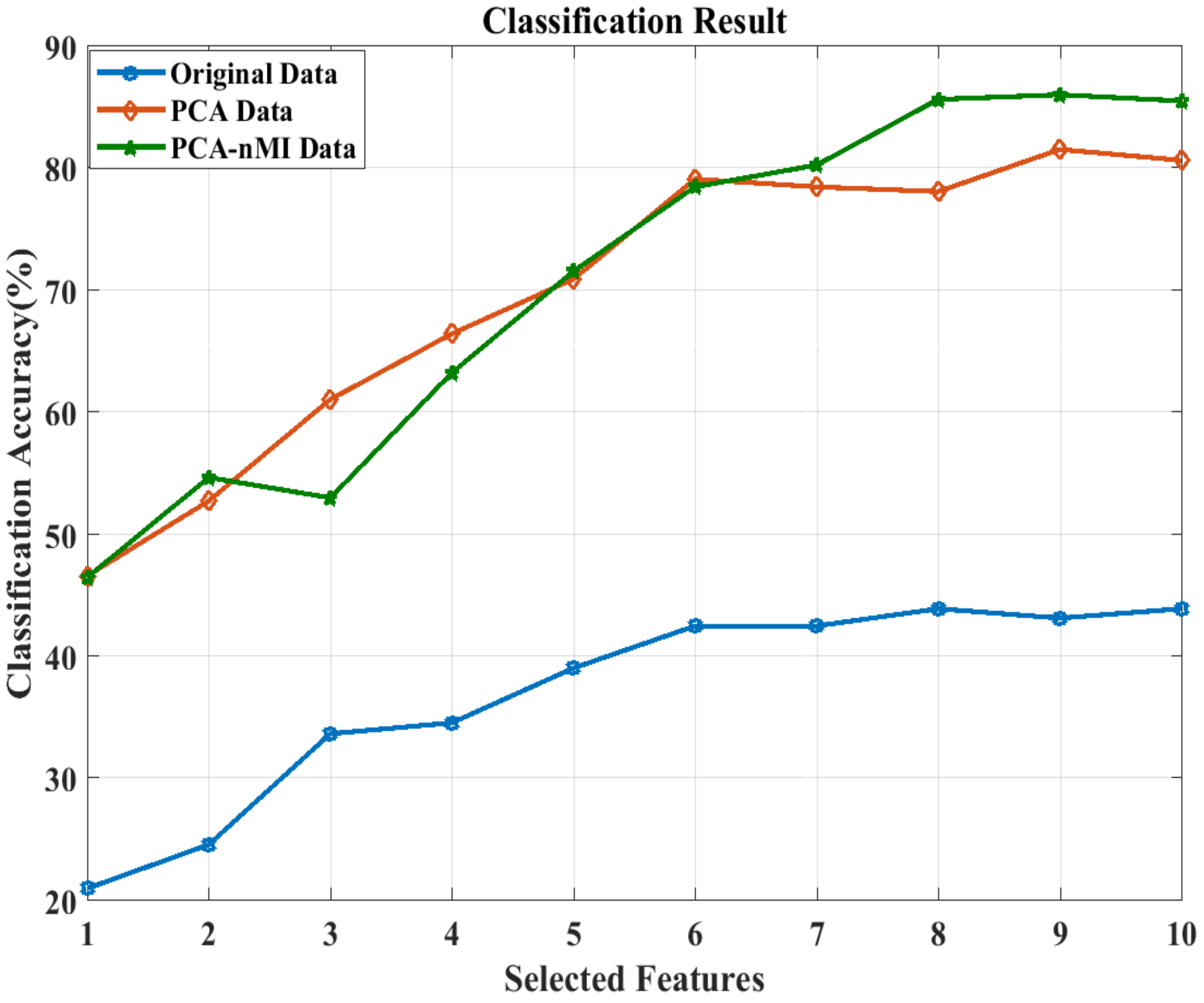


Figure: Comparison of classification result of original data, PCA data and PCA-nMI data.

Future Work

- ❖ Subset features are selected after PCA
- ❖ Selection will be perfomed by introducing adaptive thresholding T.

Acknowledgement

Special thanks to
Professor Dr. David A. Landgrebe, Purdue University
For the dataset and MUltiSpec

References

[1]L. I. Smith, “A tutorial on Principal Component Analysis”, February 26,2002.

[2]J. A. Richards and X. Jia, Remote Sensing Digital Image Analysis 4th edition,
Verlag: Springer 2006.

[3]C. Hsu, C. Chang, and C. Lin,Department of Computer Science,National Taiwan University,
Taipei 106, Taiwan,"A Practical Guide to Support Vector Classication",
Initial version: 2003, Last updated: May 19, 2016.

[4]M. A. Hossain, X. Jia and M. Pickering, “Subspace detection using a mutual-information measure for hyperspectral image classification,” IEEE Geoscience and Remote Sensing Letters,
vol. 11, no. 2, pp. 424-428, Feb. 2014.

[5][Online]https://en.wikipedia.org/wiki/Hyperspectral_imaging

References

[6]M. A. Hossain, X. Jia and M. Pickering, " Improved feature selection based on a mutual information measure for hyperspectral image classification," *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Munich, Germany, pp. 3058-3061, Jul. 2012.

[7]M. A. Hossain, M. Pickering and X. Jia, " Unsupervised feature extraction based on a mutual information measure for hyperspectral image classification," *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*,Vancouver, Canada, pp. 1720-1723, Jul. 2011.

The End

Thank You