PRINCIPAL COMPONENT TRANSFORM

Lecture 6

Multiband Image Operations

- Operations performed by combining gray levels recorded in different bands for the same pixel

- Applications
  - Data reduction through decorrelation
  - Highlighting specific features with significant difference in response in different bands
  - The transformed data may be viewed like enhanced versions compared to originals

Principal Component Transform

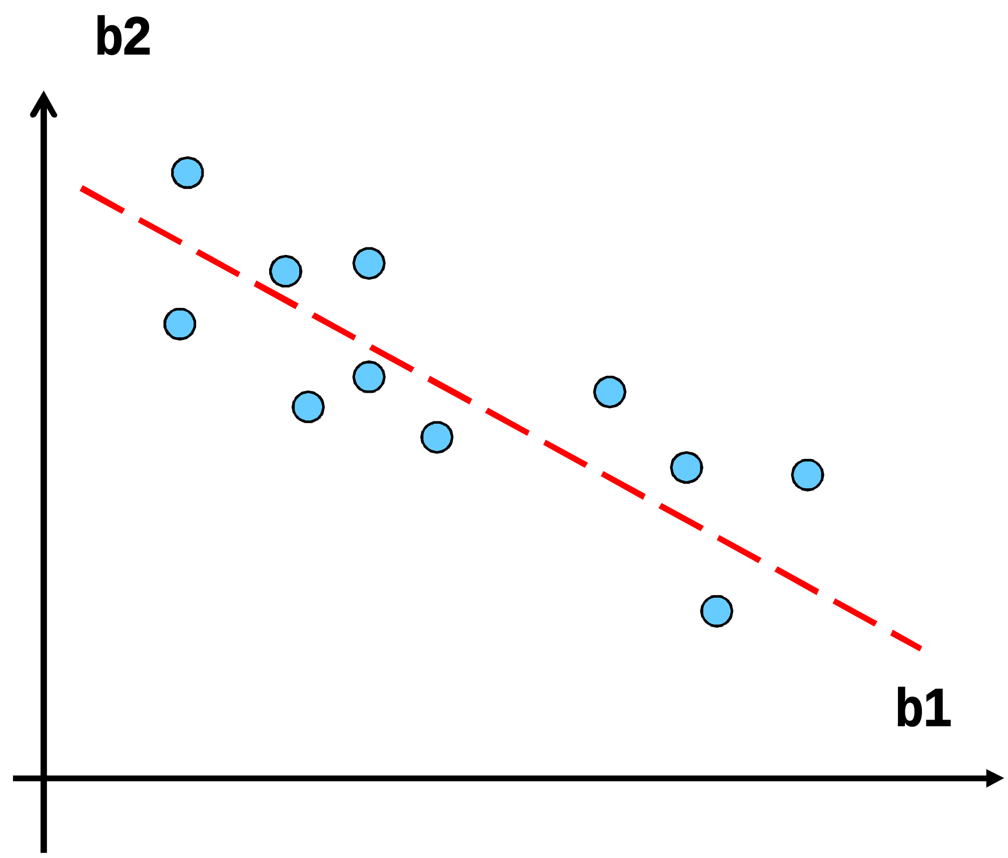
- Highlights the ***redundancy*** in the data sets due to similar response in *some* of the wavelengths

- Original bands variables represented along different coordinate axes, redundancy implies variables are ***correlated***, not independent

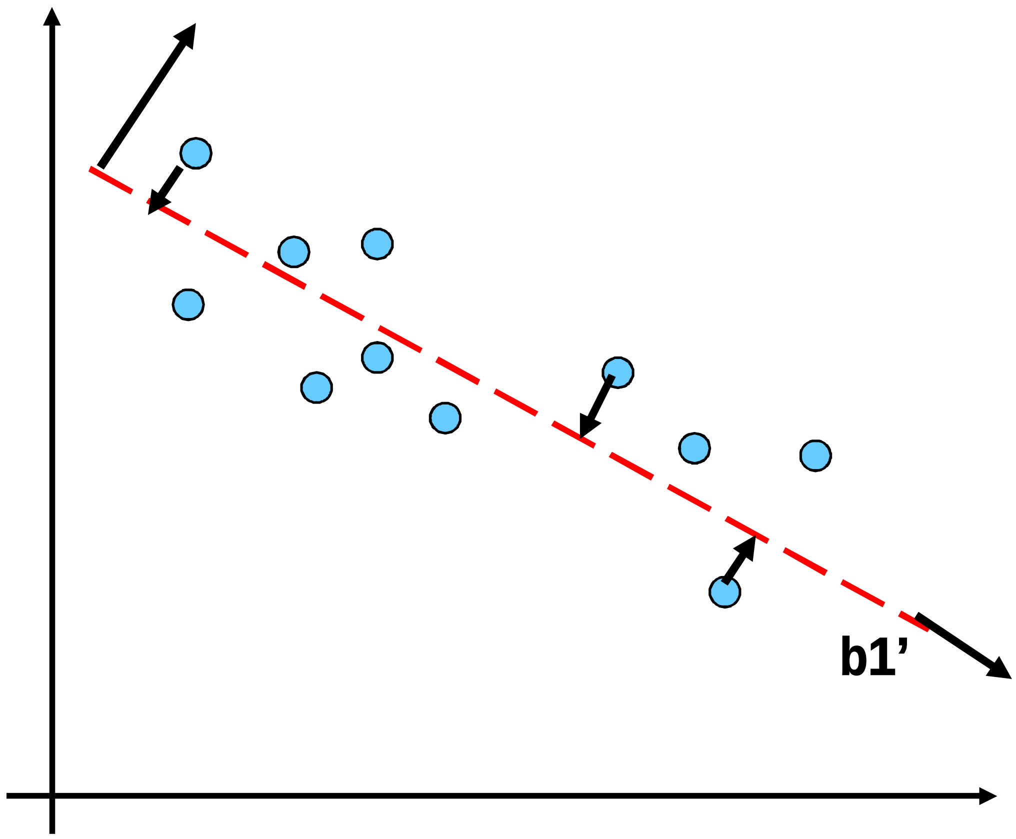
- Gray level in a band at a pixel can be predicted from the knowledge of the pixel gray level in other bands

Example of Redundancy in Data

- **Example:** Highly correlated data
- Values along band b1 leads to knowledge along band b2 of the data element
- Linear variation (nearly) between b1 and b2
- Often true in case of visible bands



- Points projected onto the line ➔a small error in the position of the point.
- Points represented by only one coordinate b1' ➔ half data reduced
- For highly correlated data, this error will be minimal



Decorrelating Multispectral Remotely Sensed Data

- How do we identify the **optimum axes** along which the remotely sensed data should be projected so that the transformed data would be uncorrelated?

- What should be the way to **rank the new axes** so that we can discard the least important dimensions of the transformed data?
- **Invertibility** of the transformation?

Statistics

Mean

$$\frac{\sum_{i=1}^{M}\sum_{j=1}^{N} g_{ij}^{k}}{M.N}$$

Variance

$$\left[\frac{\sum_{i=1}^{M}\sum_{j=1}^{N}(g_{ij}^{k}-\mu_k)^2}{M.N}\right]$$

Covariance

$$\frac{\sum_{i=1}^{M}\sum_{j=1}^{N}(g_{ij}^{k}-\mu_k)(g_{ij}^{l}-\mu_l)}{M.N}$$

Covariance Matrix

- $C = \{C_{kl} \mid k = 1, \ldots, K, l = 1, \ldots, K\}$
- K is the number of bands in which the multispectral dataset was generated
- C is a symmetric matrix
- $C_{kl} = C_{lk}$
- Diagonal elements of C are the intra-band variances
- Off-diagonal elements are the inter-band covariances

Relation between correlation and covariance

- Correlation R_{kl} =

$$\frac{\sum_{i=1}^{M}\sum_{j=1}^{N} g_{ij}^{k} g_{ij}^{l}}{M.N}$$

- It can be shown that $R_{kl} = C_{kl} + \mu_k \mu_l$

- For data with zero-mean, correlation and co-variance will be equal

Principal Component Transformation

Problem to solve:

- Find a transformation to be applied to the input multispectral image such that the covariance matrix of the result is reduced to a diagonal matrix

- Further, we should find an axis $\underline{\boldsymbol{v}}$ such that the variance of the projected coordinates ($z_k = \underline{\boldsymbol{v}_k}^t\, \underline{\mathbf{x}}$) is maximum.

Solving

Given the transformed vector

$$z_k = \underline{\boldsymbol{v}}_{\underline{\boldsymbol{k}}}^{t}\,\underline{\mathbf{x}}$$

The variance σ_z^2 =

$$\frac{\sum_{i=1}^{M}\sum_{j=1}^{N} v^t (x_{ij} - \mu_k)(x_{ij} - \mu_l)^t v}{M.N}$$

This simplifies to $\sigma_z^2 = \underline{\boldsymbol{v}}^t C \underline{v}$

C, the covariance matrix is a positive, semi-definite, real symmetric matrix.

Finding vector v

- To maximize the projected variance σ_z^2, find a $\underline{\boldsymbol{v}}$ such that $\underline{\boldsymbol{v}}^t C \underline{v}$ is maximum, subject to the constraint $\underline{\boldsymbol{v}}^t \underline{v} = 1$. Combining the maximization function with the constraint, we can write
- $\underline{\boldsymbol{v}}^t C \underline{v} - \lambda(\underline{\boldsymbol{v}}^t \underline{v} - 1) = \text{maximum}$
- Differentiating w.r.t. $\underline{\boldsymbol{v}}$,

$$\frac{\partial}{\partial v}\left[\mathbf{v}^t C \mathbf{v} - \lambda(\mathbf{v}^t \mathbf{v} - 1)\right] = 0$$

Eigenvector/Eigenvalue

The derivative results in

$$C\underline{\boldsymbol{v}} = \lambda\underline{\boldsymbol{v}} \text{ (Verify!)}$$

Therefore, $\underline{\boldsymbol{v}}$ is an *eigenvector of* C

$$\underline{\boldsymbol{v}}^t C \underline{v} = \underline{\boldsymbol{v}}^t(\lambda\underline{\boldsymbol{v}}) = \lambda\underline{\boldsymbol{v}}^t\underline{\boldsymbol{v}} = \lambda$$

This implies that $\underline{\boldsymbol{v}}$ is the eigenvector of C with the largest eigenvalue

Therefore all the eigenvectors with decreasing eigenvalues lead to axes with decreasing variance along them.

Example

Covariance Matrix

34.89	55.62	52.87	22.71
55.62	105.95	99.58	43.33
52.87	99.58	104.02	45.80
22.71	43.33	45.80	21.35

Eigenvalues

253.44	7.91	3.96	0.89

Eigenvectors

0.34	−0.61	0.71	−0.06
0.64	−0.40	−0.65	−0.06
0.63	0.57	0.22	0.48
0.28	0.38	0.11	−0.88

Transformation

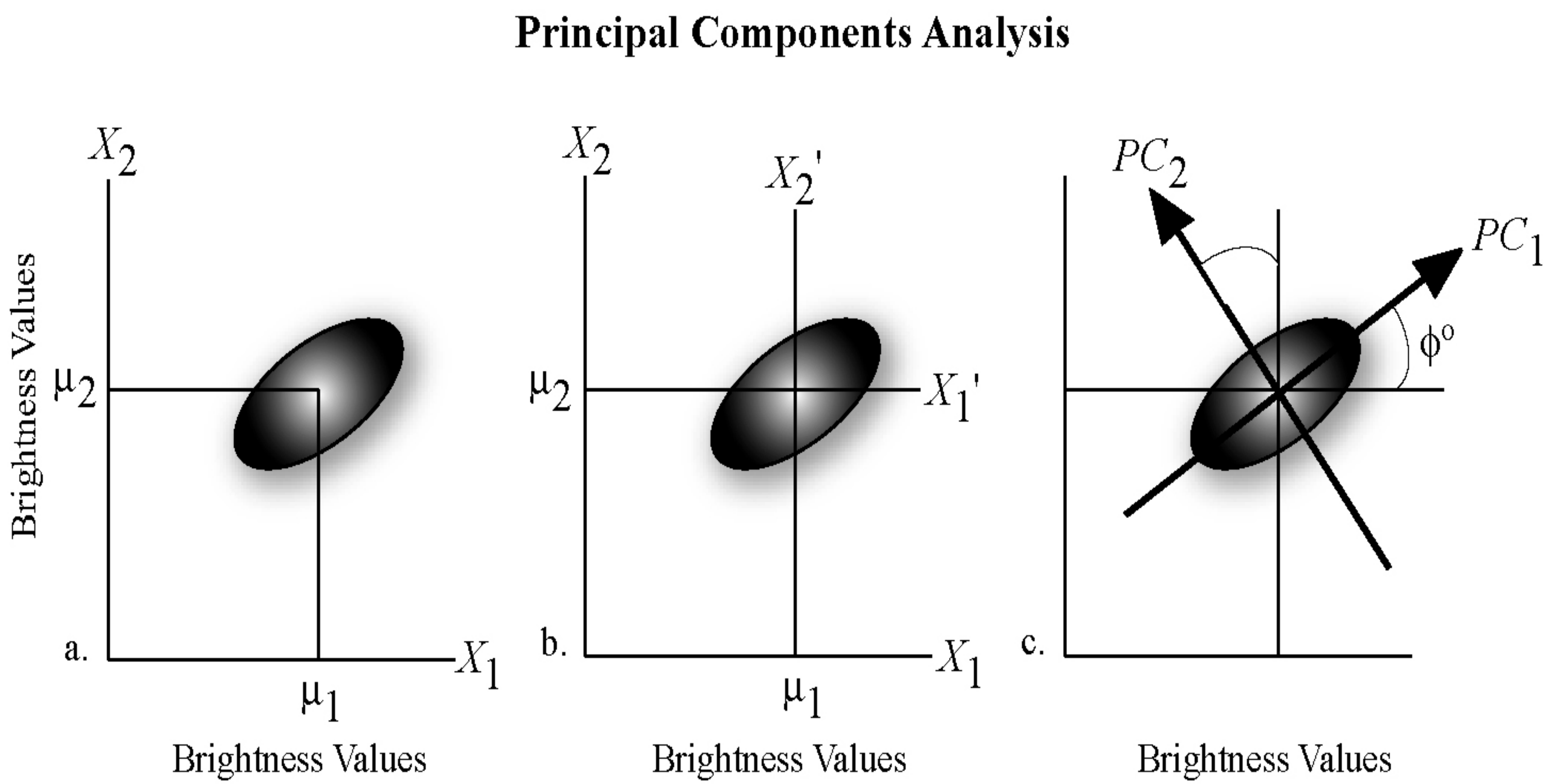
- New component value = dot product of eigenvector and pixel vector
- (i,j) ➔ pixel position
- n eigenvectors for n principal components
- 1st principal component ➔ dot product of pixel vector with eigenvector corresponding to largest eigenvalue

Principal Components

For n input bands, n principal components are computed

The utility of the principal components gradually decreases from 1st towards the last

e.g., For Landsat TM, last three PCs are generally of very little value



From J.R. Jensen's lecture notes at Univ. South Carolina; used with permission

PCT applications

- For IRS / IKONOS images, out of four bands, 2-3 principal components capture most of the useful information. The last 1-2 bands are redundant.

- Advantages
  - Smaller data volume to handle
  - Principal components appear to be enhanced versions of the originals, having contributions from all the four input bands

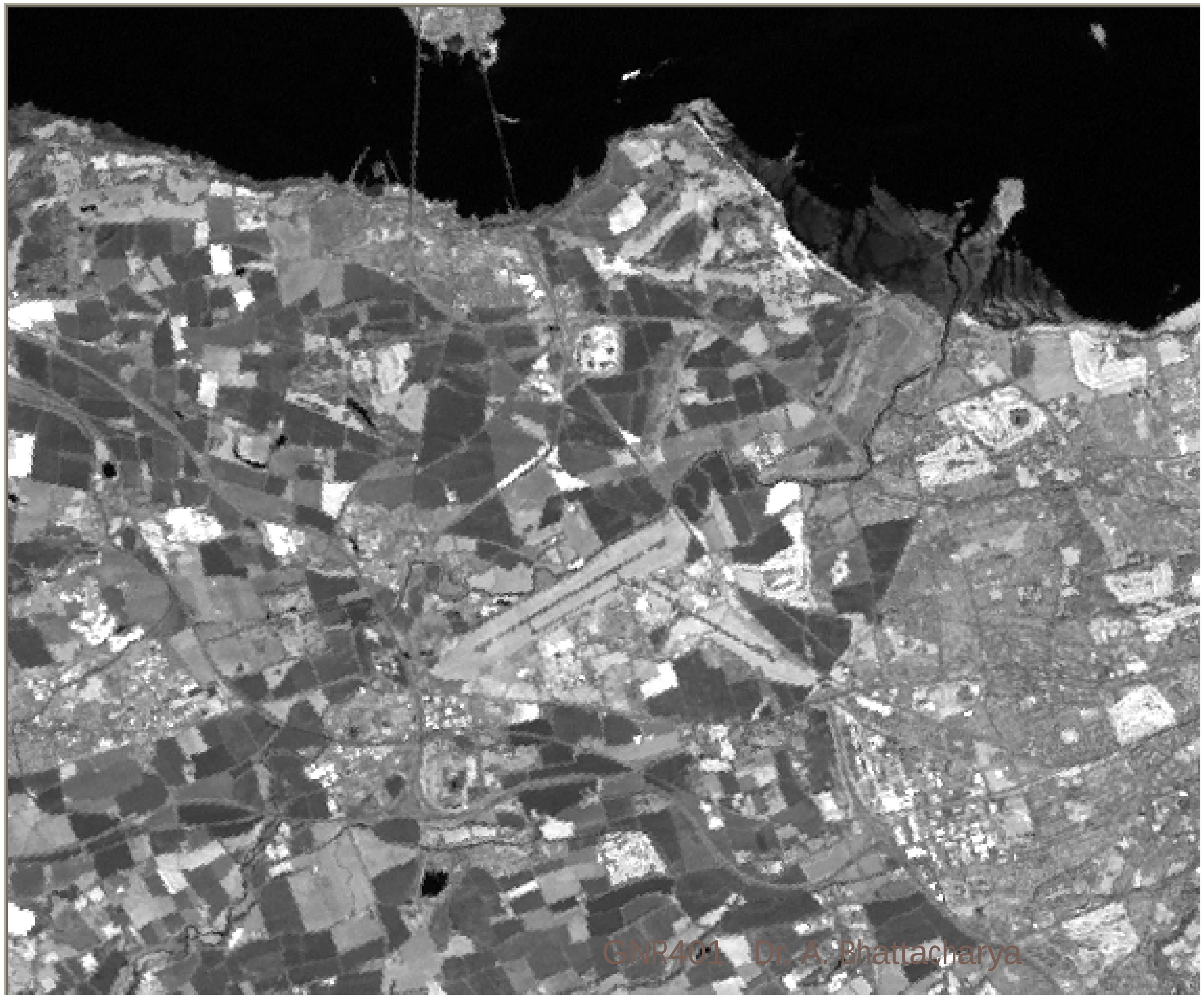
- Application scientists use composites of PC 1-2-3 for interpretation of various features such as geology

Band 1 (Blue)

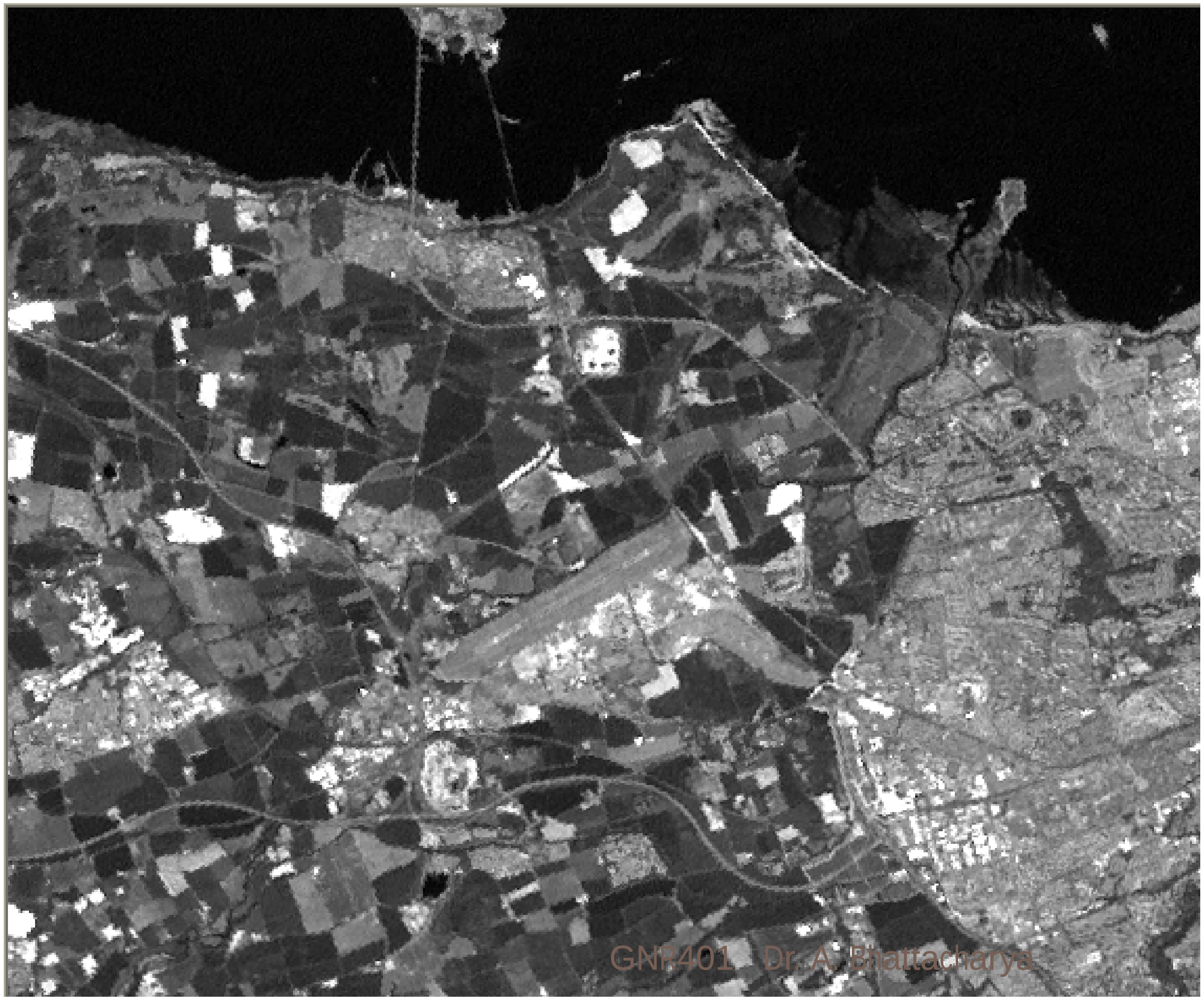
Band 2 (Green)

Band 3 (Red)

Band 4 (NIR)

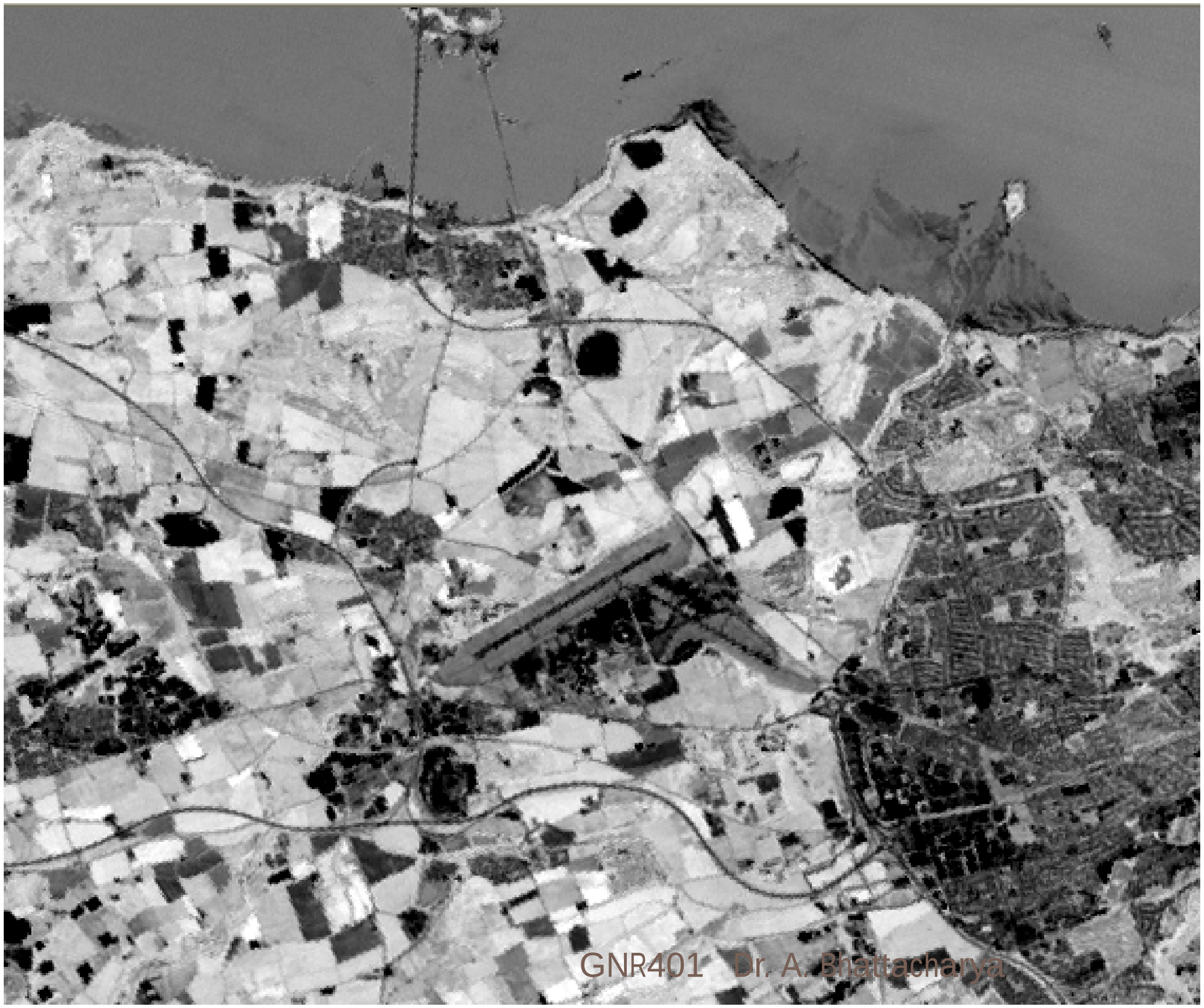


Band 5 (SWIR)



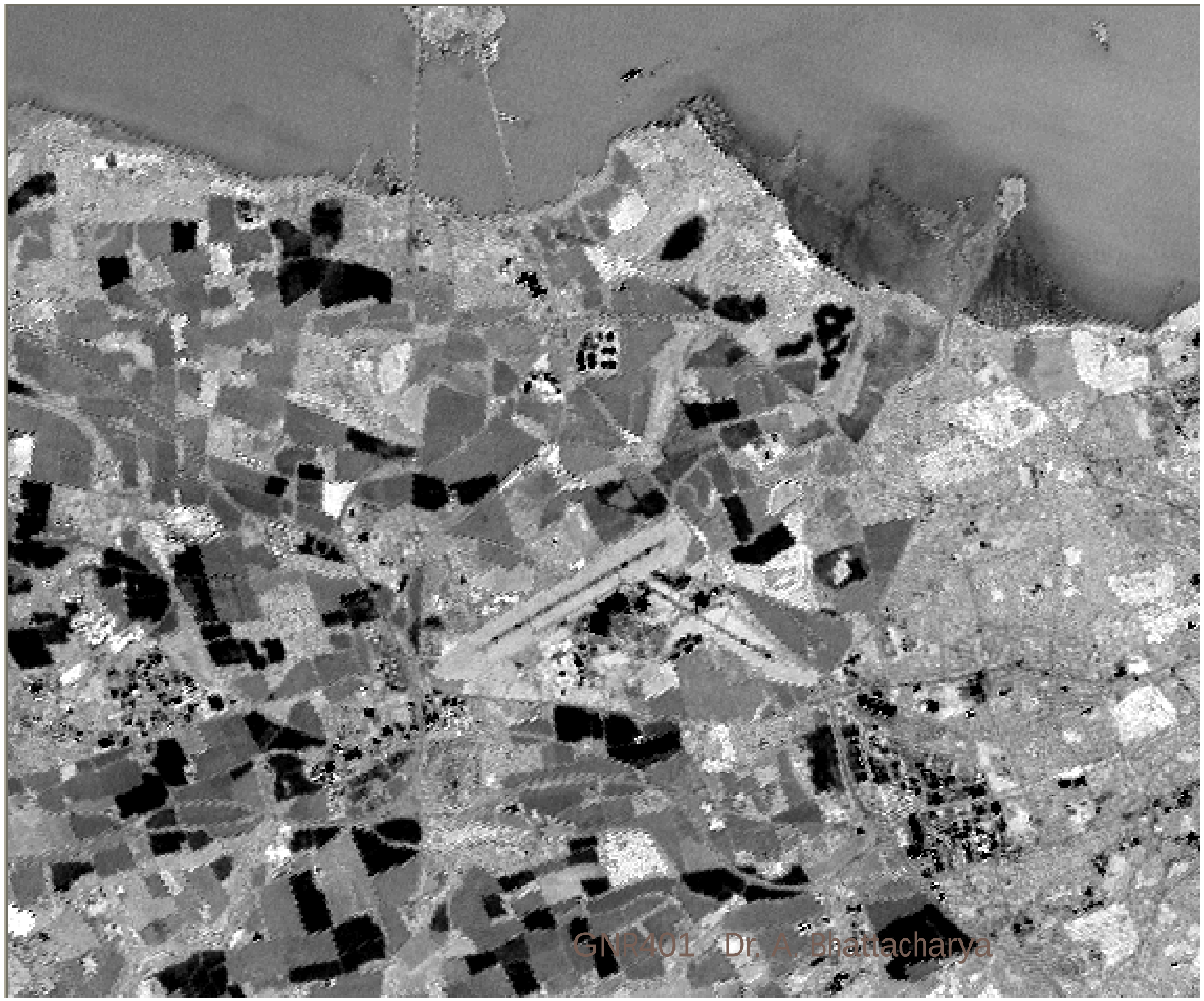
Band 7 (SWIR)

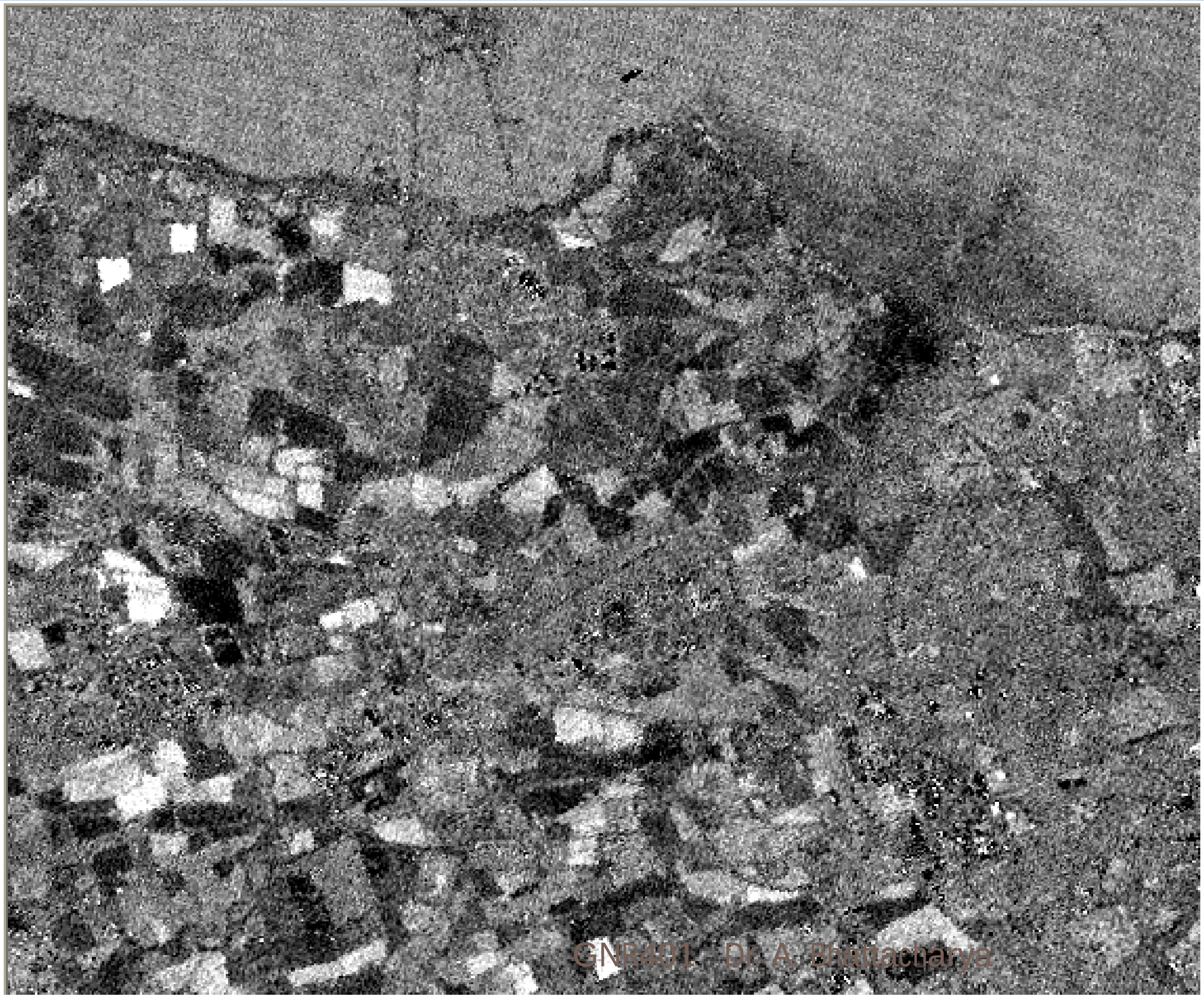
PC1



PC2

PC3





PC6

- Assumption behind PCA is that the data points **x** are multivariate Gaussian
- Often this assumption does not hold
- However, it may still be possible that a transformation $\phi(\mathbf{x})$ is still Gaussian, then we can perform PCA in the space of $\phi(\mathbf{x})$