IMAGE DATA ANALYSIS (6CFU)

MODULE OF

REMOTE SENSING

(9 CFU)

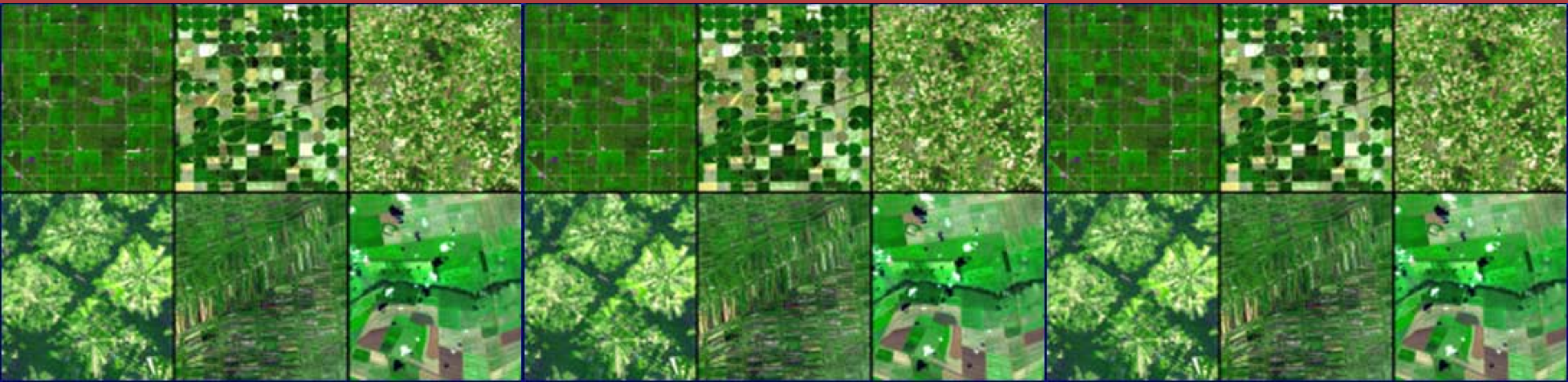
A.Y. 2022/23

MASTER OF SCIENCE IN COMMUNICATION TECHNOLOGIES AND MULTIMEDIA

MASTER OF SCIENCE IN COMPUTER SCIENCE, LM INGEGNERIA INFORMATICA

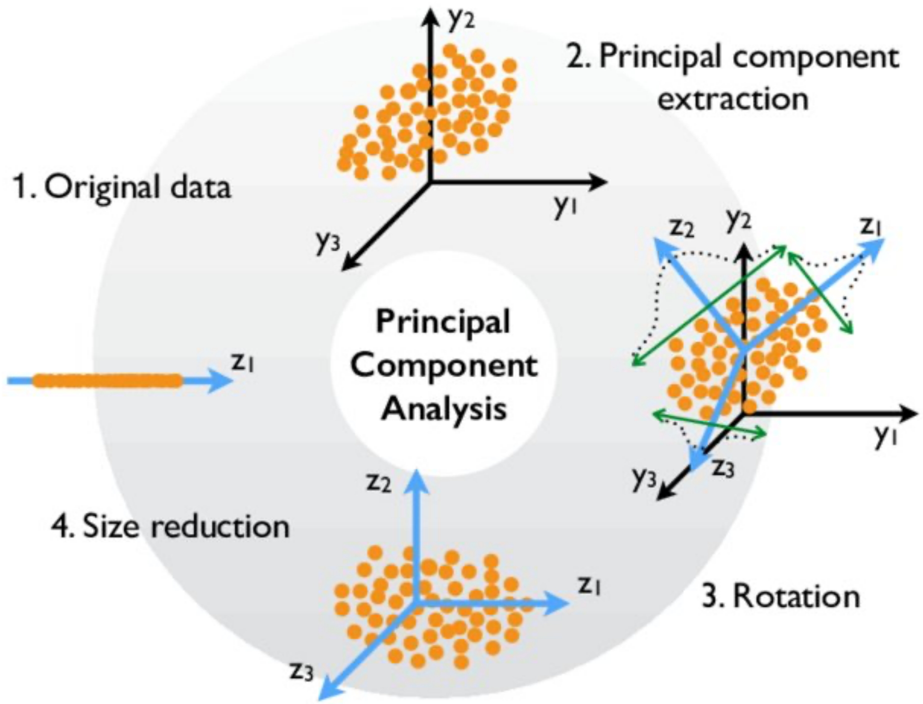
PROF. ALBERTO SIGNORONI

MULTISPECTRAL TRANSFORMATIONS OF IMAGE DATA



Introduction

- The multispectral/multichannel (vectorial) character of most remote sensing image data renders it amenable to **spectral transformations** that **generate new sets of image components** or bands.
  - These components then represent an **alternative description of the data**, in which the new components of a pixel vector are related to its old brightness values in the original set of spectral bands via a linear combination.
- The transformed image
  - may make evident features not discernable in the original data
    - this has significance for **image enhancement**
    - or as **preconditioning** of image data **prior to classification** techniques
  - or alternatively it might be possible to preserve the essential information content of the image (for a given application) with a reduced number of the transformed dimensions.
    - this has significance for **displaying data** in the three dimensions available on a **colour** monitor or in colour hardcopy,
    - and for **transmission and storage** of data.
- Despite specialized transform exists for Remote Sensing data we will only see a very general transform, the **principal components transformation (a.k.a. Principal Component Analysis, PCA)**
  - With **PCA** we transform a set of P correlated variables into a smaller number (K<P) of uncorrelated variables called **principal components** while keeping as much of the variability in the original data as possible.



THE PRINCIPAL COMPONENTS TRANSFORMATION

The principal components transform defined in the following is also known as the Principal Component Analysis (PCA), the Karhunen-Loève transform (KLT) or the Hotelling transform.

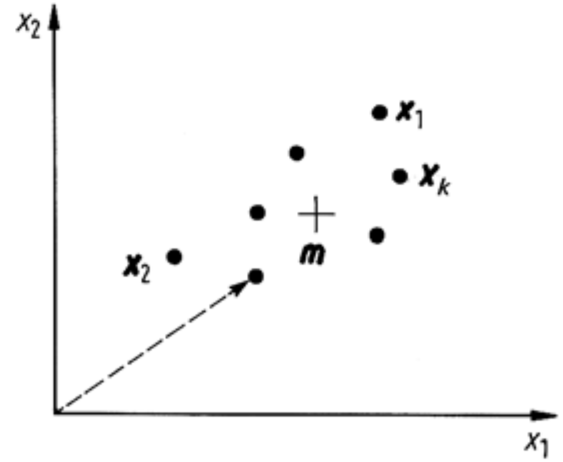
The Principal Components Transformation

- The **multispectral** (or multidimensional or multi-channel) nature of many imaging data (in remote sensing, biomedical, industrial fields) can be accommodated by constructing a **vector space** with as many axes (or dimensions) as there are spectral components associated with each pixel.
  - In the case of **Landsat Thematic Mapper** data, it will have **seven dimensions**
  - For **RGB** color images it will be **three dimensional**
  - For **hyperspectral** data there may be **several hundred axes**
  - **…**

- A particular ***pixel*** *in an image is plotted as a* ***point in such a space*** with *coordinates* that correspond to the *brightness values* of the pixels in the appropriate spectral components.
  - For simplicity, the treatment to be developed in this topic will be based upon a two-dimensional multispectral space (say visible red and infrared) since the diagrams are then easily understood and the mathematical detail is readily assimilated.
  - The results derived however are perfectly general and apply to data of any dimensionality.

The Mean Vector and Covariance Matrix

- The positions of pixel points in multispectral space can be described by vectors, whose components are the individual spectral responses in each band.
  - Consider a multispectral space with a large number of pixels plotted (see Figure), with each pixel described by its appropriate vector ***x*.**
  - The **mean position** of the pixels in the space is defined by the expected value of the pixel vector ***x***, according to



$$m = \mathscr{E}(x\} = \frac{1}{K}\sum_{k=1}^{K} x_k$$

where ***m*** is the mean pixel vector and the x_k are the individual pixel vectors of total number *K;* $\mathscr{E}$ is the expectation operator.

  - While the mean vector is useful to define the average or expected position of the pixels in multispectral vector space, it is of value to have available a means by which their scatter or spread is described. This is the role of the **covariance matrix** which is defined as

$$\Sigma_x = \mathscr{E}\{(x - m)(x - m)^t\}$$

in which the superscript '*t* ' denotes vector transpose.

The Mean Vector and Covariance Matrix

- An *unbiased estimate* of the covariance matrix is given by

$$\Sigma_x = \frac{1}{K-1}\sum_{k=1}^{K}(\boldsymbol{x}_k - \boldsymbol{m})(\boldsymbol{x}_k - \boldsymbol{m})^t$$

- The covariance matrix is one of the most important mathematical concepts in the analysis of multispectral data and in general multiband or multi-component data.
  - If there is correlation between the responses in a pair of spectral bands the corresponding *off-diagonal element* in the covariance matrix will be large by comparison to the diagonal terms.
  - On the other hand, if there is a little correlation, the off-diagonal terms will be close to zero.
- This behaviour can also be described in terms of the correlation matrix *R* whose elements are related to those of the covariance matrix by

$$\varrho_{ij} = v_{ij}/\sqrt{v_{ii}v_{jj}}$$

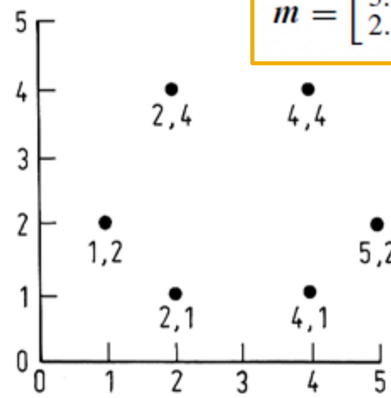
where ϱ_{ij} *is an element of the correlation matrix and* v_{ij} *etc. are elements of the* covariance matrix; v_{ii} *and* v_{jj} *are the variances of the i*-th and *j*-th bands of data.

The ϱ_{ij} describes the correlation between band *i* and band *j*.

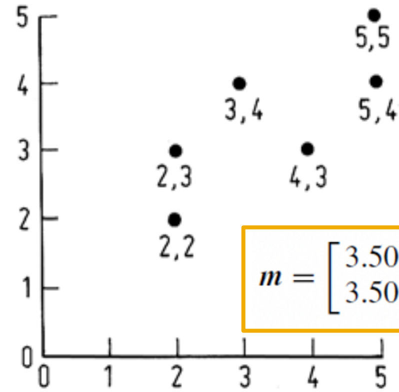
An example

- **Figure a** shows *little correlation* between the two components: in other words, both components are necessary to describe where a pixel lies in the space.
- The data shown in **Figure b** however exhibits a *high degree of correlation* between its two components, evident in the elongated spread of the data at an angle to the axes.

$$m = \begin{bmatrix} 3.00 \\ 2.33 \end{bmatrix}$$



a



b

$$m = \begin{bmatrix} 3.50 \\ 3.50 \end{bmatrix}$$

x	$x-m$	$[x-m]\,[x-m]^t$
$\begin{bmatrix} 1 \\ 2 \end{bmatrix}$	$\begin{bmatrix} -2.00 \\ -0.33 \end{bmatrix}$	$\begin{bmatrix} 4.00 & 0.66 \\ 0.66 & 0.11 \end{bmatrix}$
$\begin{bmatrix} 2 \\ 1 \end{bmatrix}$	$\begin{bmatrix} -1.00 \\ -1.33 \end{bmatrix}$	$\begin{bmatrix} 1.00 & 1.33 \\ 1.33 & 1.77 \end{bmatrix}$
$\begin{bmatrix} 4 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 1.00 \\ -1.33 \end{bmatrix}$	$\begin{bmatrix} 1.00 & -1.33 \\ -1.33 & 1.77 \end{bmatrix}$
$\begin{bmatrix} 5 \\ 2 \end{bmatrix}$	$\begin{bmatrix} 2.00 \\ -0.33 \end{bmatrix}$	$\begin{bmatrix} 4.00 & -0.66 \\ -0.66 & 0.11 \end{bmatrix}$
$\begin{bmatrix} 4 \\ 4 \end{bmatrix}$	$\begin{bmatrix} 1.00 \\ 1.67 \end{bmatrix}$	$\begin{bmatrix} 1.00 & 1.67 \\ 1.67 & 2.79 \end{bmatrix}$
$\begin{bmatrix} 2 \\ 4 \end{bmatrix}$	$\begin{bmatrix} -1.00 \\ 1.67 \end{bmatrix}$	$\begin{bmatrix} 1.00 & -1.67 \\ -1.67 & 2.79 \end{bmatrix}$

whereupon $m = \begin{bmatrix} 3.00 \\ 2.33 \end{bmatrix}$ $\Sigma_x = \begin{bmatrix} 2.40 & 0 \\ 0 & 1.87 \end{bmatrix}$

and $R = \begin{bmatrix} 1.00 & 0 \\ 0 & 1.00 \end{bmatrix}$

where R is the correlation matrix.

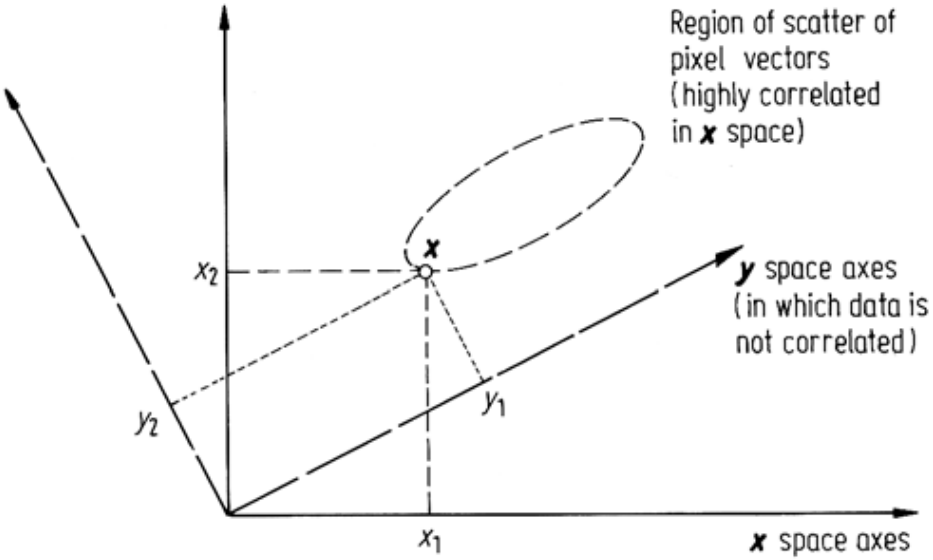
$$m = \begin{bmatrix} 3.50 \\ 3.50 \end{bmatrix} \quad \Sigma_x = \begin{bmatrix} 1.900 & 1.100 \\ 1.100 & 1.100 \end{bmatrix}$$

$$R = \begin{bmatrix} 1.000 & 0.761 \\ 0.761 & 1.000 \end{bmatrix}$$

Thus components 1 and 2 of the data in Figure b are 76% correlated.

A Zero Correlation, Rotational Transform

- It is fundamental to the development of the *principal components analysis* to ask whether there is a *new co-ordinate system in the multispectral vector space in which the data can be represented* ***without correlation*** (see Figure); in other words, such that the covariance matrix in the new co-ordinate system is *diagonal.*



If the vectors describing the pixel points are represented as $\boldsymbol{y}$ in the new co-ordinate system then it is desired to find a linear transformation G of the original co-ordinates, such that

$$\boldsymbol{y} = G\boldsymbol{x} = D^t\boldsymbol{x}$$

subject to the constraint that the covariance matrix of the pixel data in $\boldsymbol{y}$ space is diagonal. Expressing G as D^t will make the comparison of principal components with other transformation operations, treated later, much simpler.

In $\boldsymbol{y}$ space the covariance matrix is, by definition,

$$\Sigma_y = \mathscr{E}\{(\boldsymbol{y} - \boldsymbol{m}_y)(\boldsymbol{y} - \boldsymbol{m}_y)^t\}$$

where $\boldsymbol{m}_y$ is the mean vector expressed in terms of the $\boldsymbol{y}$ co-ordinates. It is shown readily that

$$\boldsymbol{m}_y = \mathscr{E}\{\boldsymbol{y}\} = \mathscr{E}\{D^t\boldsymbol{x}\} = D^t\mathscr{E}\{\boldsymbol{x}\} = D^t\boldsymbol{m}_x \text{ [1]}$$

[1] $\mathscr{E}\{D^t x\} = \frac{1}{K}\sum_{k=1}^{K} D^t x_k = D^t \frac{1}{K}\sum_{k=1}^{K} x_k = D^t m_x$

i.e. D^t, being a matrix of constants, can be taken outside

A Zero Correlation, Rotational Transform

where $\boldsymbol{m}_x$ is the data mean in $\boldsymbol{x}$ space. Therefore

$$\Sigma_y = \mathscr{E}\{(D^t\boldsymbol{x} - D^t\boldsymbol{m}_x)(D^t\boldsymbol{x} - D^t\boldsymbol{m}_x)^t\}$$

which can be written as

$$\Sigma_y = D^t\mathscr{E}\{(\boldsymbol{x} - \boldsymbol{m}_x)(\boldsymbol{x} - \boldsymbol{m}_x)^t\}D \ ^2$$

i.e. $\Sigma_y = D^t\,\Sigma_x\,D$

where Σ_x is the covariance of the pixel data in $\boldsymbol{x}$ space. Since Σ_y must, by demand, be diagonal, D can be recognised as the matrix of eigenvectors of Σ_x, provided D is an orthogonal matrix. Thus we can recognize, from basic linear algebra that we are dealing with the diagonalization of a matrix. As a result, Σ_y can then be identified as the diagonal matrix of eigenvalues of Σ_x,

$$\Sigma_y = \begin{bmatrix} \lambda_1 & 0 & & & \\ 0 & \lambda_2 & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & \lambda_N \end{bmatrix}$$

where N is the dimensionality of the data. Since Σ_y is, by definition, a covariance matrix and is diagonal, its elements will be the variances of the pixel data in the respective transformed co-ordinates. It is arranged such that $\lambda_1 > \lambda_2 > \ldots \lambda_N$ so that the data exhibits maximum variance in y_1, the next largest variance in y_2 and so on, with minimum variance in y_N.

[2] Since $[A\zeta]^t = \zeta^t A^t$ (reversed law of matrices). Note also $[A\zeta]^{-1} = \zeta^{-1}A^{-1}$.

An example (contd)

Before proceeding it is of value at this stage to pursue further the examples of **Figure** to demonstrate the computational aspects of principal components analysis. Recall that the original $\boldsymbol{x}$ space covariance matrix for the highly correlated image data of **Figure** b is

$$\Sigma_x = \begin{bmatrix} 1.90 & 1.10 \\ 1.10 & 1.10 \end{bmatrix}$$

To determine the principal components transformation it is necessary to find the eigenvalues and eigenvectors of this matrix. The eigenvalues are given by the solution to the characteristic equation

$$|\Sigma_x - \lambda \boldsymbol{I}| = 0, \quad \boldsymbol{I} \text{ being the identity matrix.}$$

i.e. $$\begin{vmatrix} 1.90-\lambda & 1.10 \\ 1.10 & 1.10-\lambda \end{vmatrix} = 0$$

or $\lambda^2 - 3.0\lambda + 0.88 = 0$

which yields $\lambda = 2.67$ and 0.33

As a check on the analysis it may be noted that the sum of the eigenvalues is equal to the trace of the covariance matrix, which is the sum of its diagonal elements.

The covariance matrix in the appropriate $\boldsymbol{y}$ co-ordinate system (with principal components as axes) is therefore

$$\Sigma_y = \begin{bmatrix} 2.67 & 0 \\ 0 & 0.33 \end{bmatrix}$$

Note that the first principal component, as it is called, accounts for $2.67/(2.67 + 0.33) \equiv 89\%$ of the total variance of the data in this particular example. It is now of interest to find the actual principal components transformation matrix $G = D^t$.

An example (contd)

Note that this is the *transposed* matrix of eigenvectors of Σ_x. Consider first, the eigenvector corresponding to $\lambda_1 = 2.67$. This is the vector solution to the equation

$$[\Sigma_x - \lambda_1 \boldsymbol{I}]\, \boldsymbol{g}_1 = 0$$

with $\boldsymbol{g}_1 = \begin{bmatrix} g_{11} \\ g_{21} \end{bmatrix} \equiv \boldsymbol{d}_1^t$ for the two dimensional example at hand.

Substituting for $\Sigma_x$and λ_1 gives the pair of equations

$$\begin{aligned} -0.77g_{11} + 1.10g_{21} &= 0 \\ 1.10g_{11} - 1.57g_{21} &= 0 \end{aligned}$$

which are not independent, since the set is homogeneous. It does have a non-trivial solution however because the coefficient matrix has a zero determinant. From either equation it can be seen that

$$g_{11} = 1.43g_{21} \tag{6.6}$$

At this stage either g_{11} or g_{21} would normally be chosen arbitrarily, and then a value would be computed for the other. However the resulting matrix G has to be orthogonal so that $G^{-1} \equiv G^t$. This requires the eigenvectors to be normalised, so that

$$g_{11}^2 + g_{21}^2 = 1 \tag{6.7}$$

This is a second equation that can be solved simultaneously with (6.6) to give

$$g_1 = \begin{bmatrix} 0.82 \\ 0.57 \end{bmatrix}$$

An example (contd)

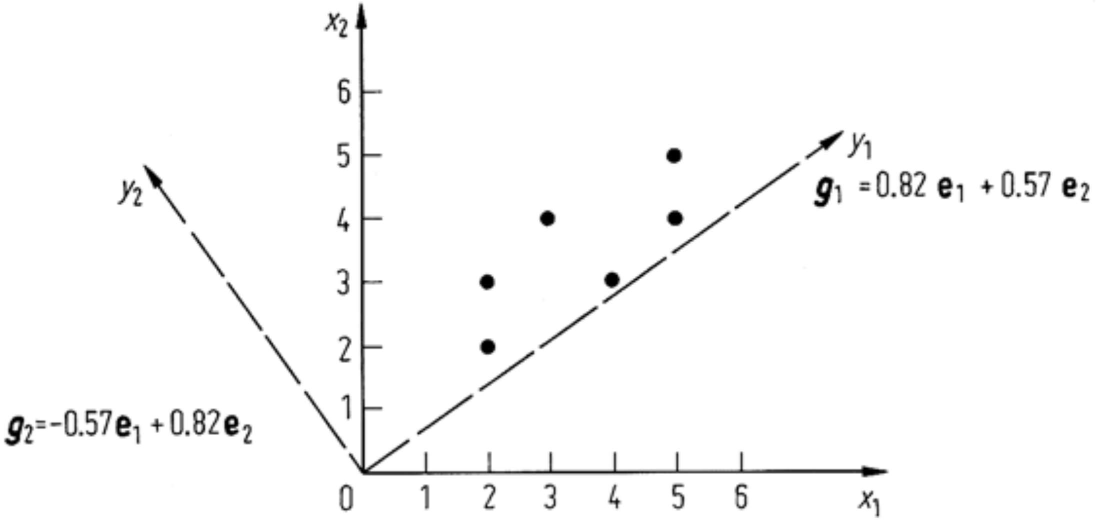
In a similar manner it can be shown that the eigenvector corresponding to $\lambda_2 = 0.33$ is

$$g_2 = \begin{bmatrix} -0.57 \\ 0.82 \end{bmatrix}$$

The required principal components transformation matrix therefore is

$$G = D^t = \begin{bmatrix} 0.82 & -0.57 \\ 0.57 & 0.82 \end{bmatrix}^t = \begin{bmatrix} 0.82 & 0.57 \\ -0.57 & 0.82 \end{bmatrix}$$

Now consider how these results can be interpreted. First of all, the individual eigenvectors g_1 and g_2 are vectors which define the principal component axes in terms of the original co-ordinate space. These are shown in Figure : it is evident that the data is uncorrelated in the new axes and that the new axes are a rotation of the original set. For this reason (even in more than two dimensions) the principal components transform is classed as a rotational transform.



An example (contd)

Secondly, consider the application of the transformation matrix G to find the positions (i.e., the brightness values) of the pixels in the new uncorrelated co-ordinate system. Since $\boldsymbol{y} = G\boldsymbol{x}$ this example gives

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 0.82 & 0.57 \\ -0.57 & 0.82 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{6.8}$$

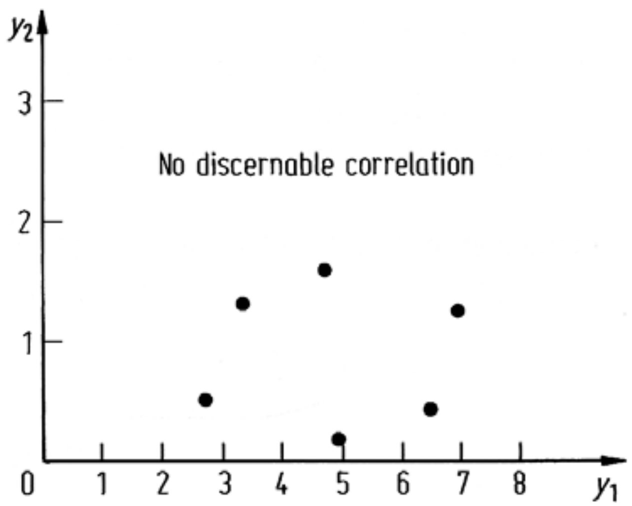
which is the actual principal components transformation to be applied to the image data. Thus, for

$$\boldsymbol{x} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 4 \\ 3 \end{bmatrix}, \begin{bmatrix} 5 \\ 4 \end{bmatrix}, \begin{bmatrix} 5 \\ 5 \end{bmatrix}, \begin{bmatrix} 3 \\ 4 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

we find

$$\boldsymbol{y} = \begin{bmatrix} 2.78 \\ 0.50 \end{bmatrix}, \begin{bmatrix} 4.99 \\ 0.18 \end{bmatrix}, \begin{bmatrix} 6.38 \\ 0.43 \end{bmatrix}, \begin{bmatrix} 6.95 \\ 1.25 \end{bmatrix}, \begin{bmatrix} 4.74 \\ 1.57 \end{bmatrix}, \begin{bmatrix} 3.35 \\ 1.32 \end{bmatrix}.$$

The pixels plotted in $\boldsymbol{y}$ space are shown in Figure . Several points are noteworthy.



Practical Considerations

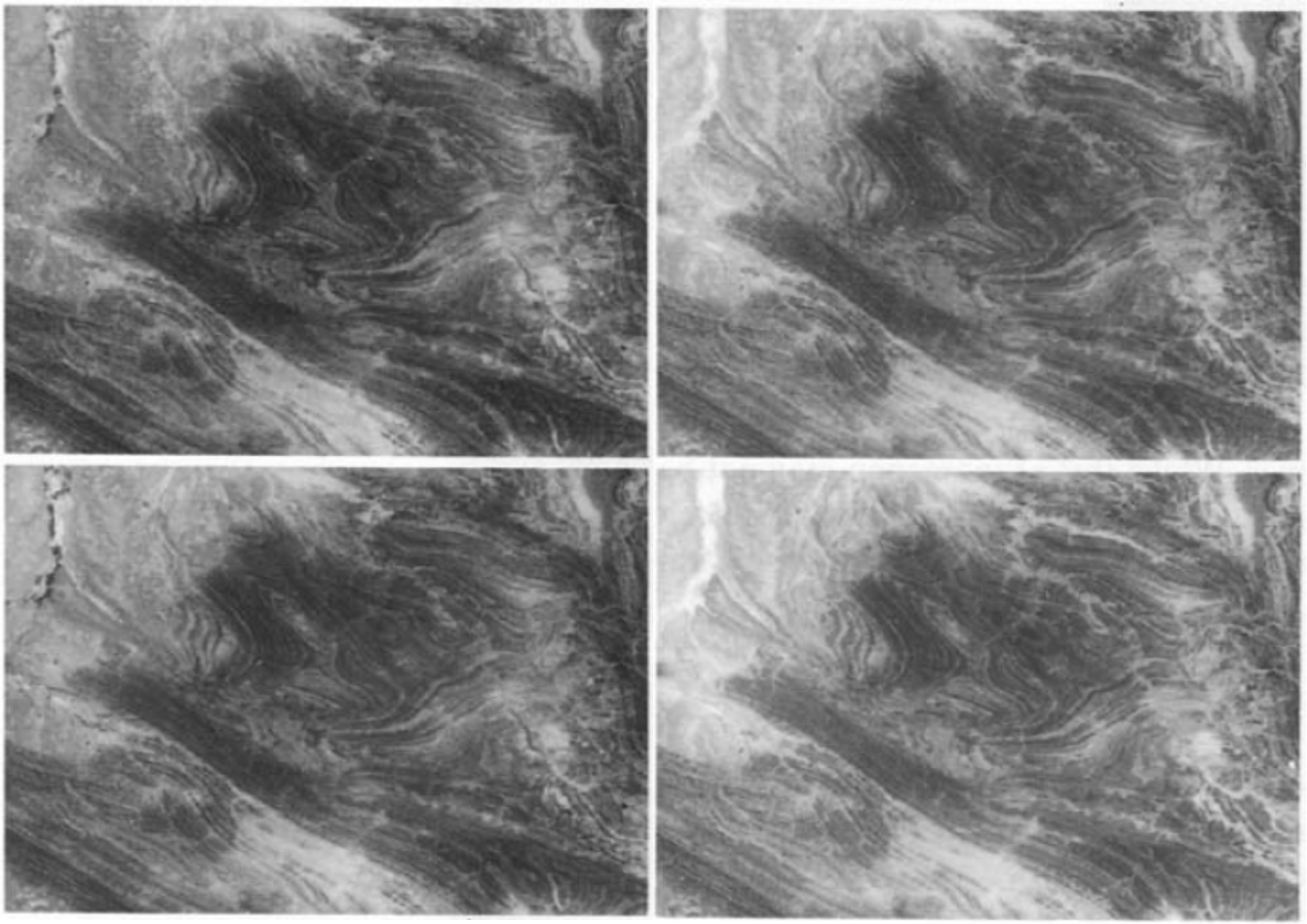
- From the above example some considerations can be drawn:

1. First, the data exhibits <u>no discernable correlation between the pair of new axes</u> (i.e., the principal components).
2. Secondly, <u>most of the data spread is in the direction of the first principal component</u>. It could be interpreted that this component contains most of the information in the image.
3. Finally, if the pair of principal component images are produced by using the $\boldsymbol{y_1}$ and $\boldsymbol{y_2}$ component brightness values for the pixels, the first principal component image will show a *high degree of contrast* whereas the second will *have limited contrast*.

   - By comparison to the first component, the second will make use of only a few available brightness levels. It will be seen, therefore, to lack the detail of the former.
   - While this phenomenon may not be particularly evident for a simple two dimensional example, it is especially noticeable in the fourth component of a principal component transformed Landsat multispectral scanner image as can be assessed in the next example.

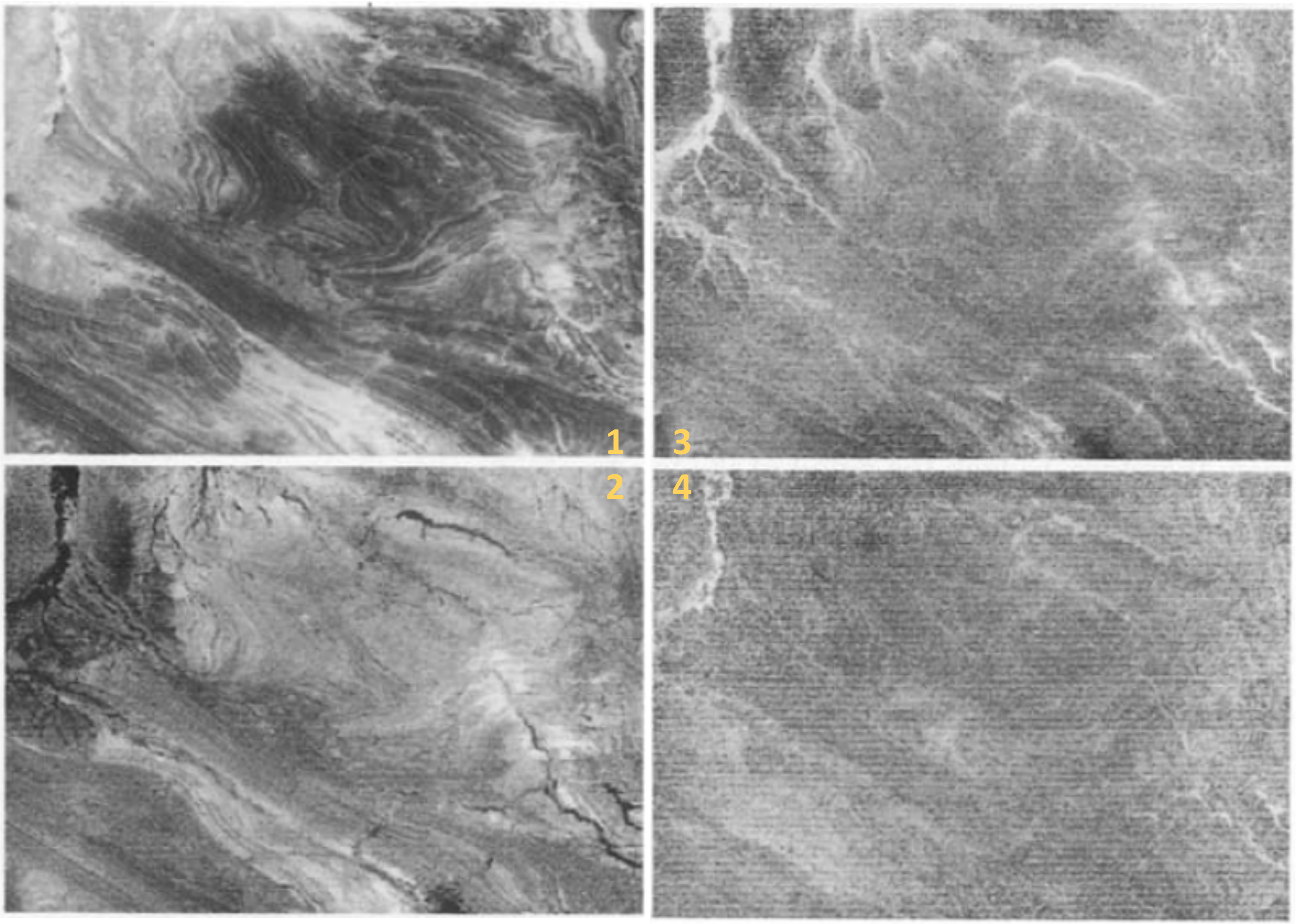
Real Example

- Four original bands acquired by the Landsat MS scanner (Andamooka, central Australia)



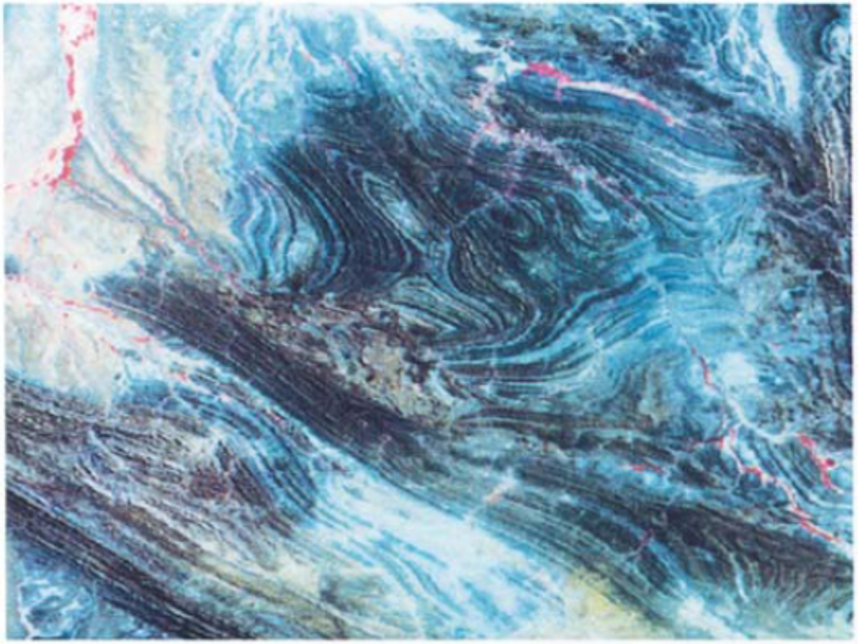
Real Example (contd)

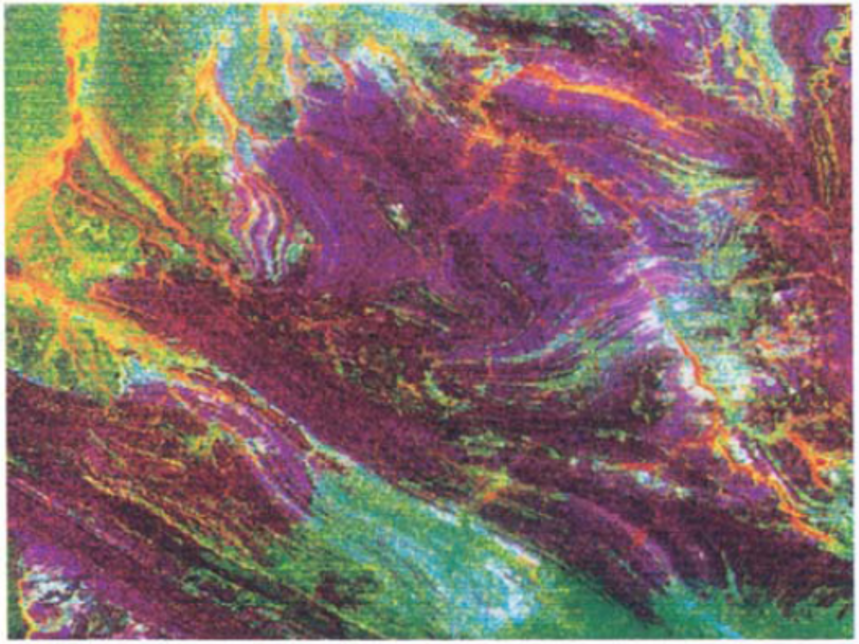
- Four principal components of the same image segment



Real Example (contd)

- Comparison of standard false colour composite (band 7 to red, band 5 to green and band 4 to blue)
- with a principal component composite (first component to red, second to green and third to blue)
- … below some pictures from the neighbor





Real Example (contd)

- The covariance matrix for this image is

$$\Sigma_x = \begin{bmatrix} 34.89 & 55.62 & 52.87 & 22.71 \\ 55.62 & 105.95 & 99.58 & 43.33 \\ 52.87 & 99.58 & 104.02 & 45.80 \\ 22.71 & 43.33 & 45.80 & 21.35 \end{bmatrix}$$

and its eigenvalues and eigenvectors are:

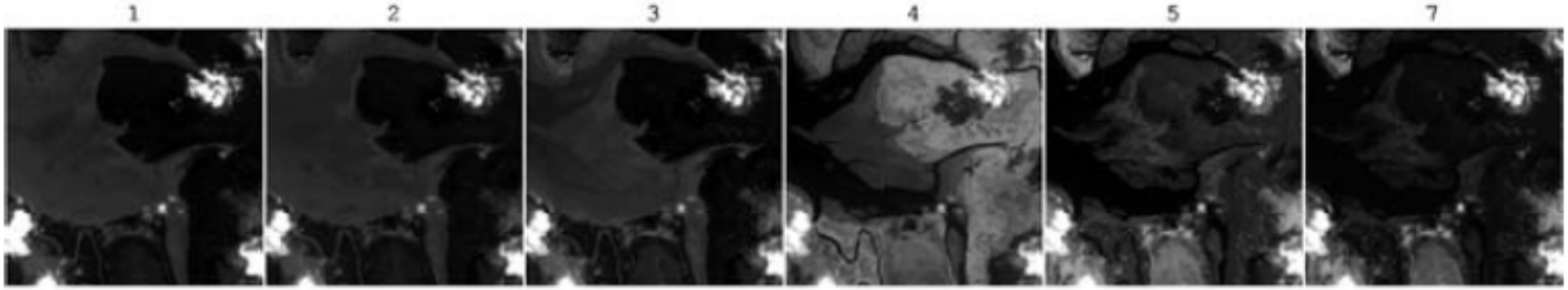
eigenvalues	253.44	7.91	3.96	0.89
eigenvector	0.34	−0.61	0.71	−0.06
components	0.64	−0.40	−0.65	−0.06
(vertically)	0.63	0.57	0.22	0.48
	0.28	0.38	0.11	−0.88

- The first principal component image will be expected therefore to contain 95% of the data variance. By comparison, the variance in the last component is seen to be negligible. It is to be expected that this component will appear almost totally as noise of low amplitude.
- The four principal component images show the information redistribution and compression properties of the transformation. By association with the numerical example it would be expected that the later components should appear dull and poor in contrast. **The high brightness and the contrasts displayed are a result of a contrast enhancement (range normalization) applied to the components for the purpose of display. This also allow to visually appreciate the poor signal to noise ratio of the last components.**

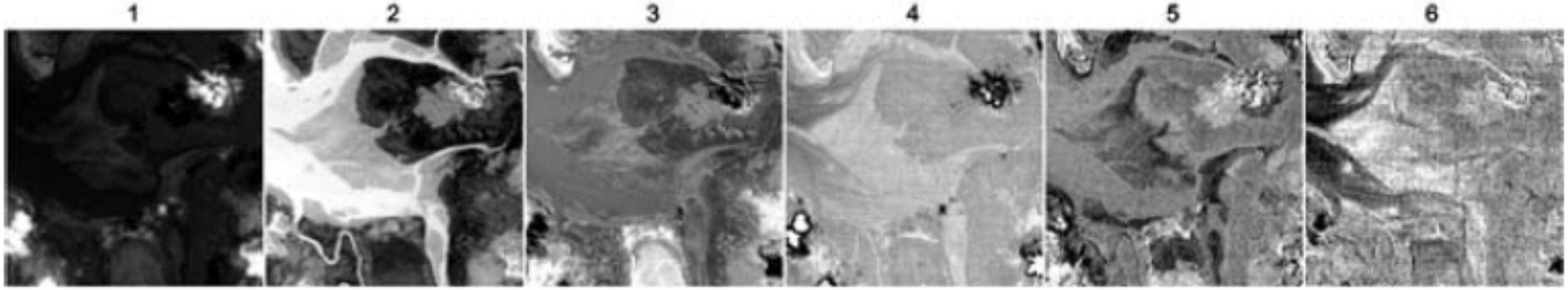
A second Real Example

- A second example of the principal components transformation is shown in Figure, this time based on the 6 reflective TM bands for a region in the Northern Territory of Australia.

- The original TM bands: visible (1,2,3), near IR, NIR (4), shortwave IR, SWIR (5,7)



- The full set of principal components



A second Real Example (contd)

- The covariance and correlation matrices for the image are:

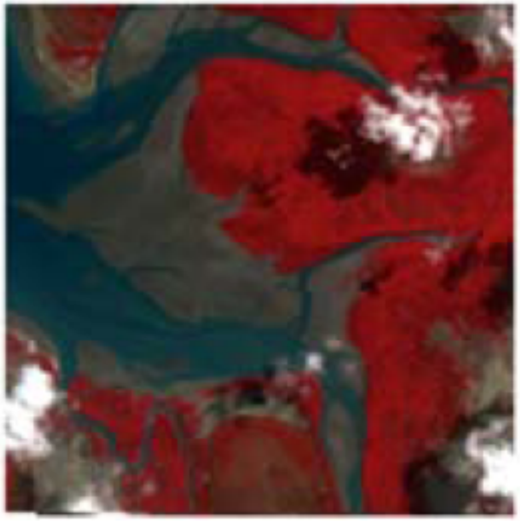
$$\Sigma_x = \begin{bmatrix} 874.98 & 550.56 & 698.00 & 335.54 & 858.15 & 551.21 \\ 550.56 & 363.82 & 454.79 & 230.30 & 558.88 & 358.38 \\ 689.00 & 454.79 & 580.63 & 288.11 & 747.97 & 471.72 \\ 335.54 & 230.30 & 288.11 & 722.46 & 742.35 & 387.61 \\ 858.15 & 558.88 & 747.97 & 742.35 & 1544.70 & 871.29 \\ 551.21 & 358.38 & 471.72 & 387.61 & 871.29 & 514.18 \end{bmatrix}$$

$$R_x = \begin{bmatrix} 1.00 & 0.98 & 0.97 & 0.42 & 0.74 & 0.82 \\ 0.98 & 1.00 & 0.99 & 0.45 & 0.75 & 0.83 \\ 0.97 & 0.99 & 1.00 & 0.44 & 0.79 & 0.86 \\ 0.42 & 0.45 & 0.44 & 1.00 & 0.70 & 0.64 \\ 0.74 & 0.75 & 0.79 & 0.70 & 1.00 & 0.98 \\ 0.82 & 0.83 & 0.86 & 0.64 & 0.98 & 1.00 \end{bmatrix}$$

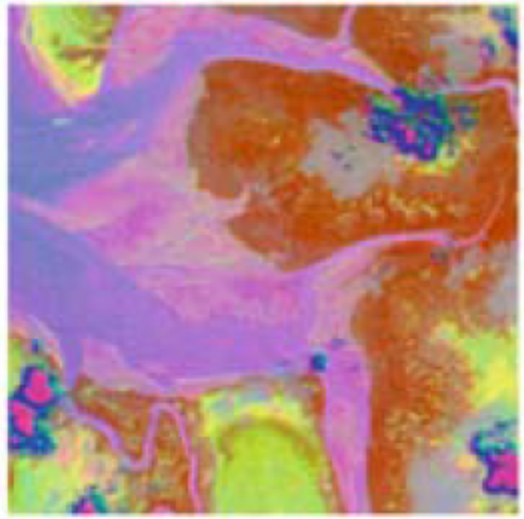
  - From the correlation matrix we directly see that the correlation among bands is high
  - Thus we can expect a significant effect of the principal component transform
  - The eigenvalues and eigenvectors are:

eigenvalues	3727.35	613.34	226.14	23.52	8.16	2.25
eigenvectors	first	second	third	fourth	fifth	sixth
	0.433	0.485	−0.307	−0.684	−0.089	0.088
	0.282	0.294	−0.218	0.369	0.094	−0.801
	0.364	0.347	−0.127	0.627	−0.153	0.561
	0.303	−0.673	-0.671	0.018	0.042	0.056
	0.615	−0.322	0.562	−0.052	−0.429	−0.129
	0.362	−0.047	0.275	−0.026	0.880	0.127

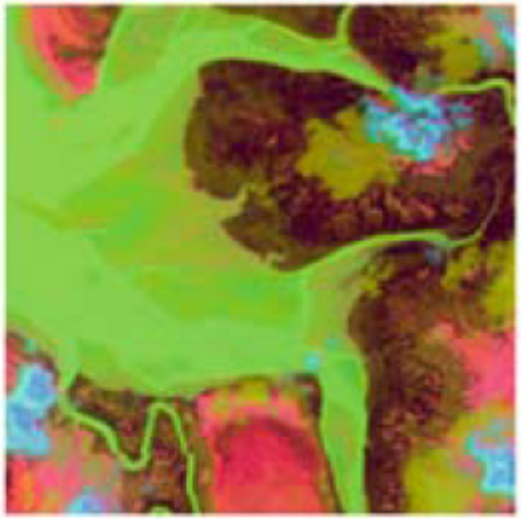
A second Real Example (contd)



a



b



c

- False color composites comparison
  - Figure a shows a colour composite formed by mapping the original bands 4, 3, and 2 to red, green and blue respectively.
  - Figure b shows PC3, PC2 and PC1 mapped to red, green and blue,
  - while Figure c shows PC4, PC3 and PC2 mapped to red, green and blue.
- Interestingly, the last **PC4, PC3, PC2** colour composite shows more detail for those ground covers whose **spectral** responses are dominant in the visible to near infrared regions, since PC4 (determined by the fourth eigenvector) is largely a difference image in the visible region (this can be seen from the linear combination coefficients in the fourth eigenvector).
- In contrast **PC1** is essentially just a total brightness image (**topographic**), as can be seen from the first eigenvector, so that it does little to enhance spectral differences.

A third real example: Landsat Thematic Mapper

TM1

A third real example: Landsat Thematic Mapper

TM2

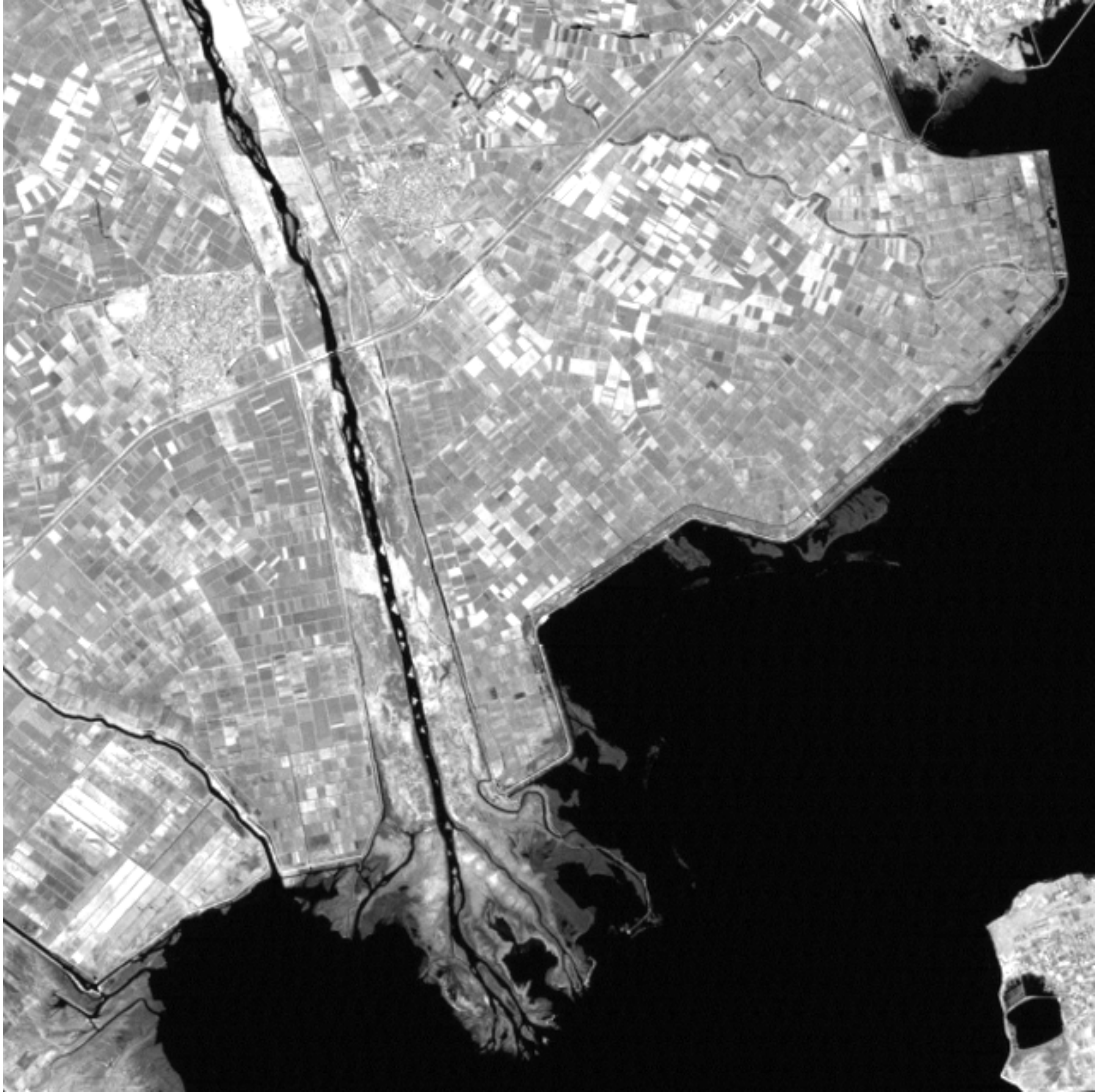
A third real example: Landsat Thematic Mapper

TM3

A third real example: Landsat Thematic Mapper

TM4

A third real example: Landsat Thematic Mapper



TM5

A third real example: Landsat Thematic Mapper

TM7

A third real example: Landsat Thematic Mapper

TM6

A third real example: Landsat Thematic Mapper

Original spectral bands

$$\mathbf{R} = \begin{bmatrix} 1 & 0.91 & 0.88 & 0.51 & 0.71 & 0.58 & 0.77 \\ 0.91 & 1 & 0.97 & 0.74 & 0.83 & 0.56 & 0.82 \\ 0.88 & 0.97 & 1 & 0.76 & 0.89 & 0.63 & 0.89 \\ 0.51 & 0.74 & 0.76 & 1 & 0.86 & 0.56 & 0.74 \\ 0.71 & 0.83 & 0.89 & 0.86 & 1 & 0.76 & 0.97 \\ 0.58 & 0.56 & 0.63 & 0.56 & 0.76 & 1 & 0.80 \\ 0.77 & 0.82 & 0.89 & 0.74 & 0.97 & 0.80 & 1 \end{bmatrix}$$

$$\sigma_i^2 =$$

66.32 53.03 214.69 797.76 1359.65 19.64 312.47

$$\sigma_i =$$

8.14 7.28 14.65 28.24 36.87 4.43 17.68

Principal Components

	PC1	PC2	PC3	PC4	PC5	PC6	PC7
σ_j^2	2543.73	192.64	68.11	9.61	5.12	3.25	1.08
σ_j	50.44	13.88	8.26	3.10	2.26	1.80	1.04

A third real example: Landsat Thematic Mapper

PC1

A third real example: Landsat Thematic Mapper

PC2

A third real example: Landsat Thematic Mapper

PC3

A third real example: Landsat Thematic Mapper

PC4

A third real example: Landsat Thematic Mapper

PC5

A third real example: Landsat Thematic Mapper

PC6

A third real example: Landsat Thematic Mapper

PC7

A third real example: Landsat Thematic Mapper

PC1
+
PC2
+
PC3

Remarks about the Principal Components Transform

- The covariance matrix used to generate the principal component transformation matrix is a *global measure of the variability of the original image segment*.

- Principal Component images are useful
  - for reducing data dimensionality,
  - condensing topographic and spectral information,
  - improving image colour presentation,
  - and enhancing some spectral features.

- Notwithstanding the negligible information content of the last, or last few, image components resulting from a principal components analysis it is sometimes important to examine all components since often local detail may appear in a later component (e.g. look at the small dark field in PC6 in the lower left part).
  - Abnormal local detail therefore may not necessarily be mapped into one of the earlier components but could just as easily appear later. This is often the case with geological structure.
  - Some variants of the principal component transform have been proposed to enforce the possibility of a better separation between topographic and specific spectral information.

The Effect of an Origin Shift

- It is evident that some principal component pixel brightnesses could be negative owing to the fact that the transformation is a simple axis rotation.
  - Clearly a combination of positive and negative brightnesses cannot be displayed.
  - Nor can negative brightness pixels be ignored since their appearance relative to the other pixels in a component serve to define detail.
  - In practice, the problem with negative values is accommodated by shifting the origin of the principal components space to yield all components with positive and thus displayable brightnesses.
  - This has no effect on the properties of the transformation as can be seen by inserting an origin shift term in the definition of the covariance matrix in the principal components axes.
  - Define $\boldsymbol{y}' = \boldsymbol{y} - \boldsymbol{y}_0$ where $\boldsymbol{y}_0$ is the position of a new origin. In the new $\boldsymbol{y}'$ co-ordinates

$$\Sigma_{y'} = \mathscr{E}\{(\boldsymbol{y}' - \boldsymbol{m}_{y'})(\boldsymbol{y}' - \boldsymbol{m}_{y'})^t\}$$

  - Now $\boldsymbol{m}_y = \boldsymbol{m}_{y'} - \boldsymbol{y}_0$ so that $\boldsymbol{y}' - \boldsymbol{m}_{y'} = \boldsymbol{y} - \boldsymbol{y}_0 - \boldsymbol{m}_y + \boldsymbol{y}_0 = \boldsymbol{y} - \boldsymbol{m}_y$.
  - Thus $\Sigma_{y'} = \Sigma_y$ – i.e. the origin shift has no influence on the covariance of the data in the principal components axes, and can be used for convenience in displaying principal component images.

Application of Principal Components in Image Enhancement and Display

- In constructing a **colour display of remotely sensed data** only three dimensions of information can be mapped to the three colour primaries of the display device.
  - For imagery with more than three bands that means the user must choose the *most appropriate subset* of three to use.
  - A less ad hoc means for colour assignment rests upon performing a principal components transform and assigning the first three components to the red, green and blue colour primaries.
    - ES. Examination of a typical set of principal component images for Landsat MSS data, such as those seen in the first example, reveals that there is very little detail in the fourth component so that, in general, it could be ignored without prejudicing the ability to extract meaningful information from the scene.
- A difficulty with principal components colour display, however, is that **there is no longer a one-to-one mapping between sensor wavelength bands and colours**.
  - **Rather each principal component** (possibly associated to a primary colour) **now represents a linear combination of spectral components**, **making human photointerpretation difficult** for many applications.
    - An exception would be in exploration geology where structural differences may be enhanced in principal components imagery, there often being little interest in the meanings of the actual colours (image texture and structure is more important).

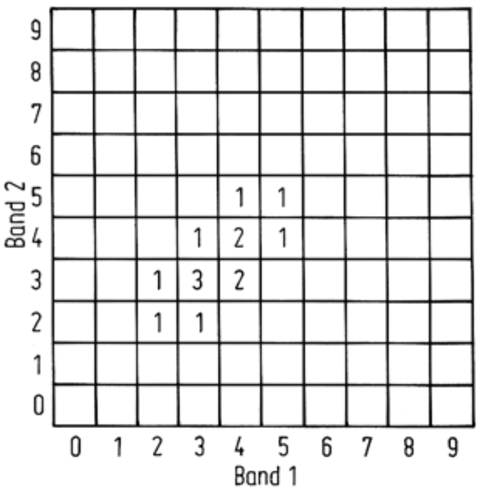
The Taylor Method of Contrast Enhancement (Decorrelation Stretch)

- Application of common contrast modification techniques (e.g. linear stretch, histogram equalization) to each of the individual components of a highly correlated vector image yields an **enhanced image in which certain highly saturated hues are missing.**
  - It is a direct result of the correlation in the image that the highly saturated colour primaries are not displayed.
  - In the display of three-dimensional correlated image data, simple contrast enhancement of each component independently will yield an image *without* highly saturated reds, blues and greens *but also without* saturated yellows, cyans and magentas (color space extremes).

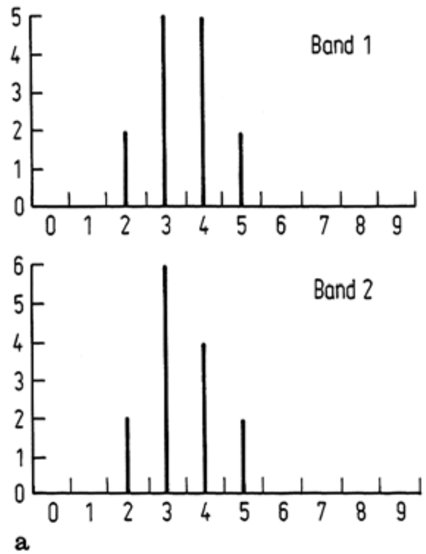
- An interesting **contrast stretching procedure** which can be used to create a modified image with *good utilisation of the range of available hues* **rests upon the use of the principal components transformation**.
  - It was developed by Taylor (1973) and has also been presented by Soha and Schwartz (1978). A more recent and general treatment has been given by Campbell (1996).
  - Also called (on other books) **Principal Component Analysis Decorrelation Stretch**, **PCA-DS**

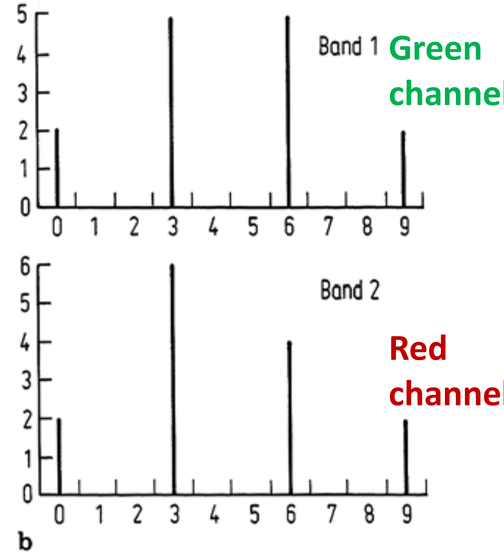
The Taylor Method of Contrast Enhancement (Decorrelation Stretch)

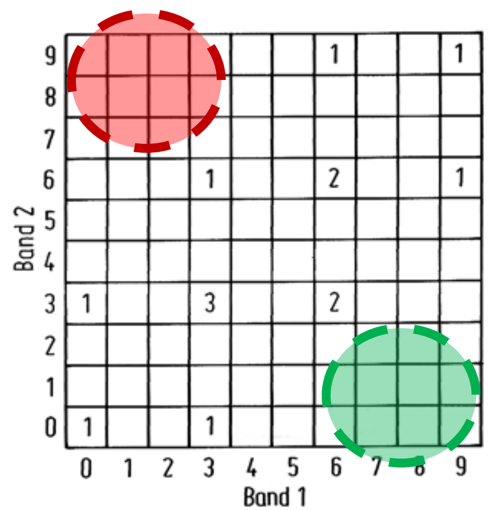
- The problem of a separate band (linear) contrast enhancement



Histogram for a hypothetical two dimensional image showing correlation in its components







- The components' correlation still remains
- Suboptimal exploitation of color channels (hue saturation not reached)

The Taylor Method of Contrast Enhancement (Decorrelation Stretch)

- The solution based on the use of the Principal Component Analysis:

1. Apply PCA to the data to transform the original bands into Principal Components
2. Apply contrast enhancement of each of the PCs (by linear stretching)
3. Apply Inverse PCA to convert back into image bands

- The procedure recommended by Taylor overcomes the lack of hue saturation because it fills the available colour space on the display more fully.

$$\Sigma_x = \mathcal{E}\{(\boldsymbol{x}-\boldsymbol{m})(\boldsymbol{x}-\boldsymbol{m})^t\}$$

Presenting non-zero off-diagonal elements (band correlation)

1.

$$\Sigma_y = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}$$

2.

$$\Sigma_y' = \begin{bmatrix} \sigma^2 & 0 & 0 \\ 0 & \sigma^2 & 0 \\ 0 & 0 & \sigma^2 \end{bmatrix} = \sigma^2 \boldsymbol{I}$$

Possible noise amplification problem can be reduced by denoising in the PC domain

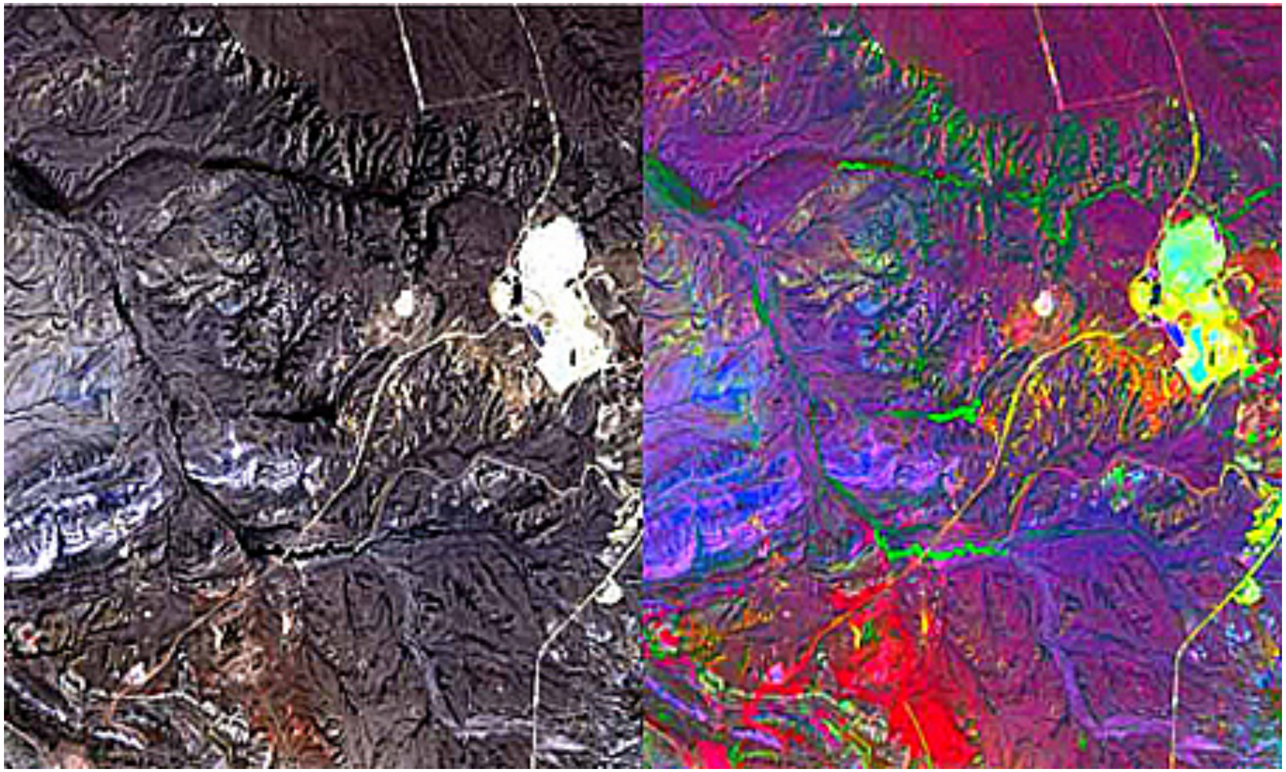
3.

$$\begin{aligned} \Sigma_x' &= G^t \mathcal{E}\{(y' - \mathcal{E}(y'))(y' - \mathcal{E}(y'))\}^t G \\ &= G^t \Sigma_y' G \\ &= G^t \sigma^2 I G \end{aligned}$$

$$\Sigma_x' = \sigma^2 I\ .$$

Decorrelation Stretch Example 1

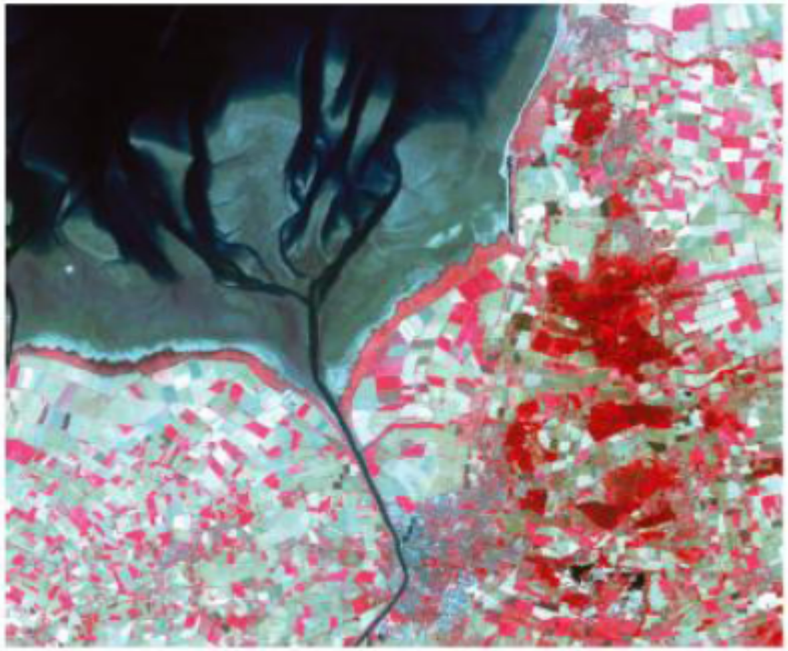
- On the left is a standard false color composite; on the right a DS image - this illustrates the ability to extract and emphasize the tonal differences not apparent in the left image (from http://rst.gsfc.nasa.gov/Sect1/Sect1_14.html)



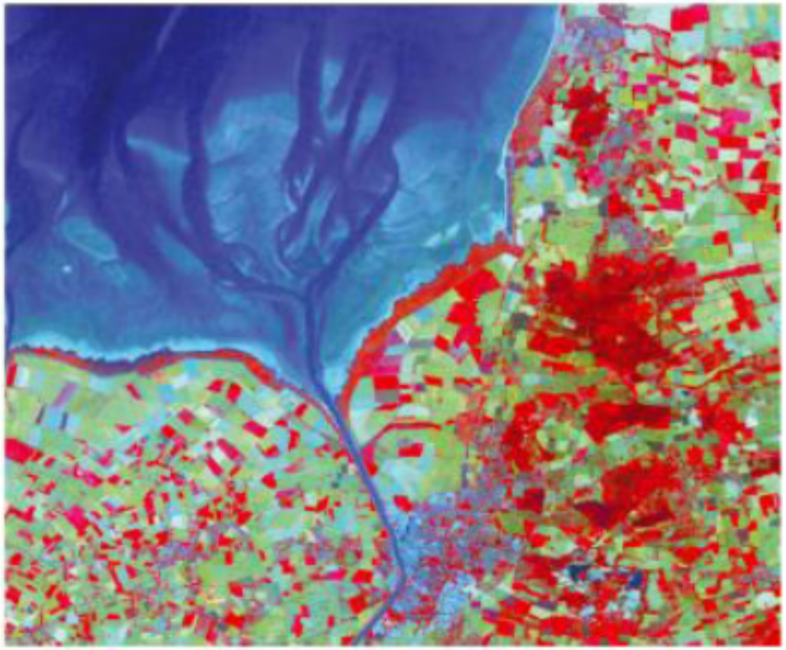
Decorrelation Stretch Example 2

a) Landsat TM bands 4,3,2 false colour composite of the coastline of The Wash, eastern England, after a 5-95% linear contrast stretch

b) The same Wash image after a decorrelation stretch based on the covariance matrix



a)



b)

Computer Processing of Remotely-Sensed Images, An Introduction, 4th edition: Paul M. Mather and Magaly Koch

Other Applications of Principal Components Analysis

- Owing to the information compression properties of the principal components transformation it lends itself to reduced representation of image data for storage or transmission.
  - In such a situation only **the uppermost significant components are retained** as a representation of an image, with the information content so lost being indicated by the sum of the eigenvalues corresponding to the components ignored.
  - Thereafter if the original image is to be restored, either on reception through a communications channel or on retrieval from memory, then *the inverse of the transformation matrix is used to reconstruct the image from the reduced set of components*.
  - Since the matrix is orthogonal its inverse is simply its transpose.
- This technique is known as **bandwidth compression** in the field of telecommunications.
  - For multispectral data it had not found great application in satellite remote sensing image processing, because hitherto image transmission has not been a consideration and available memory has not placed stringent limits on image storage.
  - With increasing use of imaging spectrometry (hyperspectral) data however, *bandwidth compression has become more important*.

- Another interesting application of principal components analysis is in the *detection of features that change with time* between images of the same region. This is exploited in *change detection techniques*.