How to extract useful information from large amounts of (spectral) data which frequently are very complex

“To obtain intelligent answers to questions which were not even posed”

To obtain information on a useful property which is hard to measure by measuring related quantities which are easy to measure
Multi-spectral data
where each pixel contains a spectrum
Example: reflectance of skin in nude mice
OBJECT IDENTIFICATION

FOREST DECLINE MONITORING

CANCER TUMOUR IDENTIFICATION

FLUORESCENCE

Reflectance

TEMPORAL DECAY

Discrimination index $Q$

$$Q = \frac{|M_Y - M_X|}{\sqrt{\sigma_X^2 + \sigma_Y^2}}$$

Who decides what is $X$ and what is $Y$?

Plant physiologist

Pathologist
Multicolour Fluorescence Imaging using a dimension-less contrast function

\[ F_c = \frac{A - k_1 D}{k_2 B} \]

Andersson-Engels et al. LSM (2000)
How to extract useful information from spectra

Example: Recordings from human blood vessels

Plaque or normal?

Testing ratios!
Using double spectral ratios to identify human bladder cancer

Identification of malignant tumours from fluorescence spectral shapes
Statistical correlation in experimental data

\[ M_x = \frac{\sum_{i=1}^{n} x_i}{n} \]

\[ \sigma_x = \sqrt{\frac{\sum_{i=1}^{n} (x_i - M_x)^2}{n-1}} \]

\[ \text{Var}(X) = \frac{\sum_{i=1}^{n} (x_i - M_x)^2}{n-1} \]

\[ \text{Cov}(x, y) = \frac{\sum_{i=1}^{n} (x_i - M_x)(y_i - M_y)}{n-1} \]

\[ r = \frac{\text{Cov}(x, y)}{\sigma_x \sigma_y} \]
Projections in generalized coordinate systems

Space projection

Quantum mechanical projection

\[ r = (x, y, z) \]

\[ \Psi = \sum_i c_i \phi_i \]
Definition of principal components

Best fit for smallest sum of squares of distances – then the variance is maximized
First and second principal components of a three-dimensional data set
Method to evaluate how many PC:s are needed
The new orthogonal coordinate system can be obtained by turning the original orthogonal system; the expansion coefficients are called **loadings**.

The coordinates of the data points in the new system are called **scores**.

**Matrix algebra** is well suited to handle the change in systems.
The concept can be expanded to not just data points but to whole spectra.

An experimental spectrum can be expanded into principal component “spectra”, which are mathematically orthogonal.
Principle Component Analysis
Singular Value Decomposition
(in different wording)
Courtesy: M. Brydegaard

- A matrix factorization method
- A coordinate transform
- Align the first dimension along the largest co-variance and so forth
- Projects data on optimal set of base functions
- Random noise are sorted in last components
- Base functions depends on data, e.g. may change if new data is included
- PCA is based on SVD but centers the data, data cannot be reconstructed after PCA.
Example – scanning fluorescence lidar imaging of the Lund cathedral
Lidar remote fluorescence monitoring system
Lund mobile Lidar system
Remote fluorescence recording set-up
Lund cathedral selected spectra (60 m stand-off)
Building up a spectrum from principal component spectra
Spectra of stones, with now infiltration by algae (chlorophyll)

Strong compression of huge amounts of data
Spectra of stones, sometimes covered by algae (then more PCs are needed!)

Loading plots of the principal components 1, 2, 3, 6
Imaging in individual principal components
Multi-variate spectral analysis classifies areas.
Multispectral imaging with ratios; Orange/Blue

(a) 448 nm filter  
(b) 600 nm filter  
(c) picture portal  
(d) 600/448
Other spectral band combinations

600/448 690/448
Rome
Coliseum
Rome Coliseum
Finding areas with similar fluorescence properties
Swedish lidar system: 60 m stand-off distance
Coliseum
Chlorophyll and surface treatment
PCA-RGB analysis of travertine
Coliseum, Rome
Example:

Consider seven fruits

(courtesy M. Brydegaard)

- Royal gala
- Red delicious
- Granny Schmith
- Orange
- Golden delicious
- Lemon
- Lime
Their reflectance spectra:

(courtesy: M. Brydegaard)
**Light intensity reflectance measurements**

*7 fruits, many spectral bands*

**Construct the data matrix:**

(courtesy: M. Brydegaard)

![Sample number](image)

<table>
<thead>
<tr>
<th>Sample number</th>
<th>( \lambda_{UV} )</th>
<th>( \lambda_{VIS} )</th>
<th>( \lambda_{NIR} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>I(_{1,1})</td>
<td>( \ldots )</td>
<td>( I_{1,1,1000} )</td>
<td>( \ldots )</td>
</tr>
<tr>
<td>I(_{2,1})</td>
<td>( \ldots )</td>
<td>( I_{2,1,1000} )</td>
<td>( \ldots )</td>
</tr>
<tr>
<td>( \ldots )</td>
<td>( \ldots )</td>
<td>( \ldots )</td>
<td>( \ldots )</td>
</tr>
<tr>
<td>I(_{10,1})</td>
<td>( \ldots )</td>
<td>( I_{10,1,1000} )</td>
<td>( \ldots )</td>
</tr>
<tr>
<td>I(_{11,1})</td>
<td>( \ldots )</td>
<td>( I_{11,1,1000} )</td>
<td>( \ldots )</td>
</tr>
<tr>
<td>( \ldots )</td>
<td>( \ldots )</td>
<td>( \ldots )</td>
<td>( \ldots )</td>
</tr>
<tr>
<td>I(_{N-1,1})</td>
<td>( \ldots )</td>
<td>( I_{N-1,1,1000} )</td>
<td>( \ldots )</td>
</tr>
<tr>
<td>I(_{N,1})</td>
<td>( \ldots )</td>
<td>( I_{N,1,1000} )</td>
<td>( \ldots )</td>
</tr>
</tbody>
</table>
Example: Agricultural Spectroscopy

Maize reflectance and fluorescence
Spectroscopy of maize

Distinguishing different types of maize
Distinguishing different types of hybrid rice for different levels of fertilizer administration
LED Multispectral microscopy malaria detection

Which red blood cells are invaded by malaria?

Brydegaard et al.
Spectrum of laser plasma

Target

Mobile lidar system

Laser beam path

Nd:YAG

OMA

Beam expander

To roof-top scanning mirror

Receiving telescope

Distinguishing different metals
Imaging LIBS, 60 m distance