The Abdus Salam
International Centre for Theoretical Physics

United Nations
Educational, Scientific and
Cultural Organization

IAEA
International Atomic Energy Agency

2017-8

Preparatory School to the Winter College on Optics in Environmental Science

26 - 30 January 2009

Multi-spectral Imaging Basics Part I: The use of multispectral imaging

Brydegaard M.

Lund Institute of Technology
Sweden

Multi-spectral Imaging Basics Part I: The use of multispectral imaging

Abdus Salam International Center for Theoretical Physics, 2009

Mikkel Brydegaard

Division of Atomic Physics

Lund University, Sweden

Today: Three steps to Multispectral Imaging:

Part I: The use of multispectral imaging

Part II: Spectroscopy, Physics and Acquisition

Part III: Color spaces, Data handling and Contrast functions

Lund University

- Lund University founded 1666
- Applied physics since 1962
- Largest university in Scandinavia

Applied Molecular Spectroscopy & Remote Sensing Atomic Physics

http://www-atom.fysik.lth.se/AMSRS/

- 40.000 under graduates
- 3.000 PhD students
- 40 researchers at Atomic Physics Division
- 3 PhD students at Applied Molecular Spectroscopy and Remote Sensing

Prof. Sune Svanberg

Gabriel Somesfalean

Märta Lewander

Mikkel Brydegaard

Zuguan Guang

Thomas Svensson

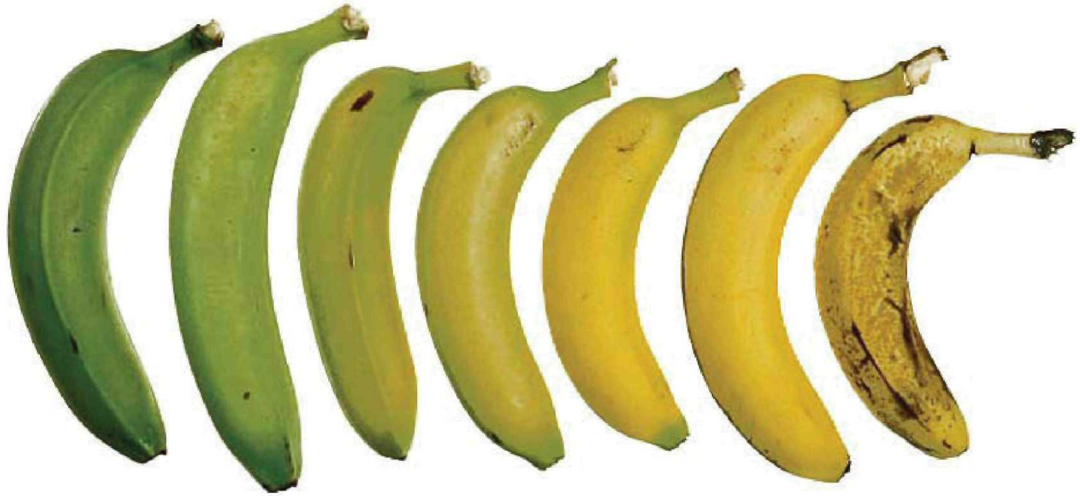
Benjamin Andersson

Hiran Jayaweera

Aboma Merdasa

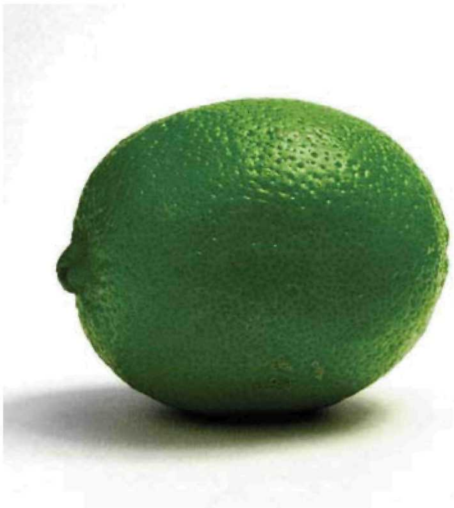
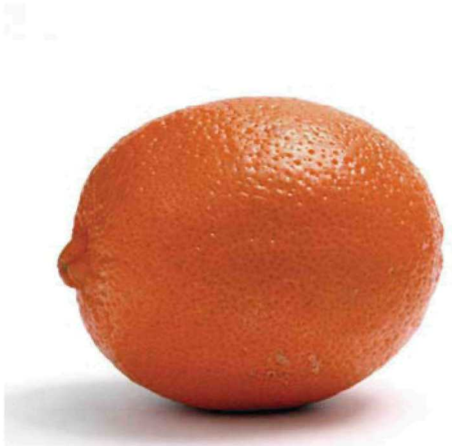
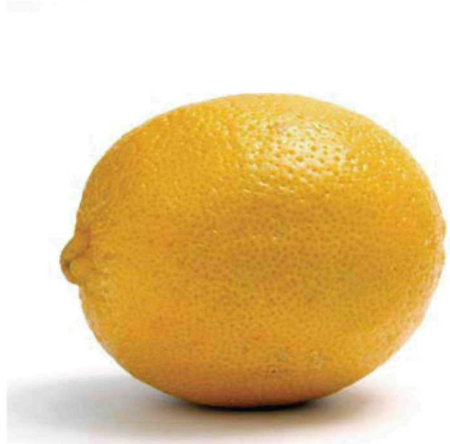
Patrik Lundhin

Spectroscopy for quality control



Which banana is the most tasty?

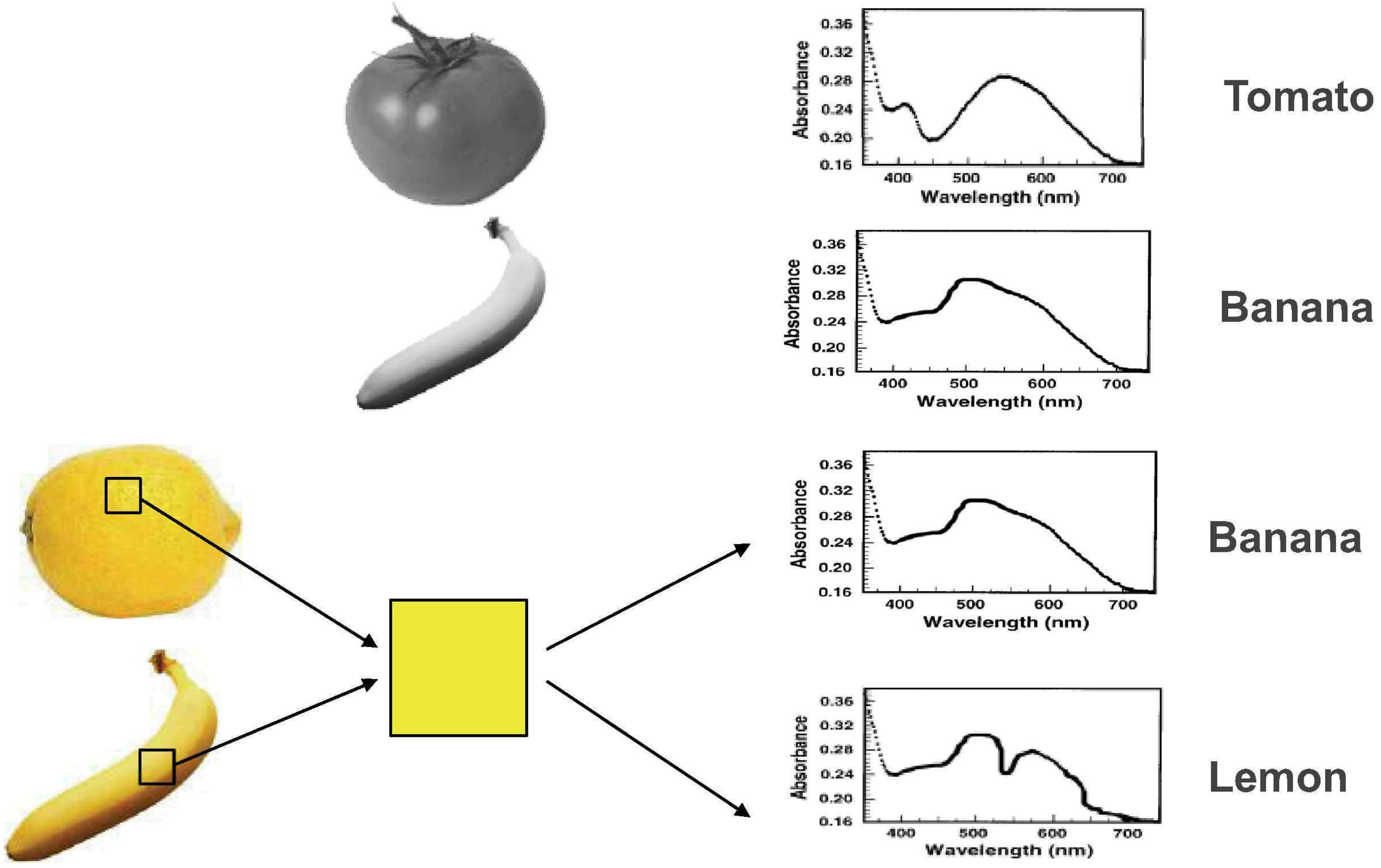
Spectroscopy for object identification

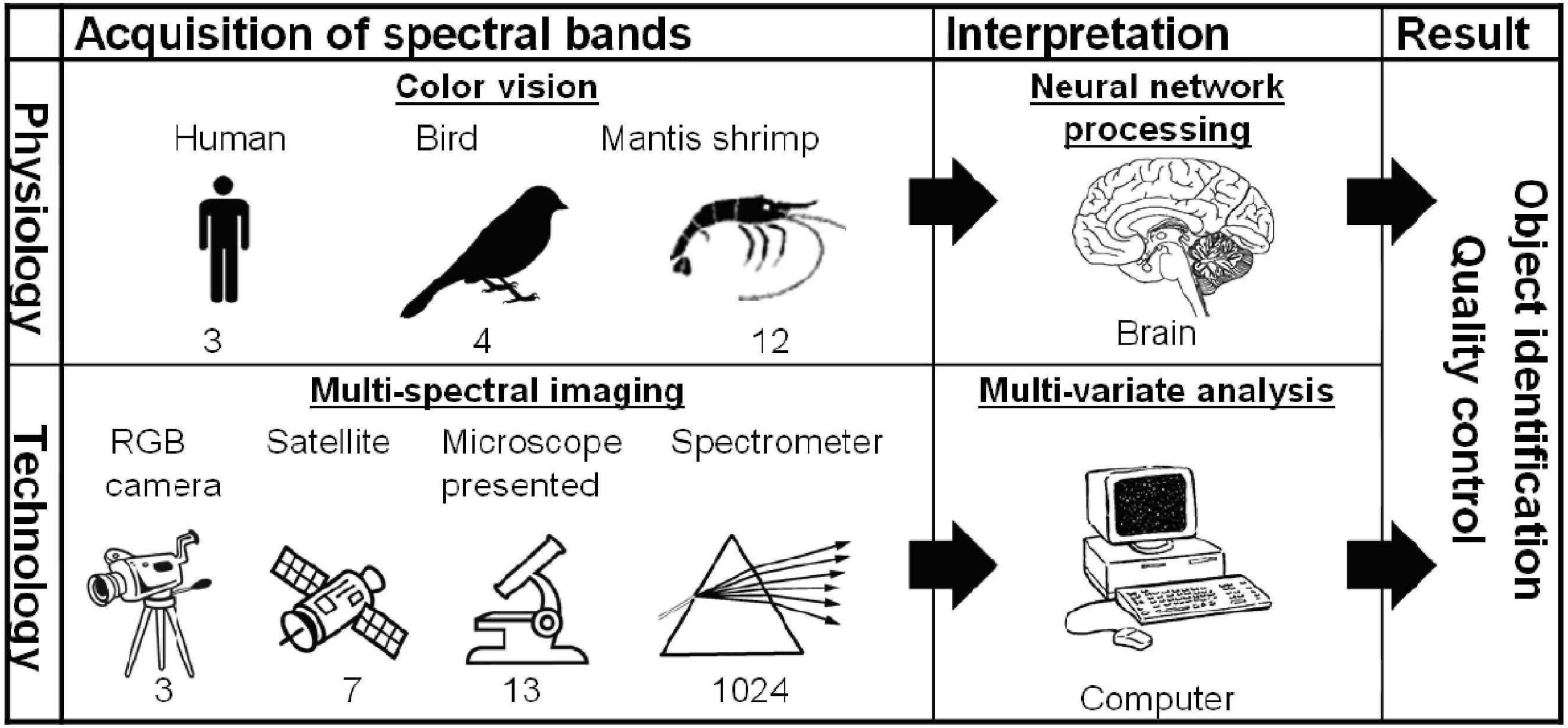
Lime, Orange and Lemon. Which is which?

Spatial signatures and spectral signatures

Spatial identification | **Spectral identification**



What to do with spectroscopy?



M. Brydegaard, AJP, 2009

Light environment and evolution of senses

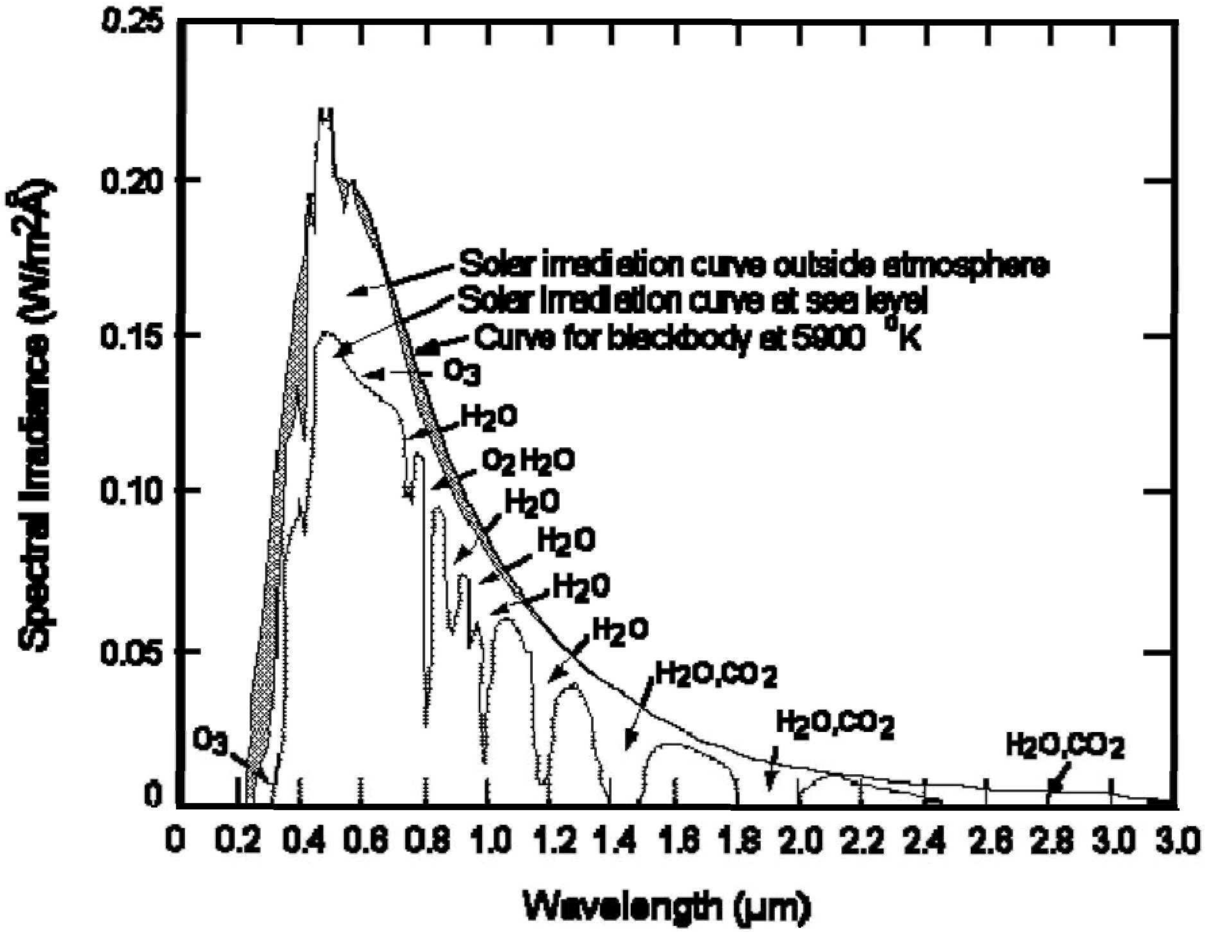
- **Black body radiation**
- **Planck's law of radiation**

$$I(\nu) = \frac{2h\nu^3}{c^2}\frac{1}{e^{\frac{h\nu}{kT}} - 1}$$

- **Differentiation leads to Wien's displacement law**

$$\lambda_{max} = \frac{b}{T}$$

- **The sun and black body radiation**
- **The solar radiation peaks around 550nm**

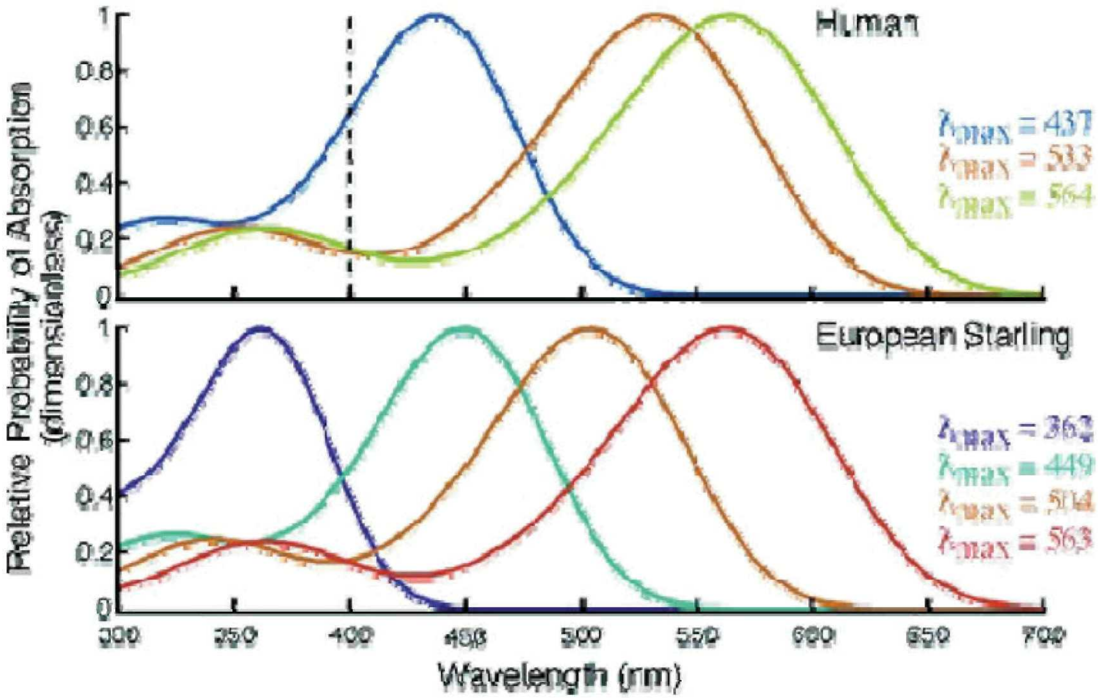


http://www.csr.utexas.edu/projects/rs/hrs/pics/irradiance.gif

Light environment and evolution of senses

- Evolution of the human eye
- Other species
- Co evolution
- Contribution I_w to a color channel w given the emission spectra E(λ) and the sensibility spectra for the colour channel $S_w(\lambda)$

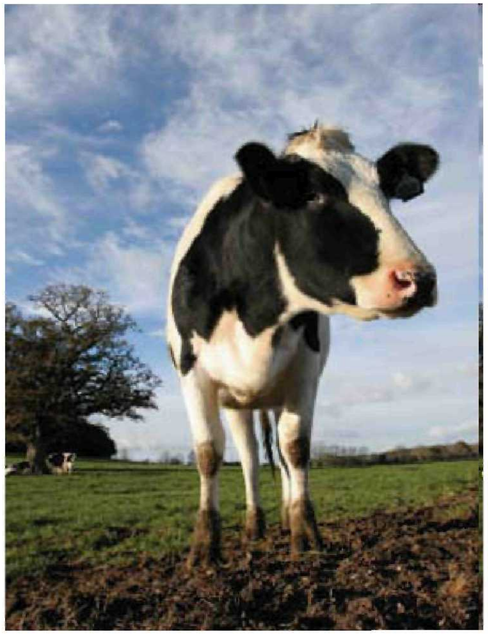
$$I_w = \int_0^{\infty} S_w(\lambda) * E(\lambda)\, d\lambda$$



Spectral world perception

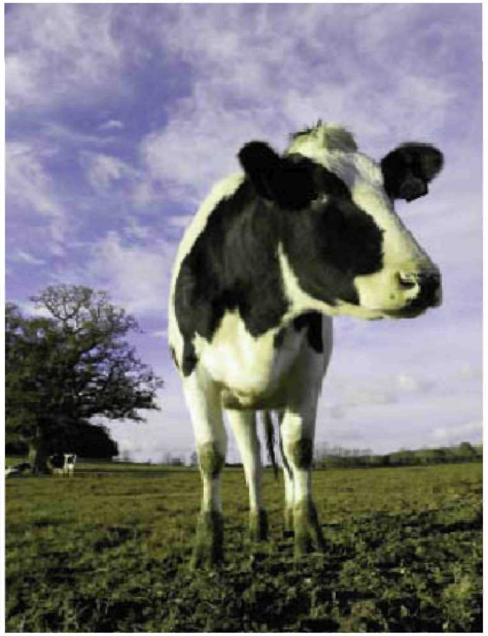
<u>Spectral bands:</u>

3



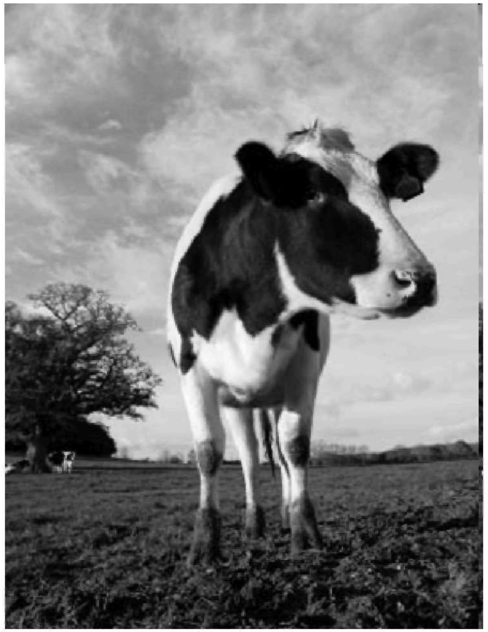
Human looking at cow

2



Cow looking at cow

1



BW observation of cow

4

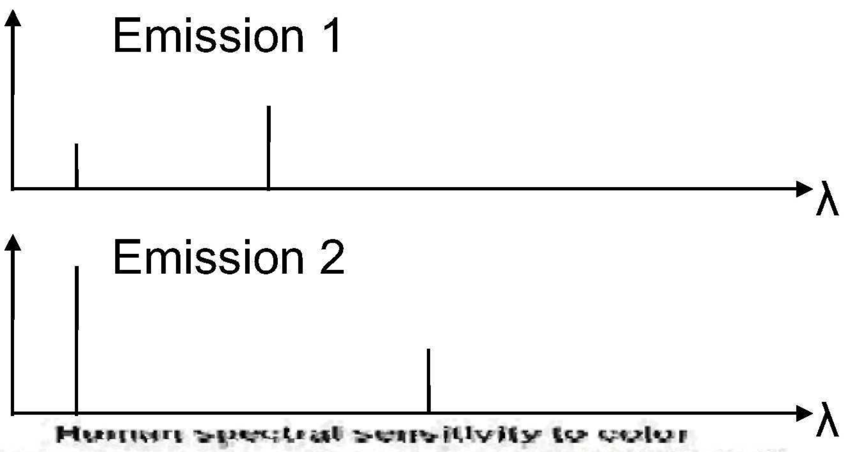
?

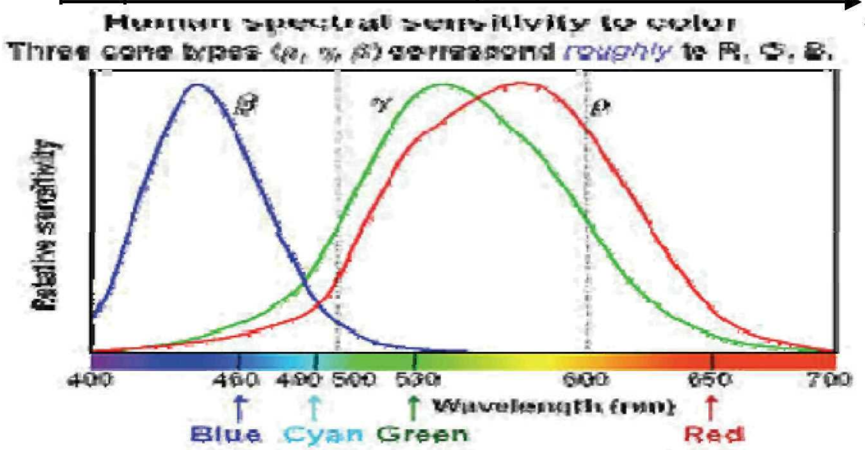
Cannot be communicated to humans* :-(

Bird looking at cow

Limitation of human spectral resolution

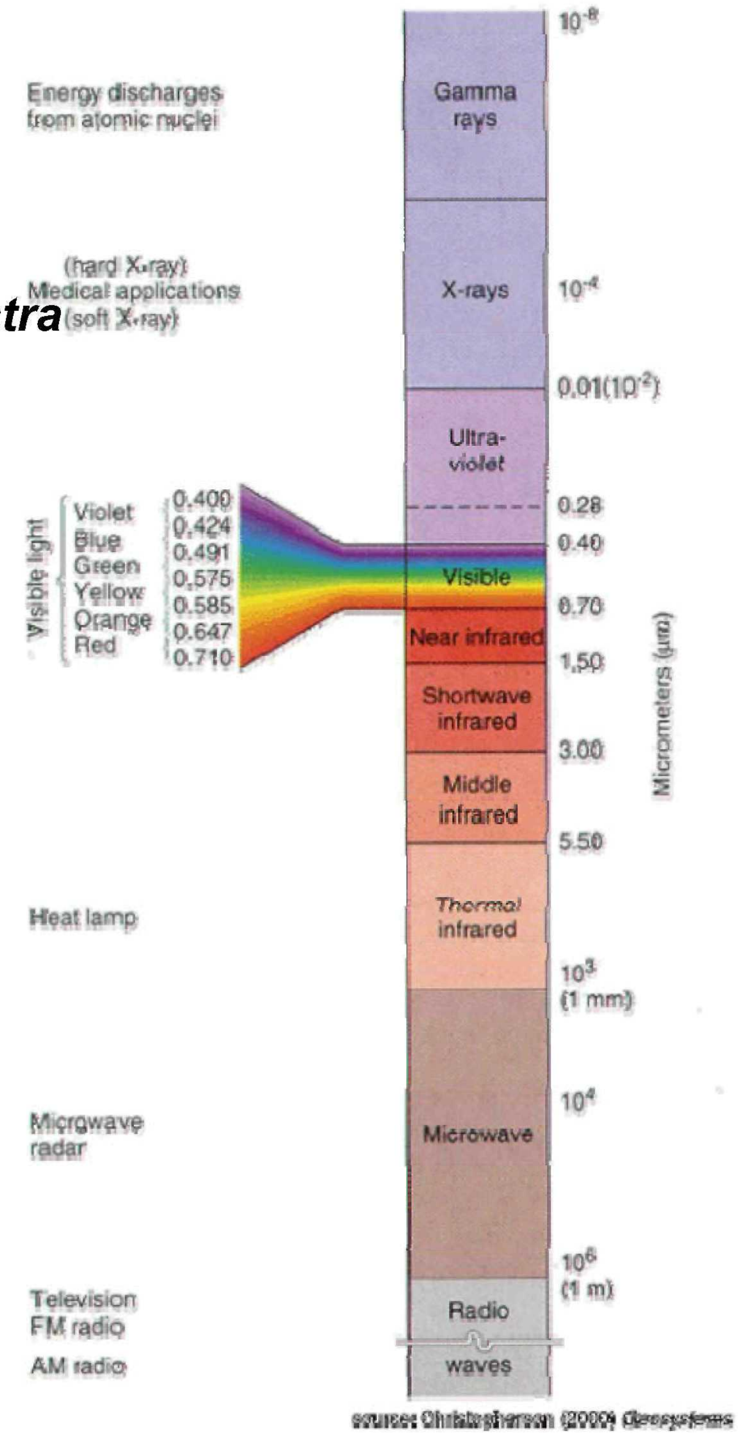
Consider two object with different emission spectra





RGB $_{\text{emission 1}}$ = [1*0+1*1 , 1*0+1*1 , 1*1+1*1] = [1 , 1 , 2]

RGB $_{\text{emission 2}}$ = [0*2+2*1 , 0*2+2*1 , 1*2+1*1] = [1 , 1 , 2]

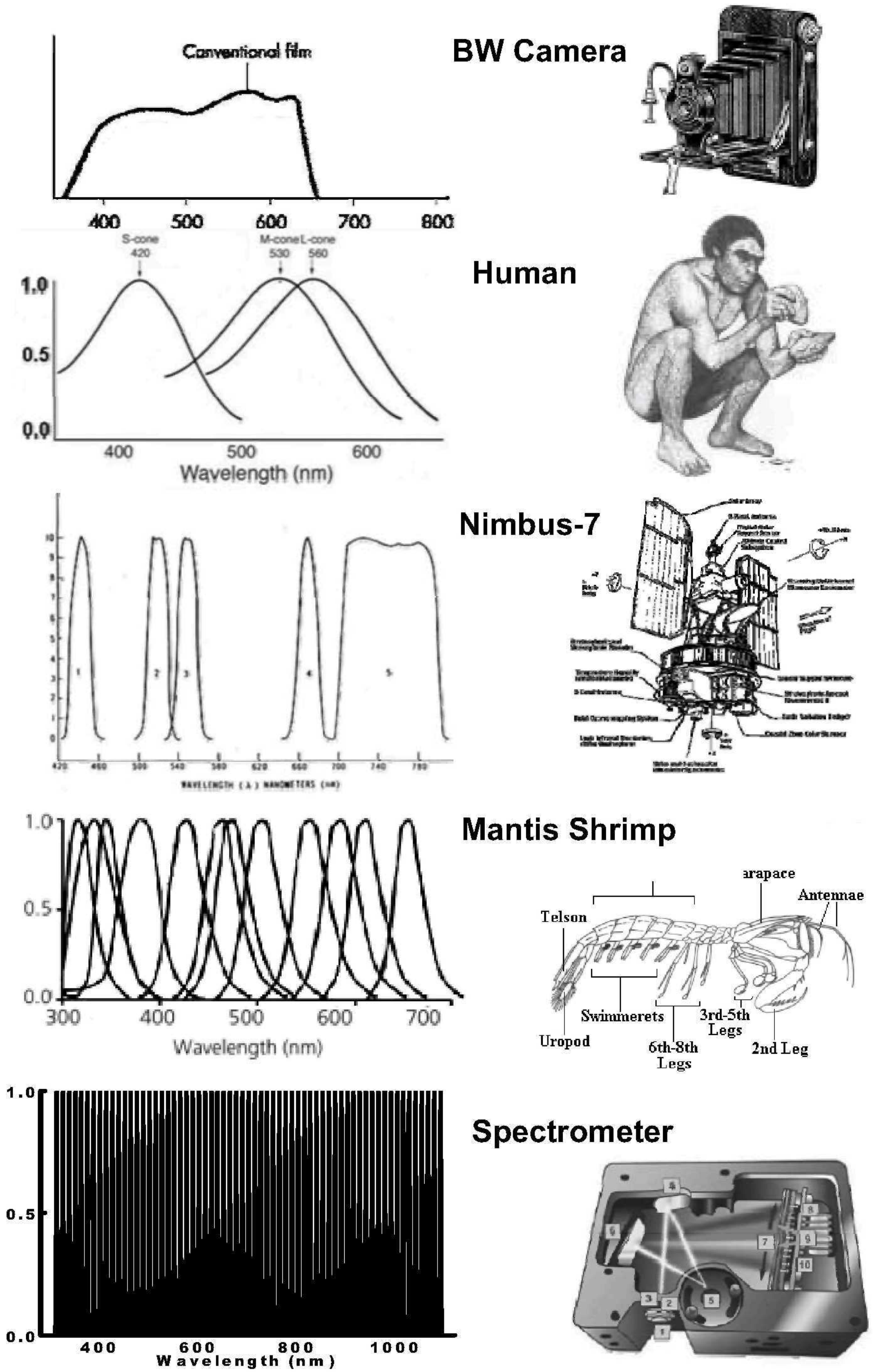


#1

Comparison of spectral bands of various imaging systems

- Spectral content is discretized by a number spectral bands.
- Bands can vary in center position, shape and width
- Paradigm can be applied to all spectroscopic methods, including laser spectroscopy
- Contribution to a band is the integral of the product of sensitivity, reflectance and illumination spectrum.

$$U_{channel(k)} = \int_0^{\infty} E_{(\lambda)} R_{(\lambda)} S_{k(\lambda)} d\lambda$$



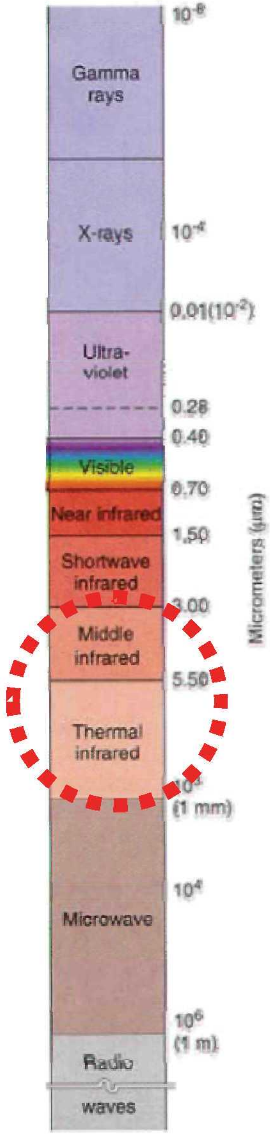
Examples of what the human eye cannot see, IR

Leaking gas from a burnt out vehicle is made visible!

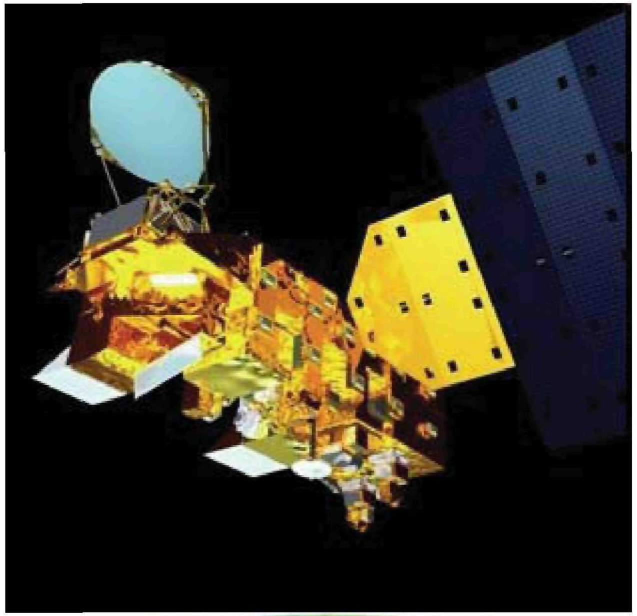
J. Sandsten, Opt. Expr. 2004

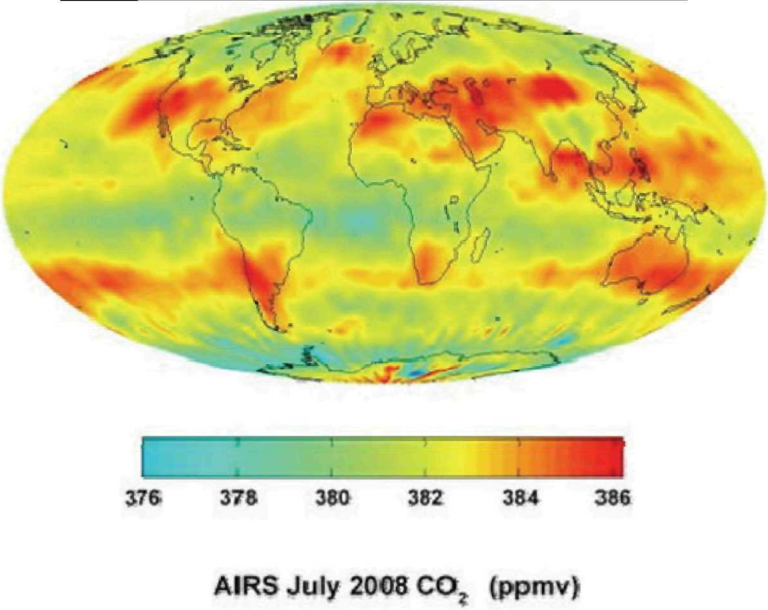
Blackbody, transmission and absorbance.

NASA

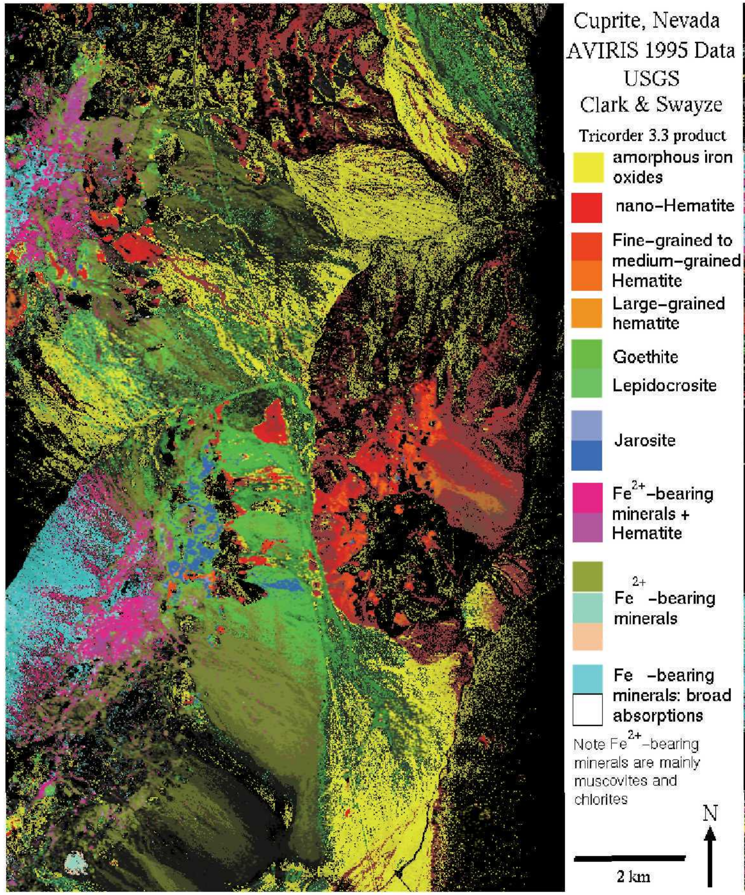


Examples of what the human eye cannot see, IR



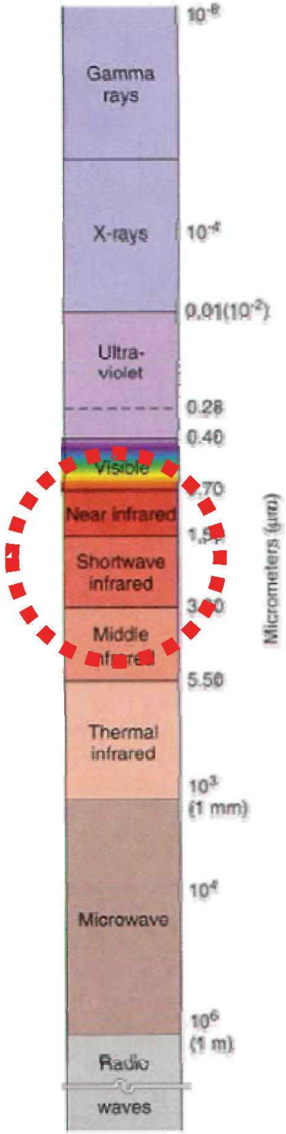


Global concentration of CO_2 measured in the IR!



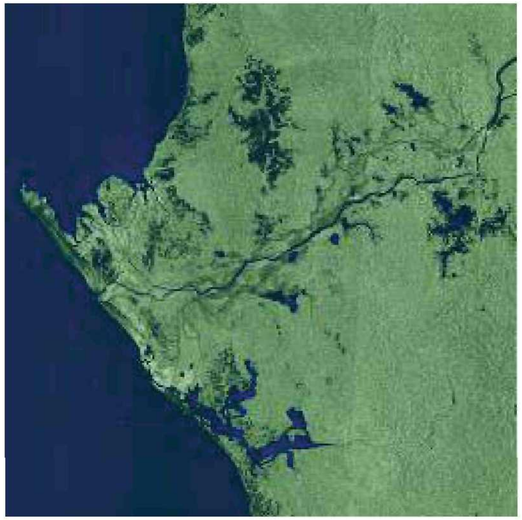
Mineral mapping from satellite covering 430nm-3µm wavelength

http://airs.jpl.nasa.gov/
http://m3.jpl.nasa.gov/

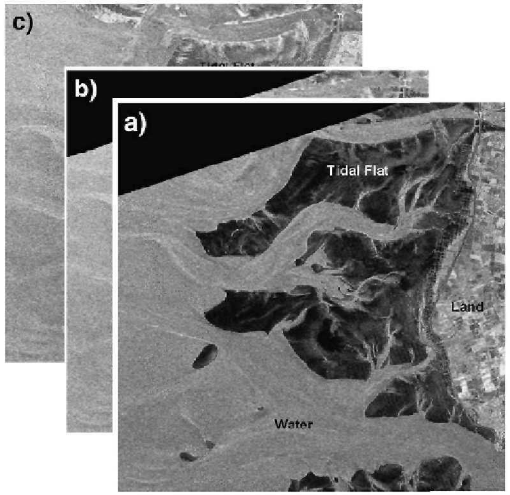


Examples of what the human eye cannot see, micro waves

Gabon at visible wavelength



Gabon at radar wavelength



L, C & X band SIR/SAR radar images of Wade Sea. M. Gade et Al.

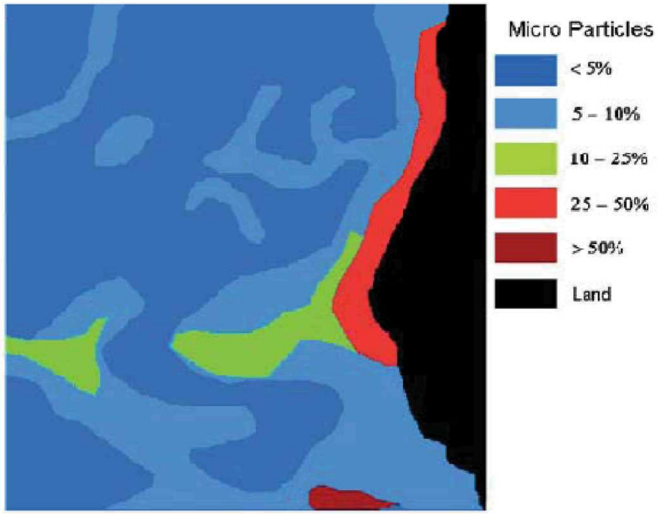
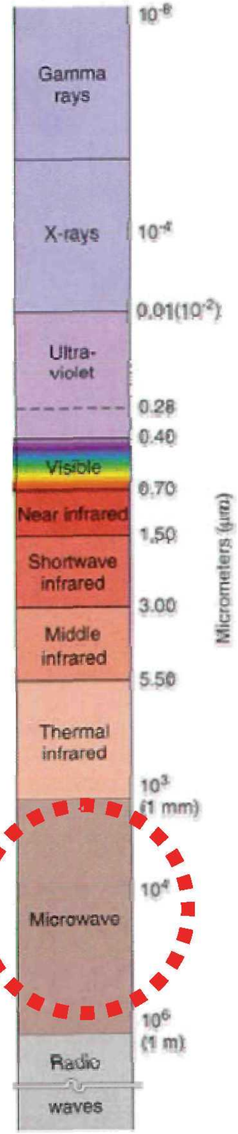


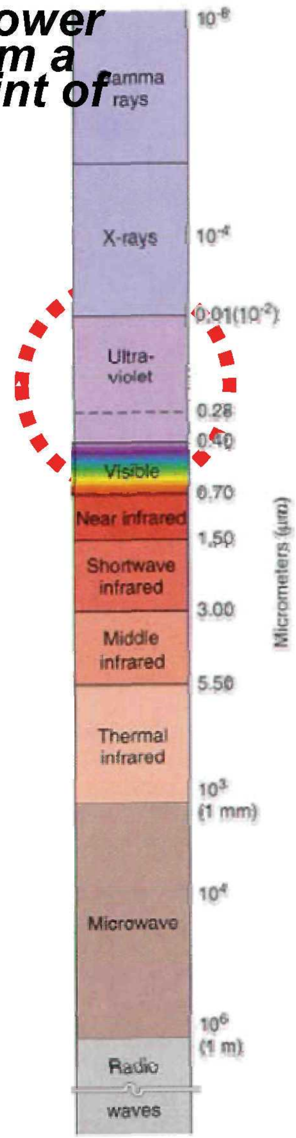
Fig. 10. Sediment classification provided by the Schleswig Holstein Wadden Sea National Park Office. The color coding denotes the percentage of micro particles (i.e., particles with diameters less than 63 μm).



Examples of what the human eye cannot see, UV

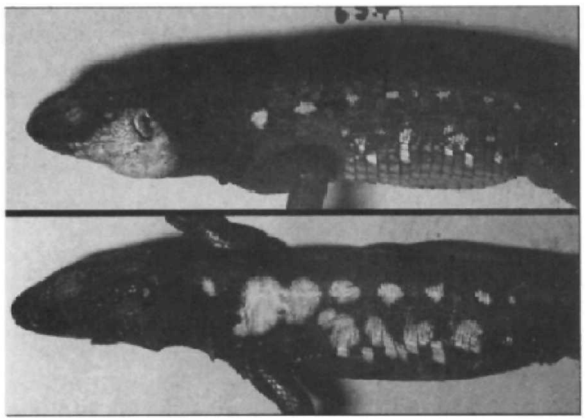
Visible Appearance | Spurious/false color | Recorded by digital camera through 18A UV filter | Greyscale, UV, record

Butter flower seen from a bees point of view

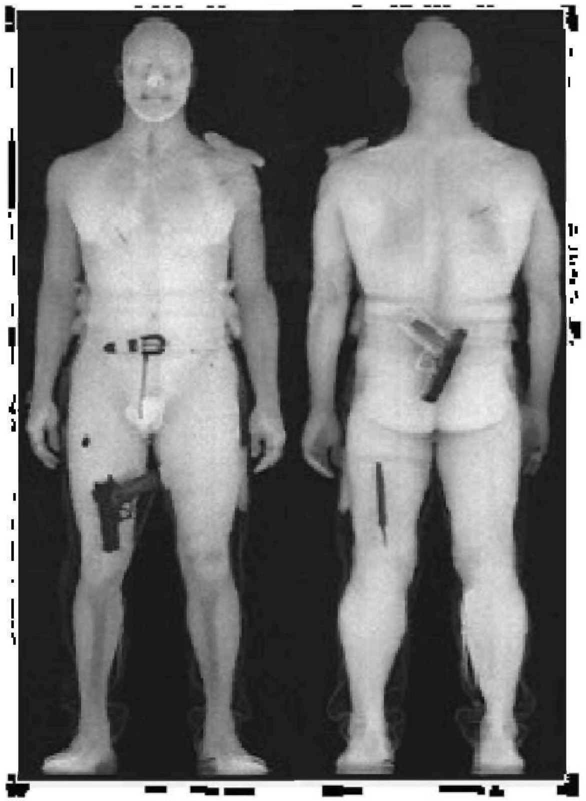


UV, Changing appearance in the visible

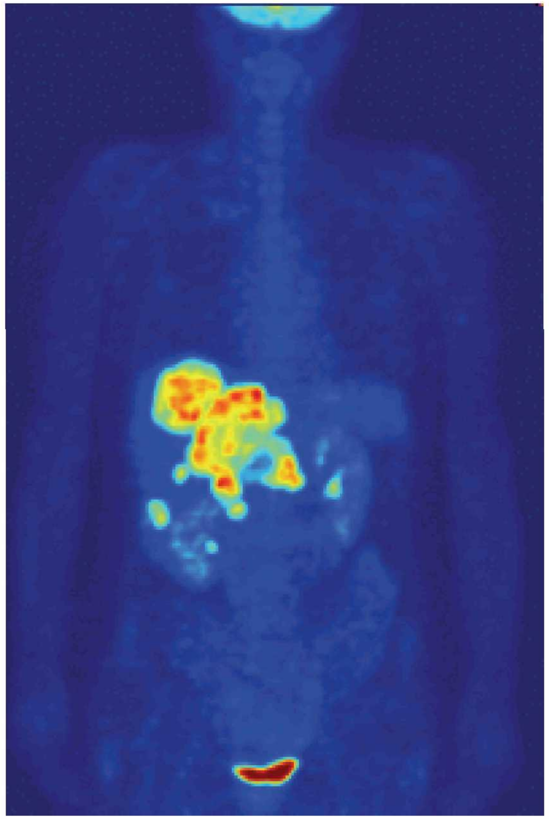
Invisible UV markings on Tenerife lizards



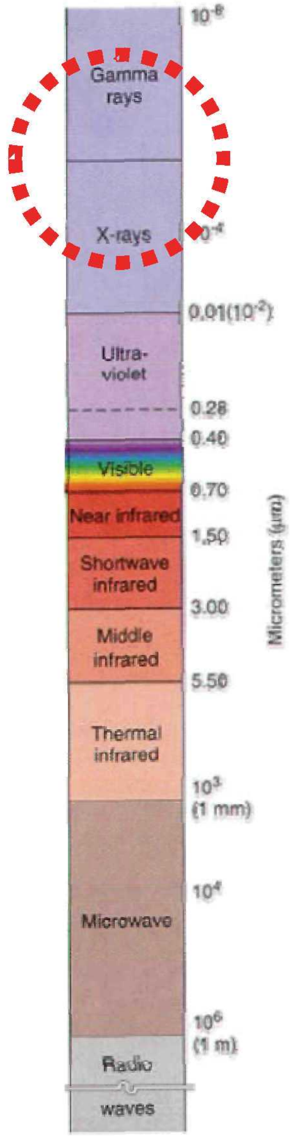
Examples of what the human eye can not see, gamma ray



Backscatter X-ray in use in Amsterdam airport



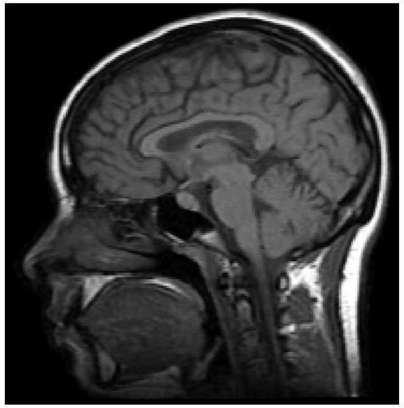
Positron Emission Tomography (PET) for drug tracing



http://en.wikipedia.org/wiki/Backscatter_X-ray

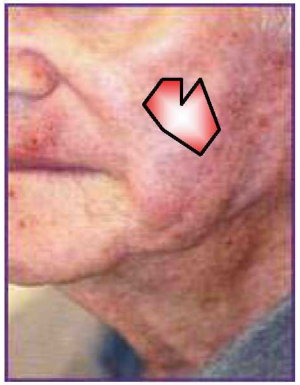
http://en.wikipedia.org/wiki/Positron_emission_tomography

Examples of what the human eye cannot see in medicine

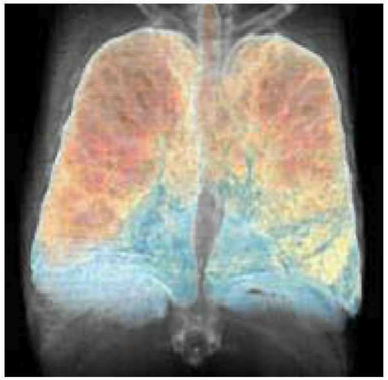


MRI

Thermography

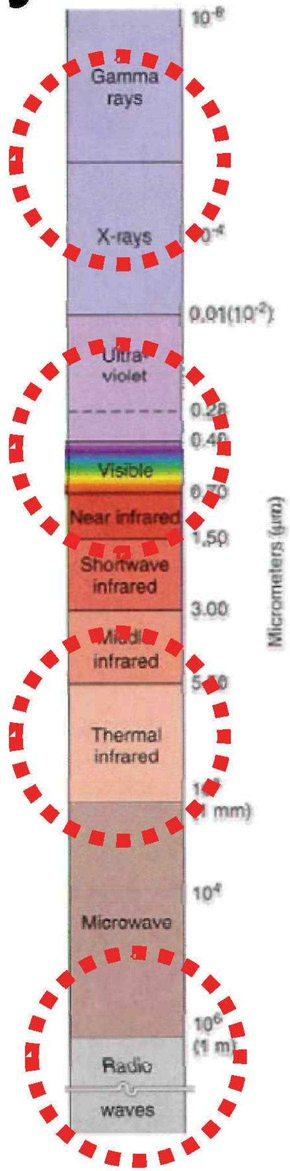


Photosensitizers

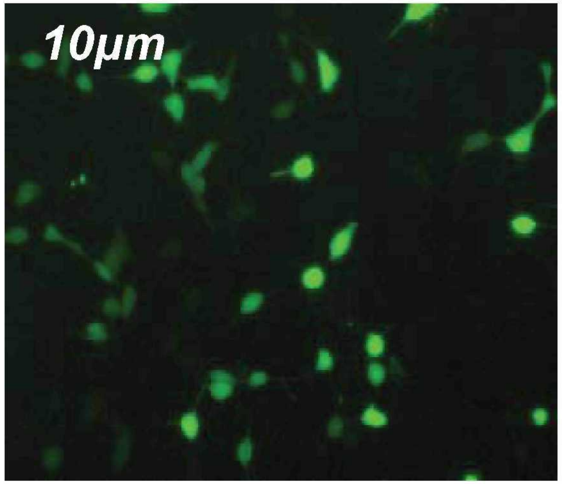


Multi-spectral X-ray

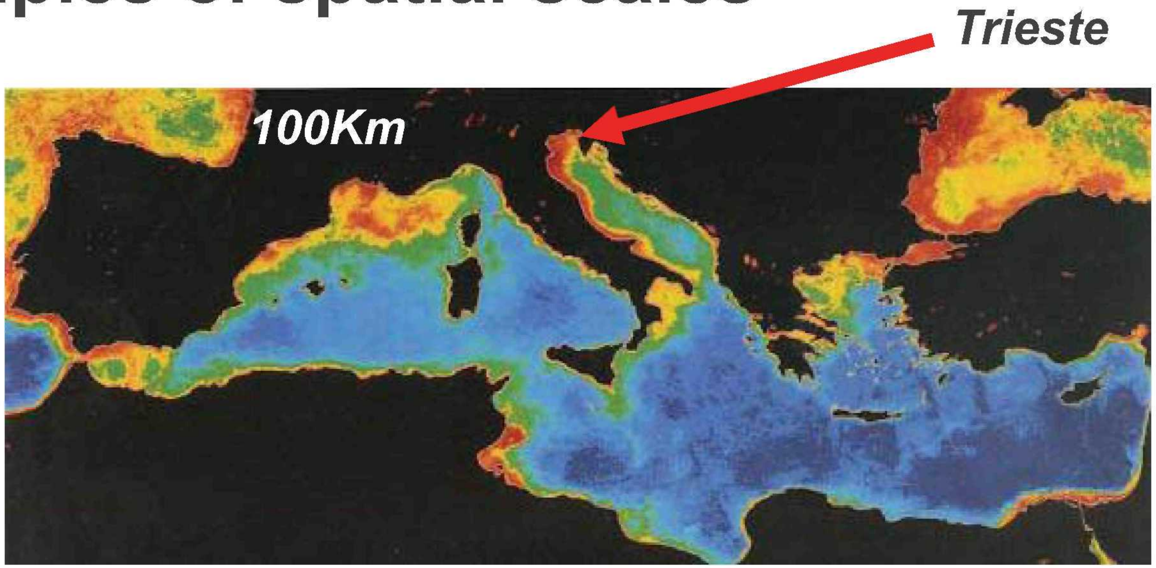
RGB Pill cameras



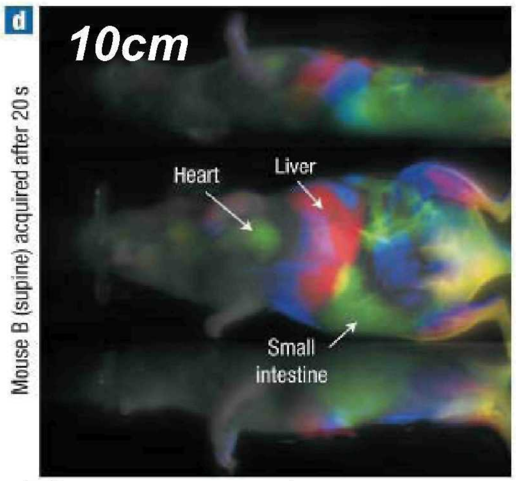
Examples of spatial scales



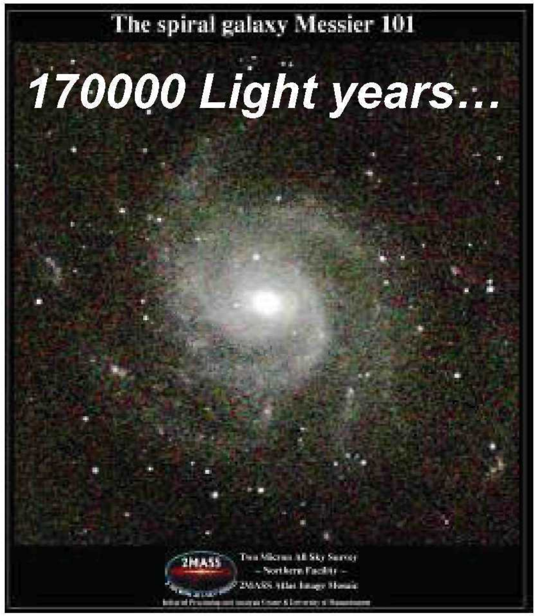
Nerve pulses observed with Ca++ fluorescence imaging



Environmental monitoring of the Mediterranean sea



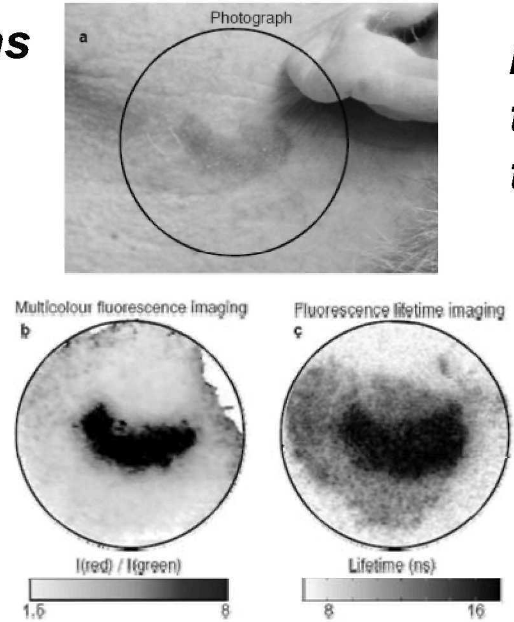
Fluorescent mouse organs, Hillman Et al. 2008



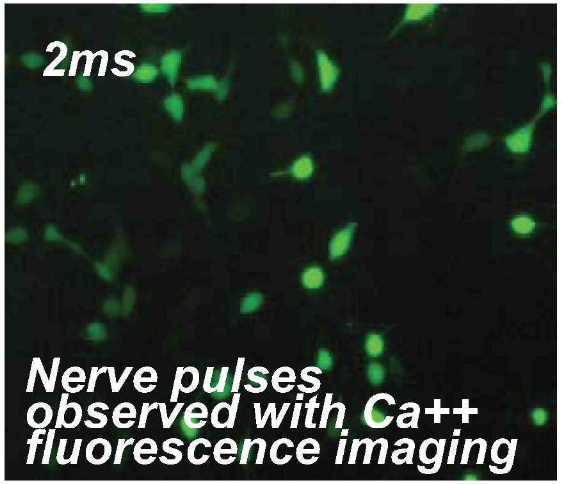
NIR image of Messier 101 spiral galaxy

Examples of temporal scales

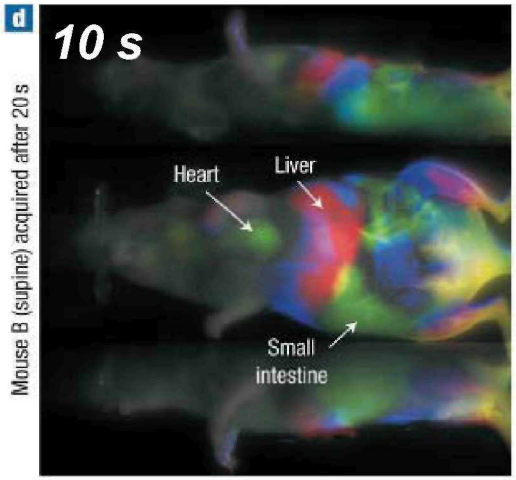
2ns



Fluorescence life time imaging for tumor detection

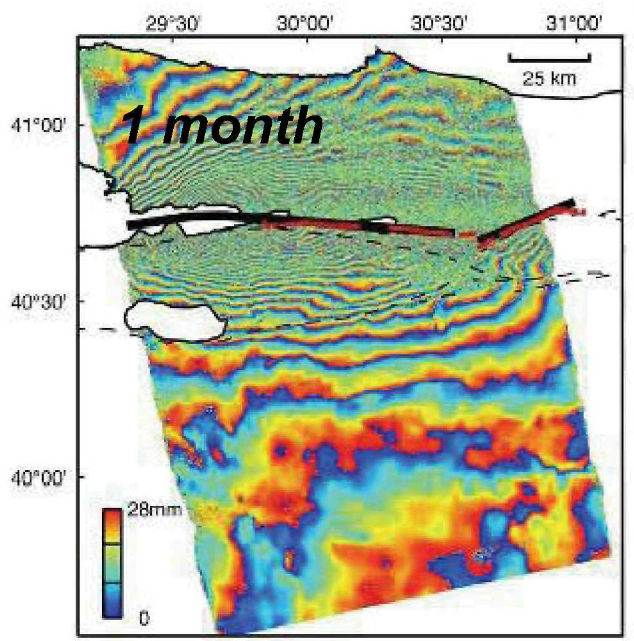


S. Andersson-Engels



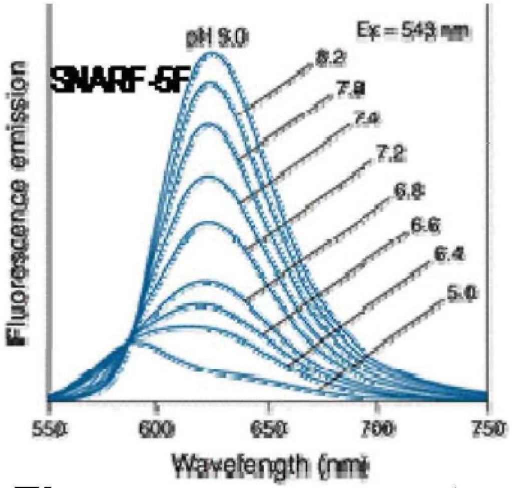
Fluorescent mouse organs, Hillman Et. al.

Turkish earth quake, Izmit 1999

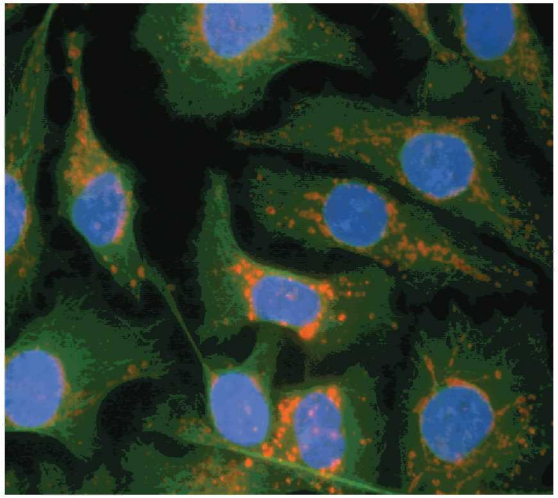


http://en.wikipedia.org/wiki/Interferometric_synthetic_aperture_radar

Indirect photosensitizers and bio-markers

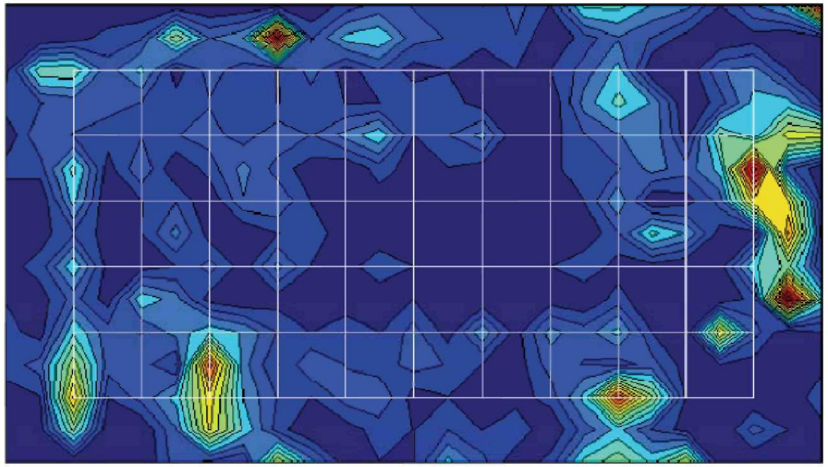


Fluorescence spectra depending on the pH



pH distribution in cells

Trained honey bee



Bee concentration and landmine distribution

J. Shaw, Opt. Expr. 2005

Multi spectral imaging in Lund

— From Astronomy to Microscopy

- From Radiowaves to Gamma rays

- **Gas correlation**
- Absorption
- Multiple narrow spectral lines
- Perfect filter fitting

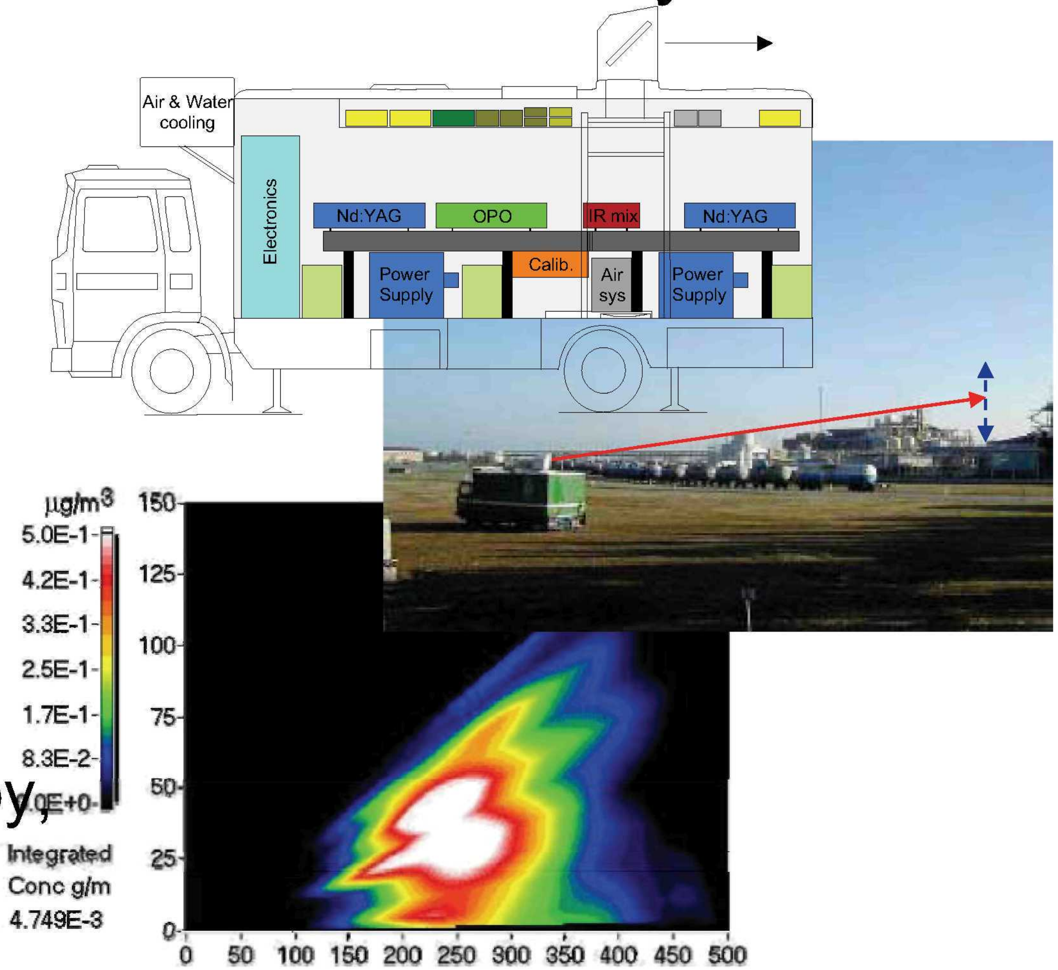
GasOptics Sweden AB

Multi spectral imaging in Lund

— From Astronomy to Microscopy

- From Radiowaves to Gamma rays

- **Lidar**
- DIAL, differential absorption lidar, narrow lines
- LIF, Laser Induces Fluorescence, broad band
- LIBS, Laser Induced breakdown Spectroscopy, elemental lines



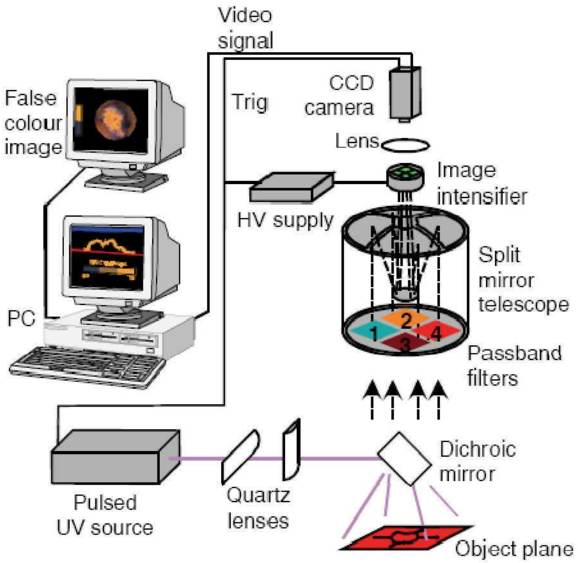
http://www-atom.fysik.lth.se/AMSRS/

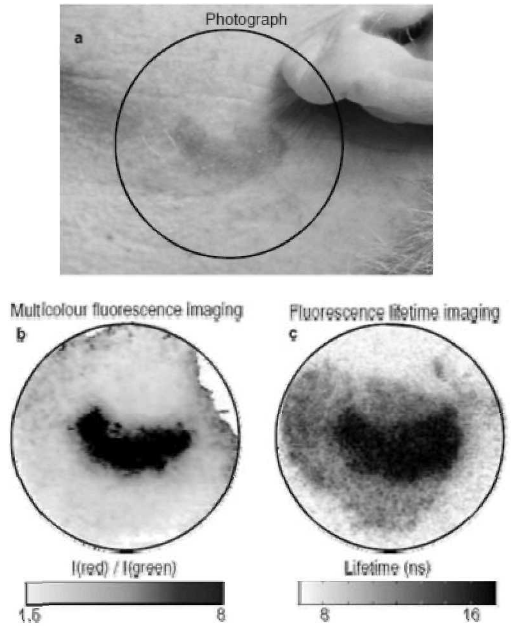
Multi spectral imaging in Lund

– From Astronomy to Microscopy

- From Radiowaves to Gamma rays

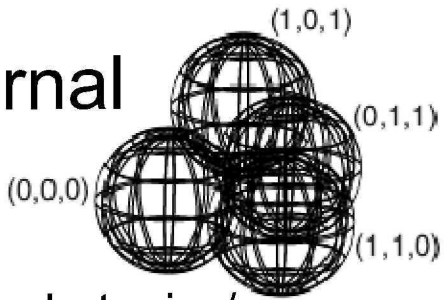
- **Tissue optics, Fluorescence imaging, DOT…**

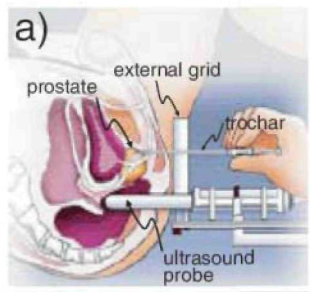


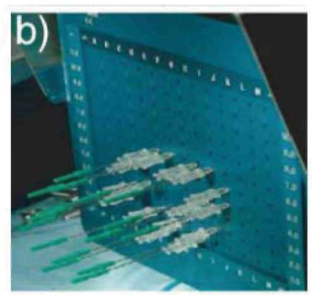


- Diagnostics and threatment

- Internal and external







http://www.atomic.physics.lu.se/biophotonics/

The ultimate multi spectral imaging conditions

- All super heroes uses multispectral imaging
- Infinitive spatial resolution
- Complete spatial dimension coverage, 3D
- Infinitive spectral resolution
- Infinitive temporal resolution
- Complete spectral region coverage

Multi-spectral Imaging Basics Part II: Spectroscopy, Physics and Acquisition

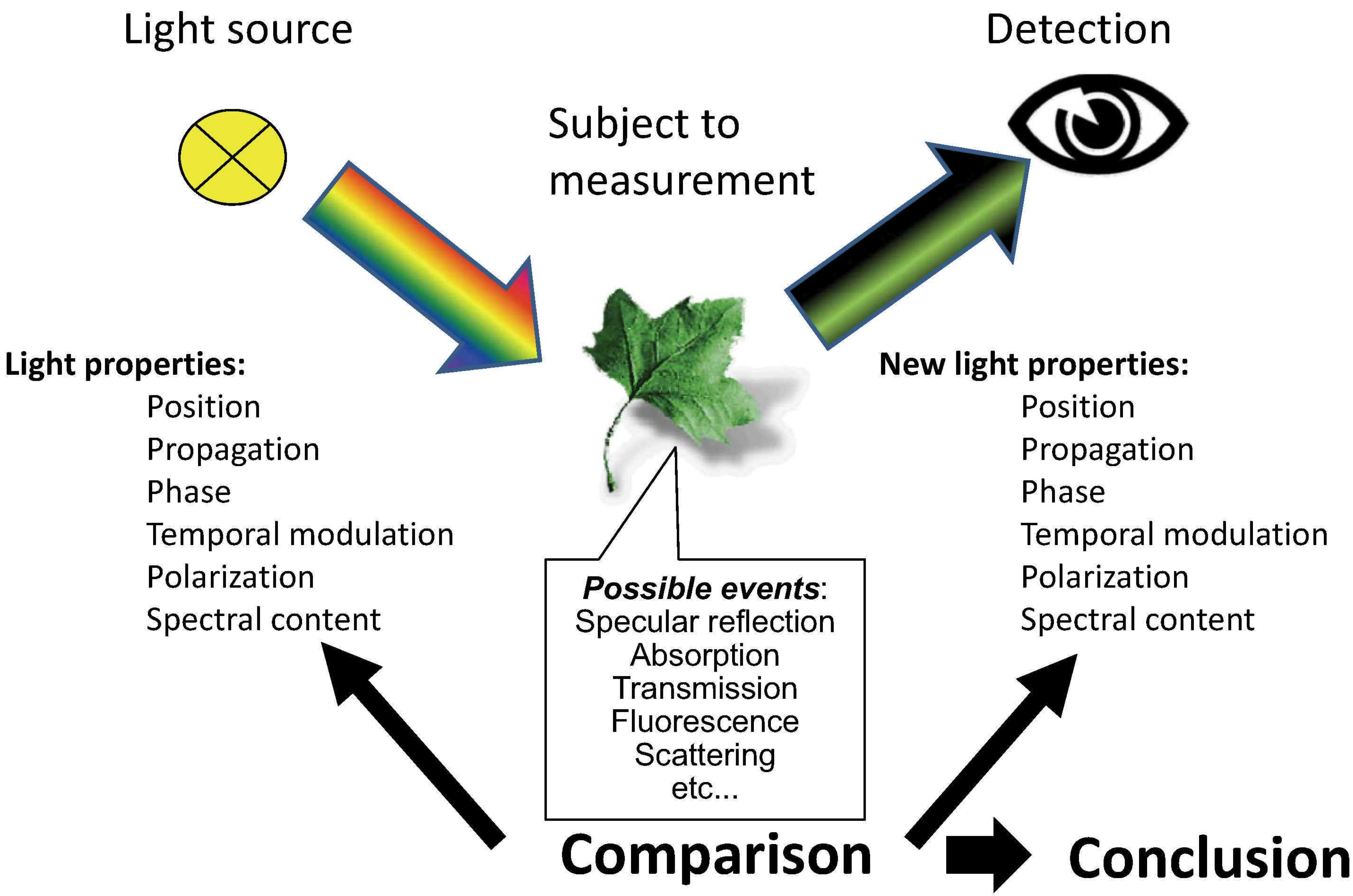
Abdus Salam International Center for Theoretical Physics, 2009

Mikkel Brydegaard

Division of Atomic Physics

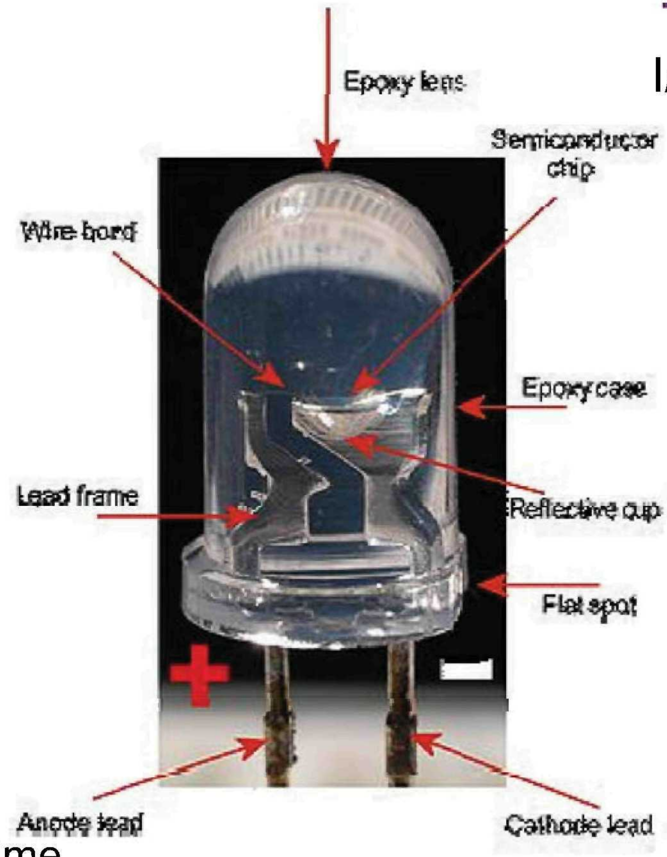
Lund University, Sweden

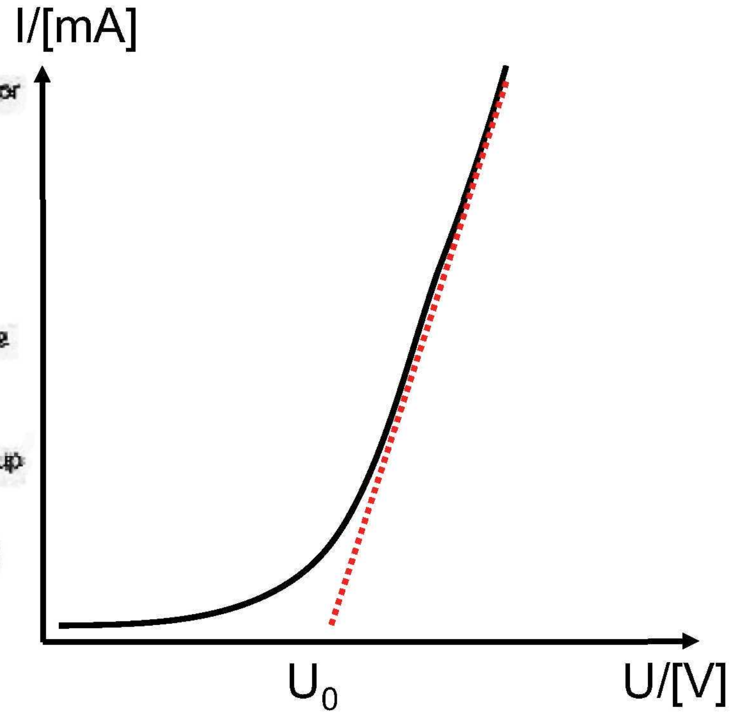
Measuring with light



The Light Emitting Diode

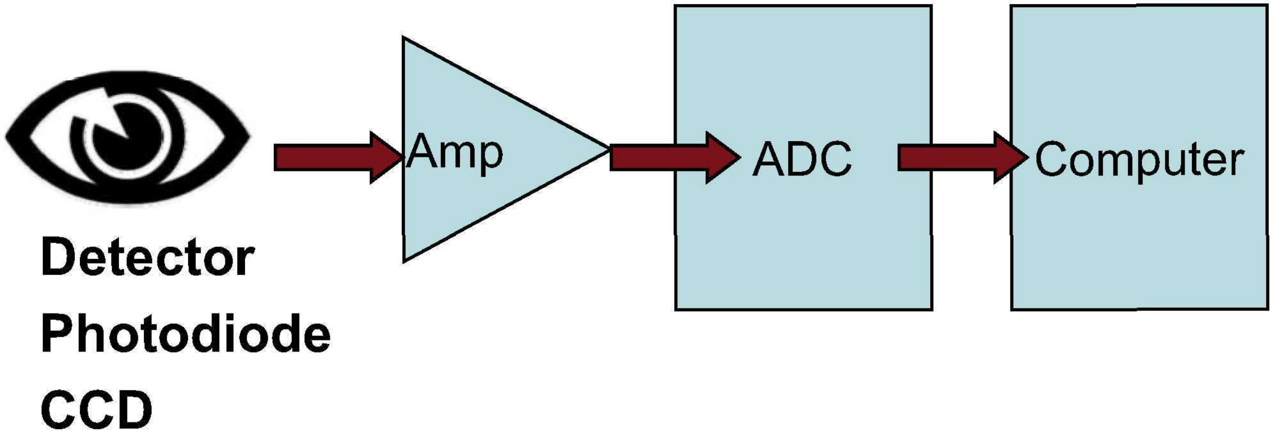
- IU characteristic
- Band gap, peak emission, U_0
- FWHM
- Divergence
- Current, intensity
- Pulsed operation
- Temperature
- Emissive yield
- Modulation, rise time and capacitances

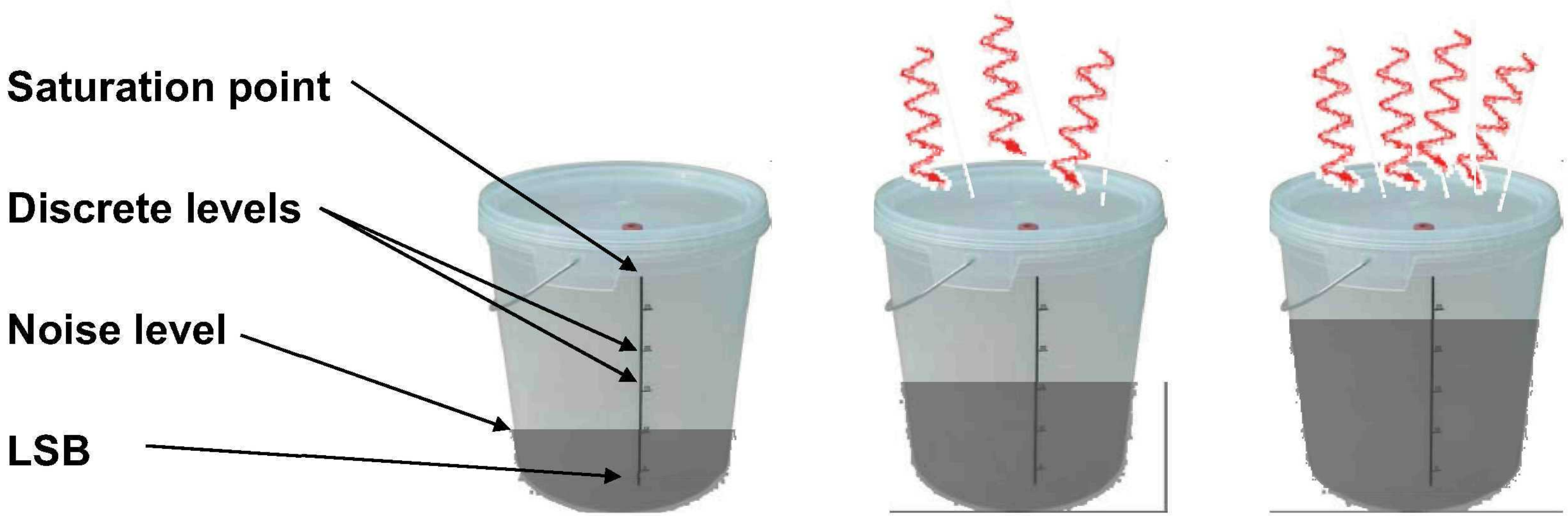




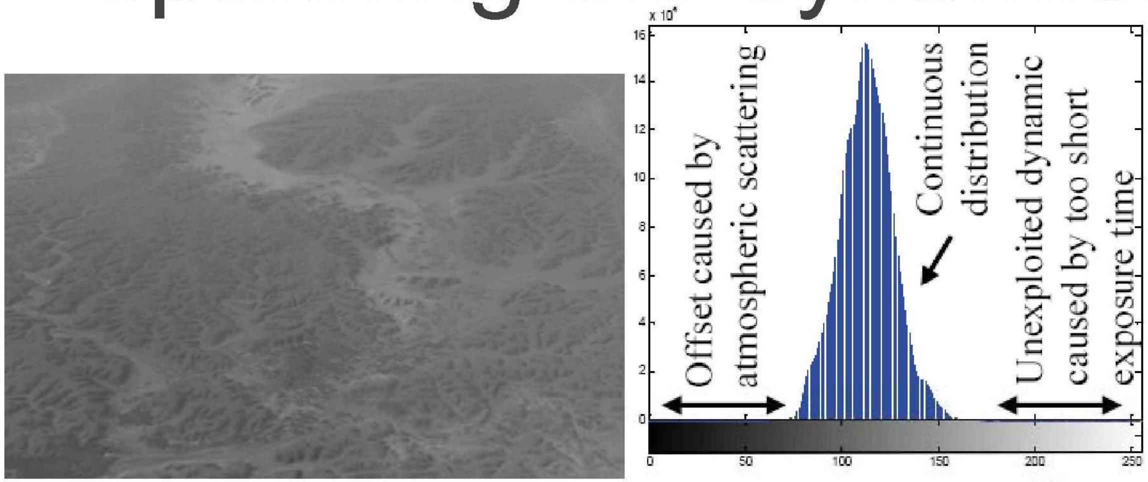
Detectors and dynamic range

- Light intensity measurements
- Light descretized by photons and bits
- Upper and lower limits
- Noise and uncertainty





Intensity histograms, spanning the dynamics.



(a) (b)

Figure 20 (a) Original image of Saharan city from an airplane (b) its histogram

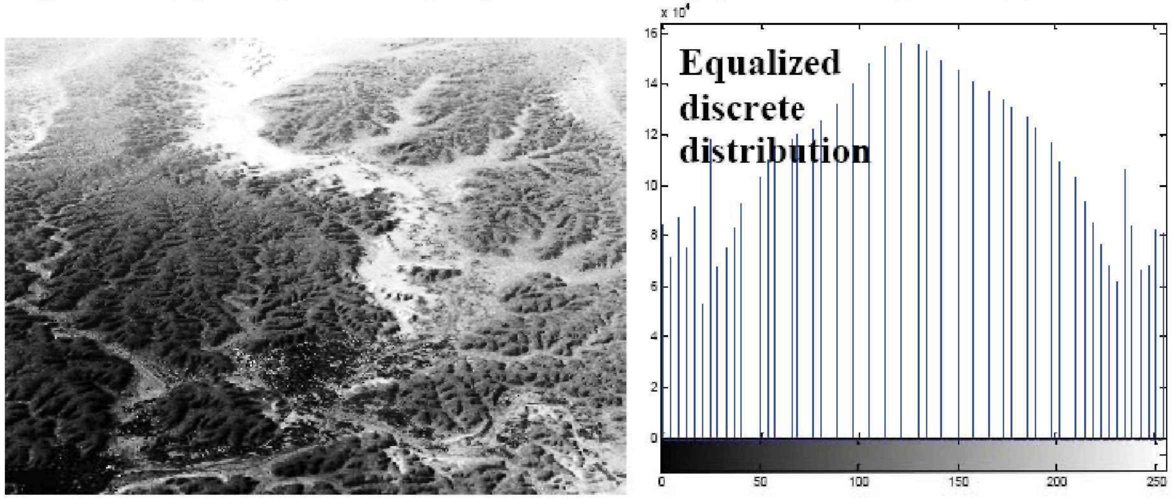
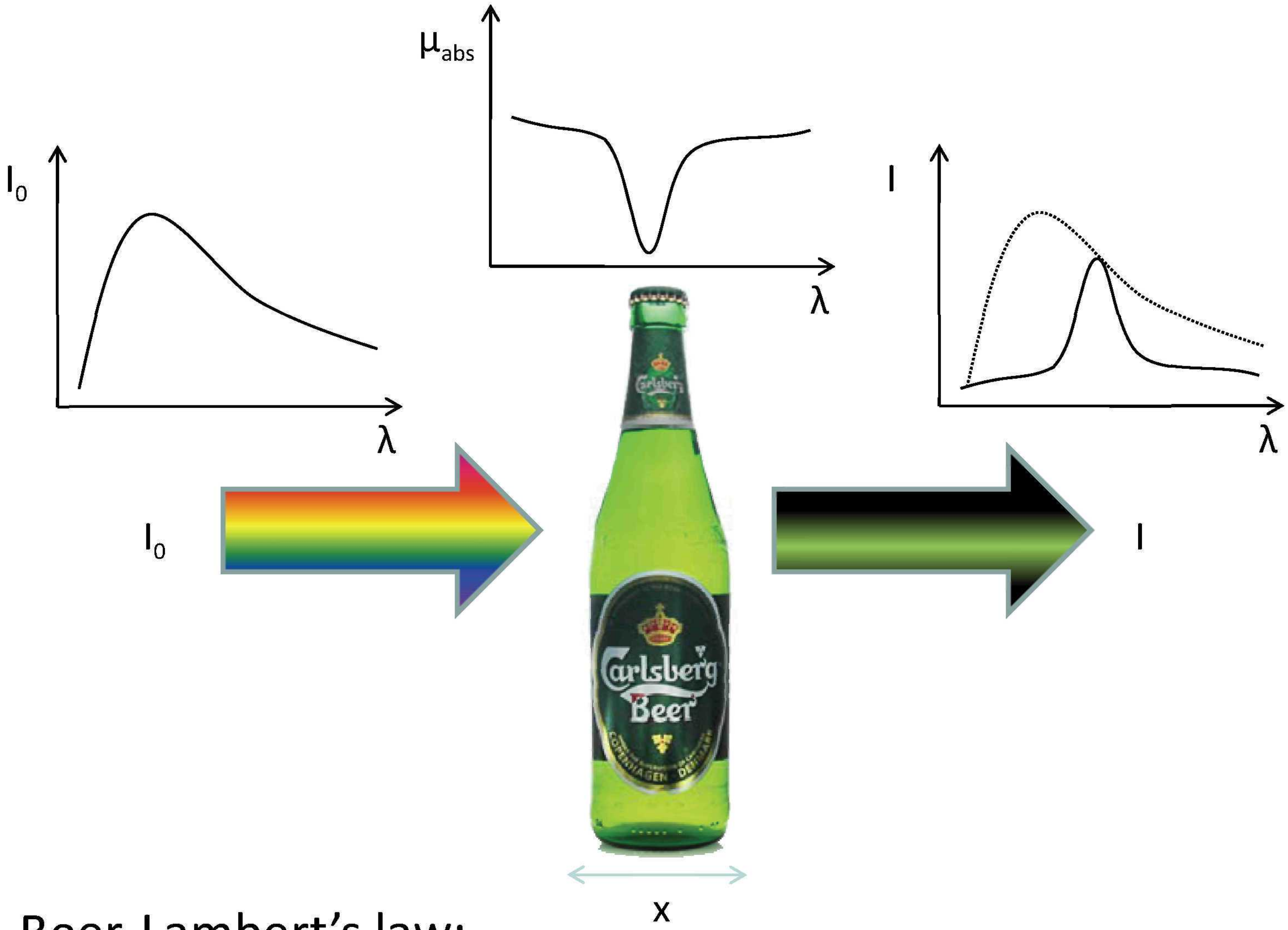


Figure 21 (a) Image subjected to histogram equalisation and (b) its histogram.

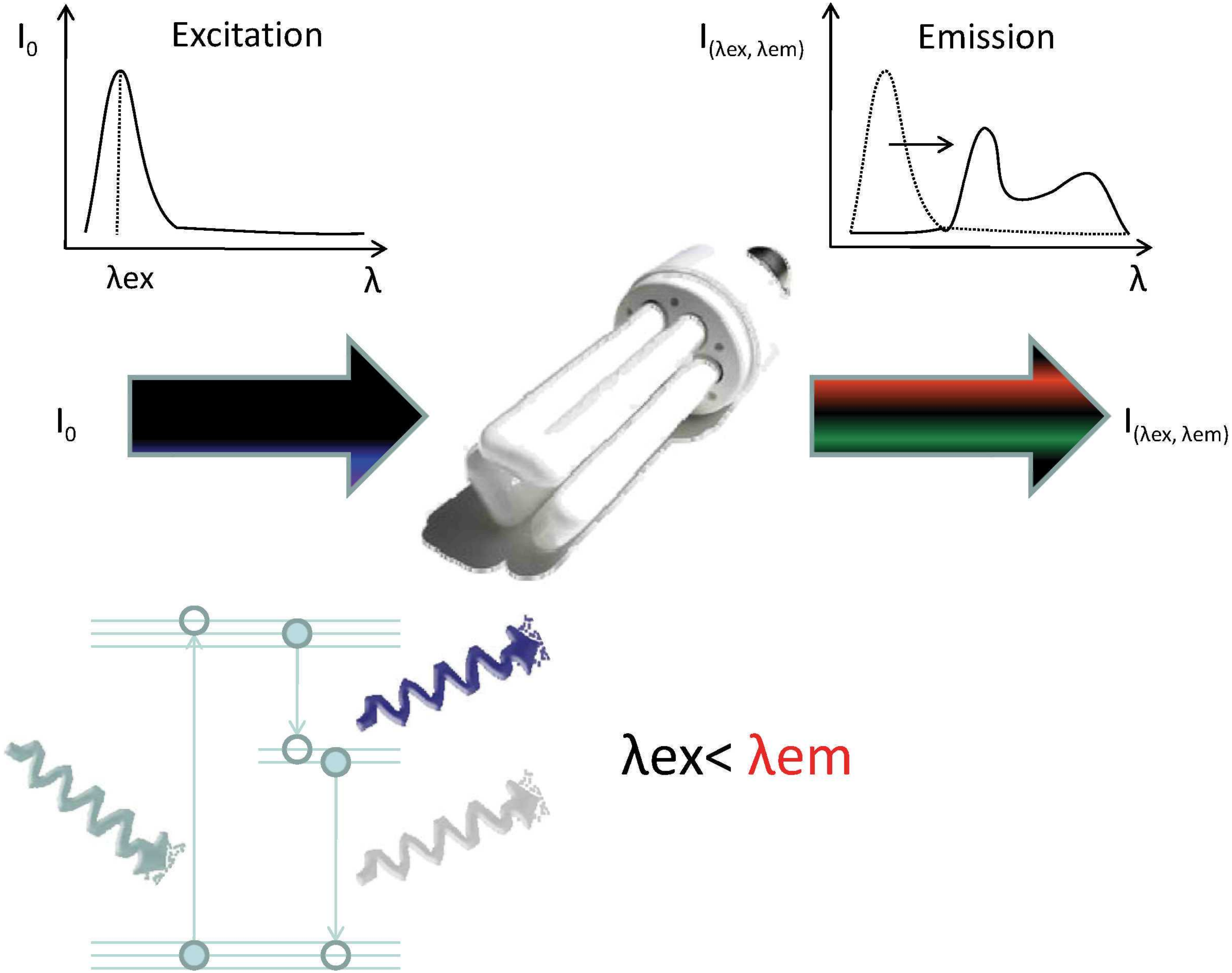
Absorbing volume



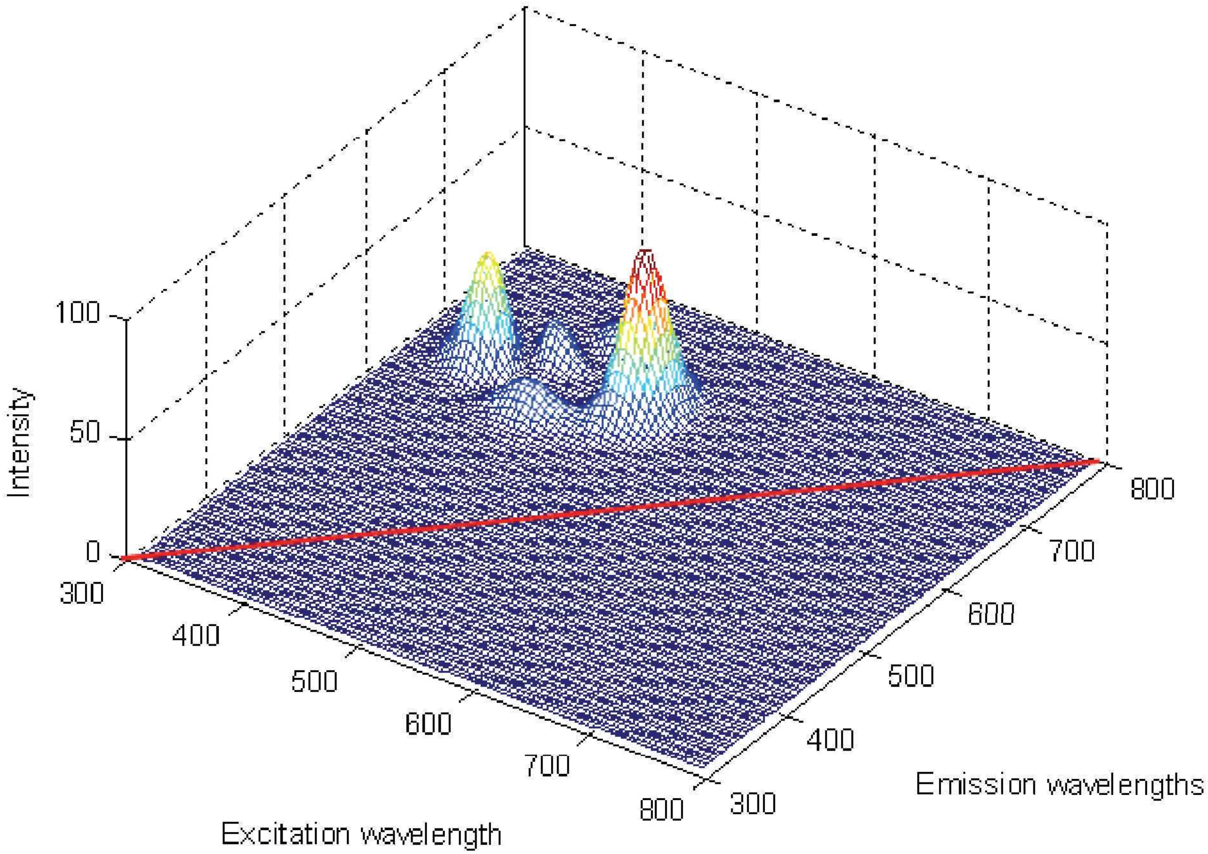
Beer-Lambert's law:

$dI_{(\lambda)}/dx = -\mu_{abs(\lambda)} \rightarrow I_{(\lambda)} = I_{0(\lambda)} e^{-\mu_{abs(\lambda)} x}$

Fluorescence

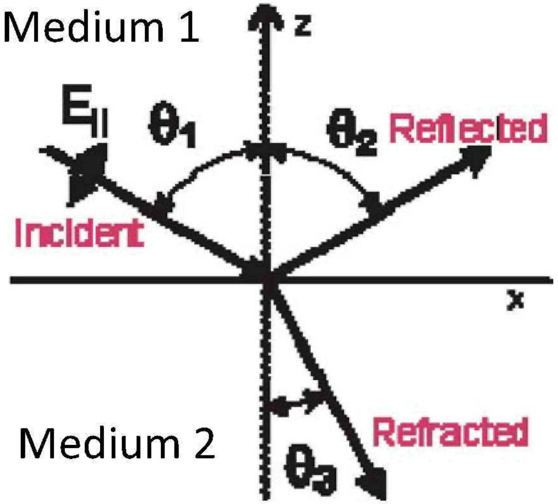


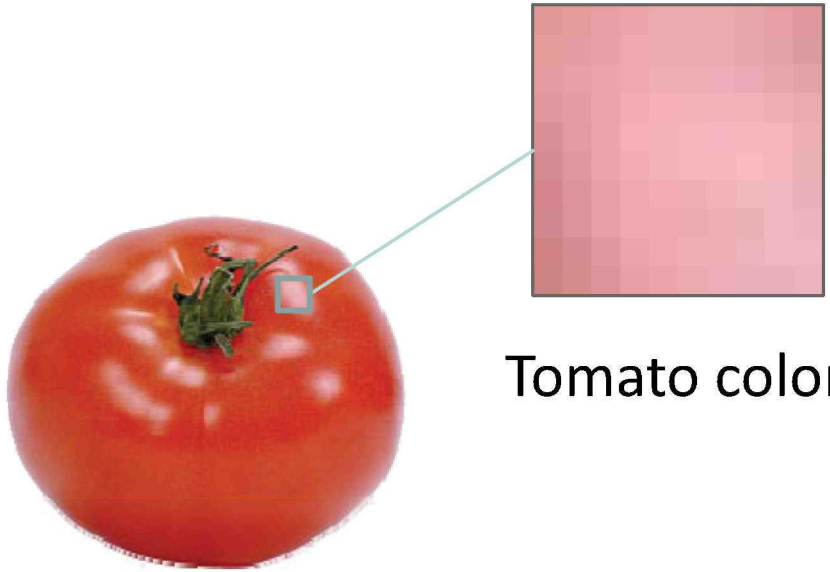
Fluorescence spectra



$$I = F(\lambda_{ex}, \lambda_{ex})$$

Reflecting surface





Tomato color?

Snell's law: $\theta_1 = \theta_2$

$$n_1 \sin\theta_1 = n_2 \sin\theta_3$$

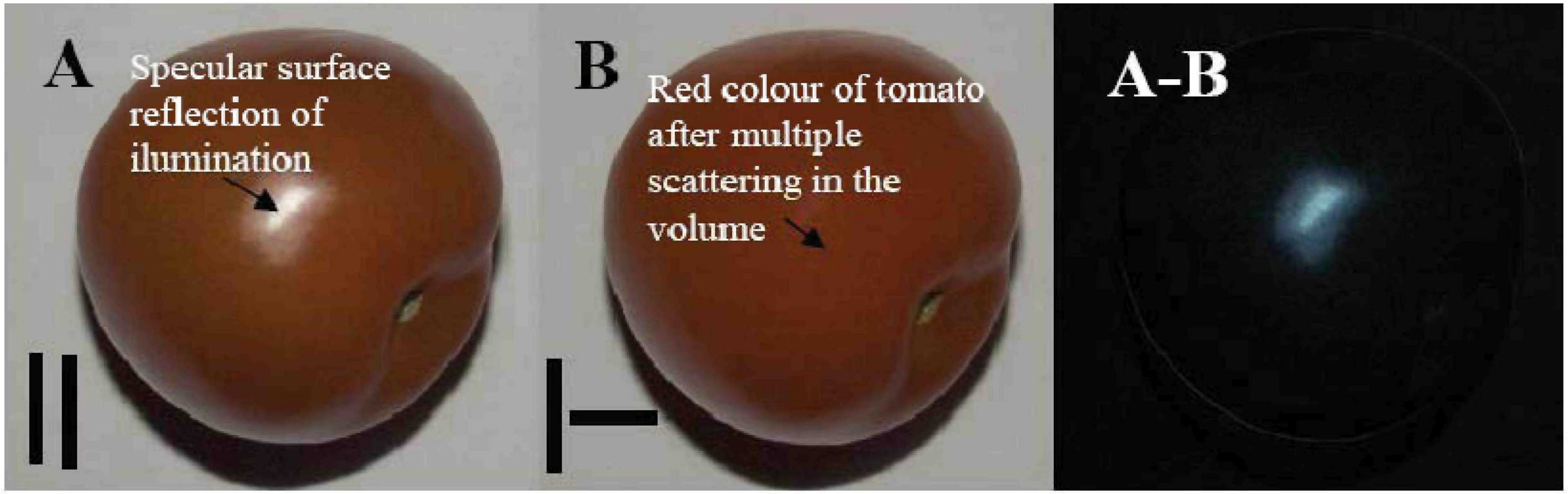
Fresnel equations:

$$R_{II} = \frac{tg^2(\theta_1 - \theta_3)}{tg^2(\theta_1 + \theta_3)}; \qquad T_{II} = \frac{\sin 2\theta_1 \sin 2\theta_3}{\sin^2(\theta_1 + \theta_3)\cos^2(\theta_1 - \theta_3)}$$

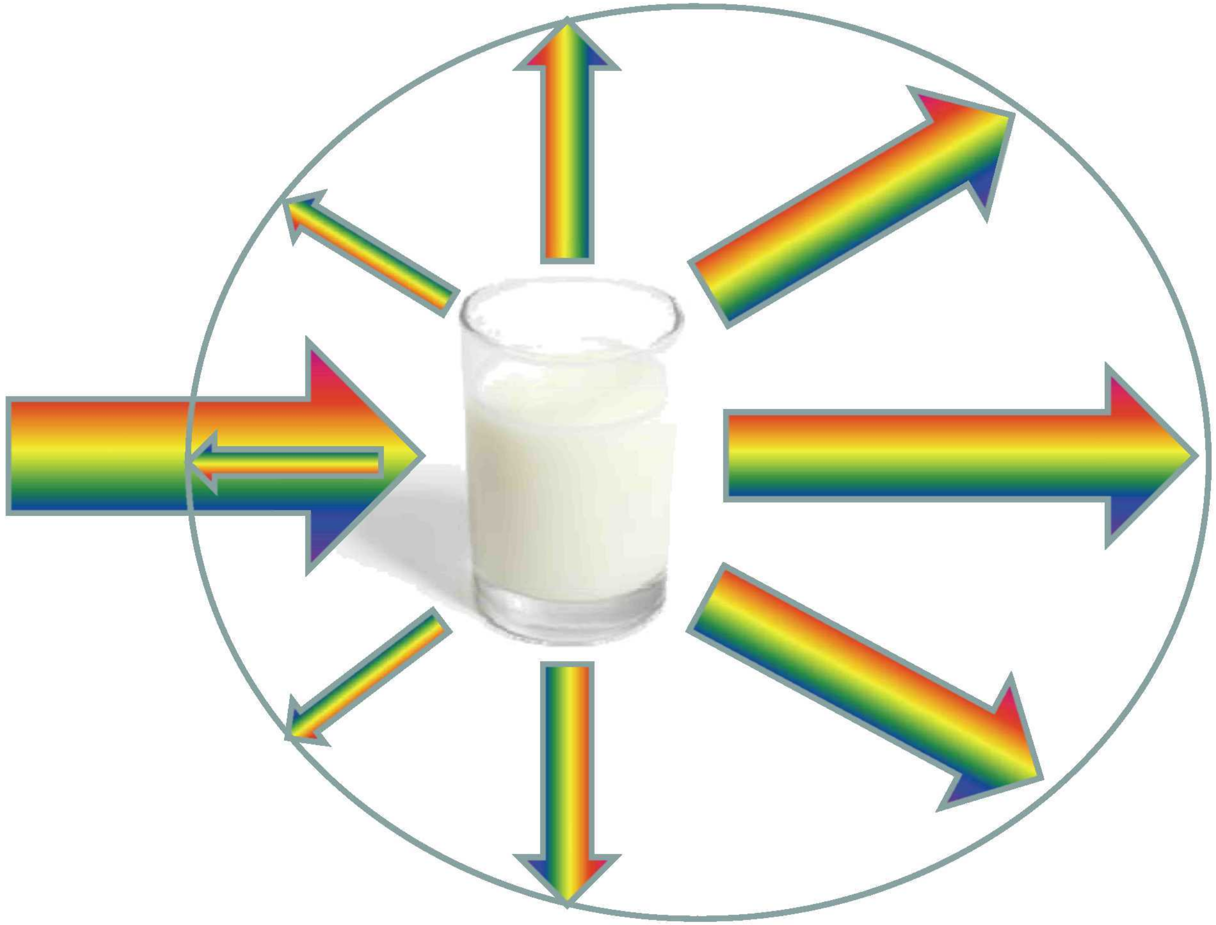
$$R_{\perp} = \frac{\sin^2(\theta_1 - \theta_3)}{\sin^2(\theta_1 + \theta_3)}; \qquad T_{\perp} = \frac{\sin 2\theta_1 \sin 2\theta_3}{\sin^2(\theta_1 + \theta_2)}$$

$$R_{II,\perp} = \left(\frac{n_2 - n_1}{n_2 + n_1}\right)^2; \qquad T_{II,\perp} = \frac{4n_2 n_1}{(n_2 + n_1)^2}$$

Specular


A
Specular surface
reflection of
ilumination
B
Red colour of tomato
after multiple
scattering in the
volume
A-B

Scattering volume



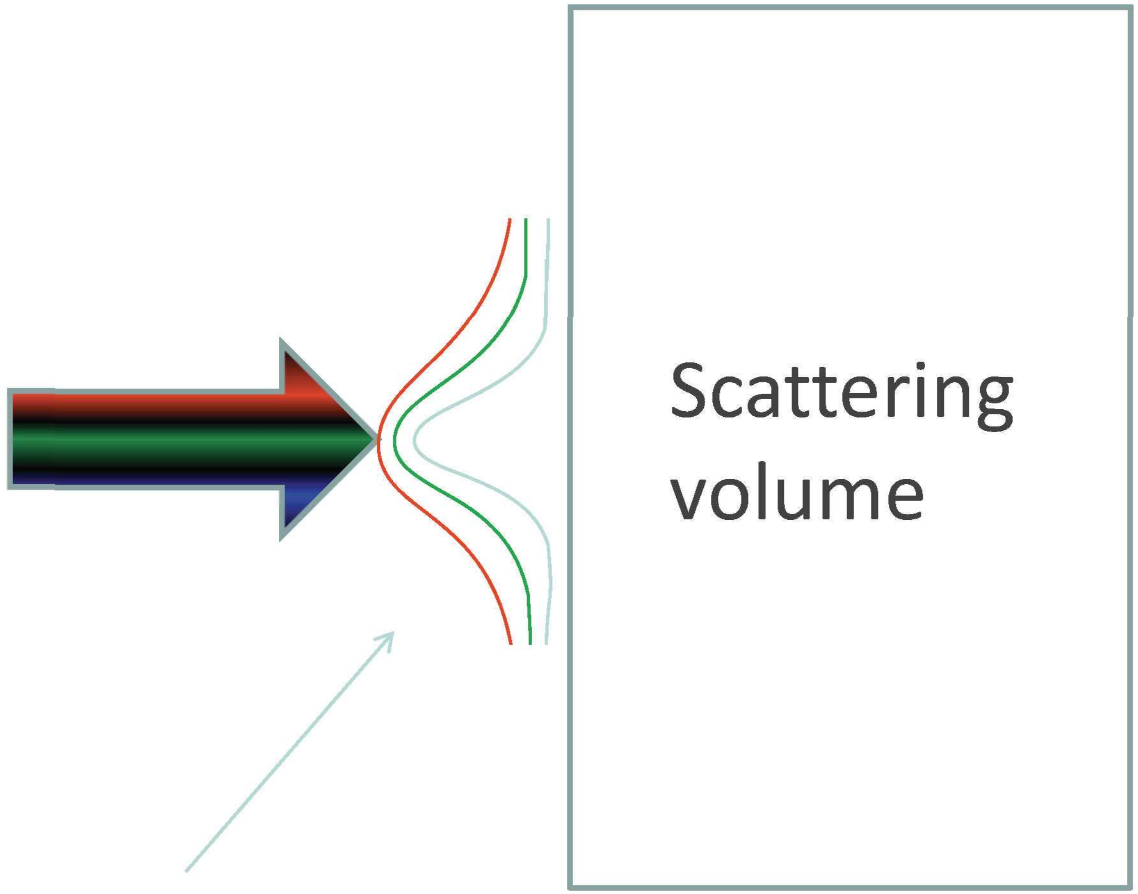
Probability of scattering:

$\mu_{sca(\lambda)}$

Dependence of incidence angle:

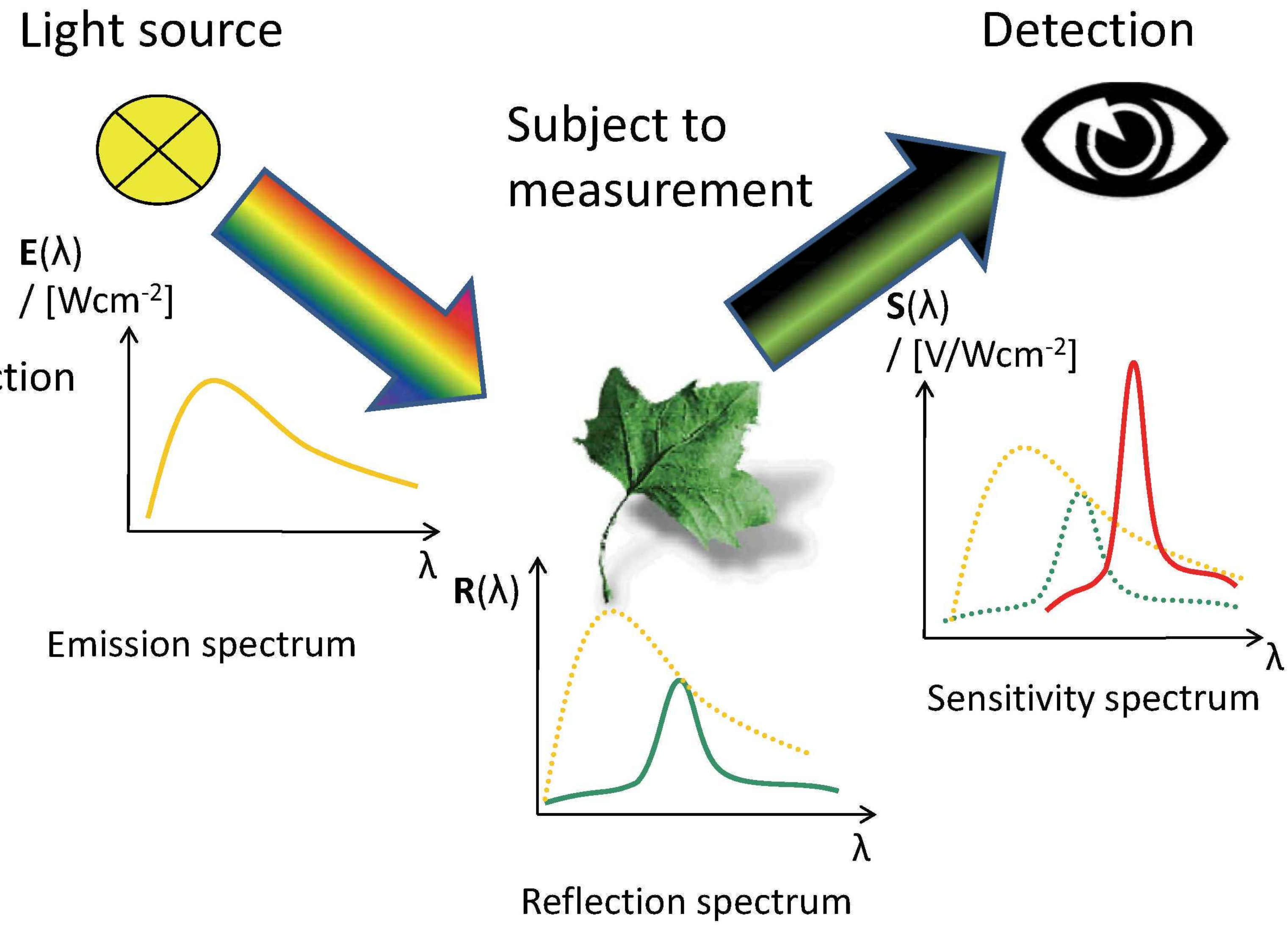
g-factor

Back-scattered light



Green's functions:
$G(\mu_{abs(\lambda)}, \mu_{sca(\lambda)}, g)$

Simplified spectroscopy



- Emission, reflection and sensitivity

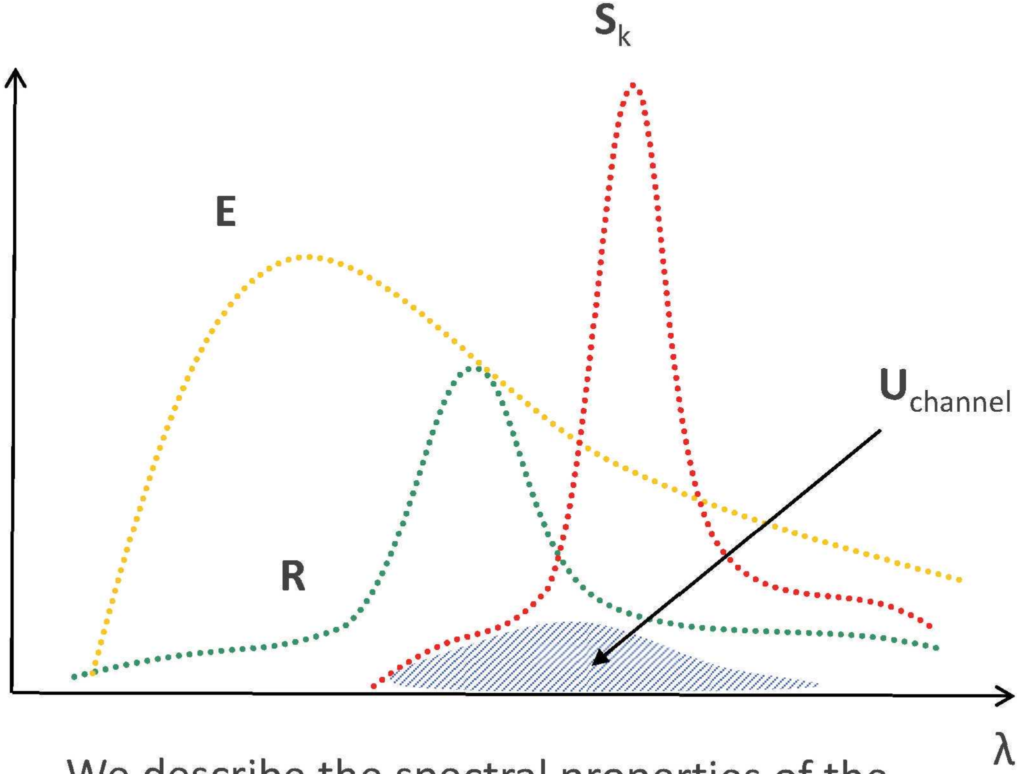
- Reflectance $R(\lambda)=I(\lambda)/I_0(\lambda)$

Contribution to a spectral band

$$U_{channel(k)} = \int_0^\infty E_{(\lambda)} R_{(\lambda)} S_{k(\lambda)} d\lambda$$

- Emission, reflection and sensitivity

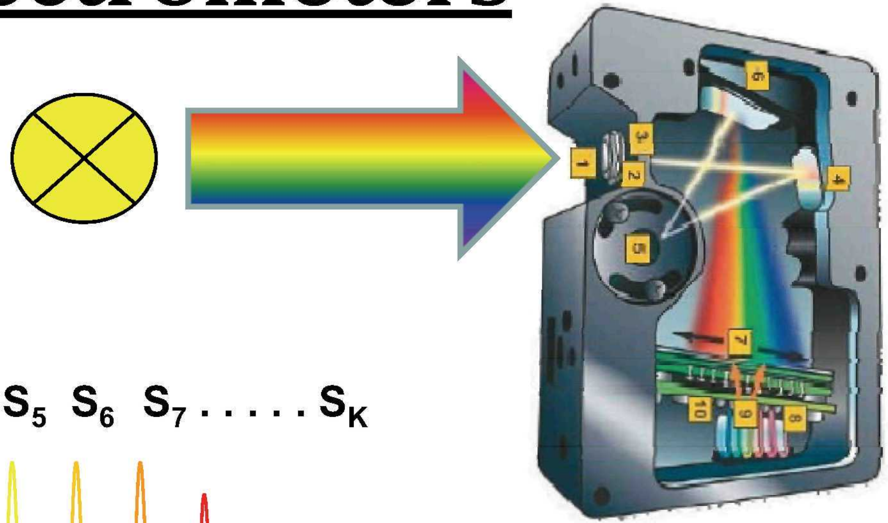
- Reflectance $R(\lambda)=I(\lambda)/I_0(\lambda)$

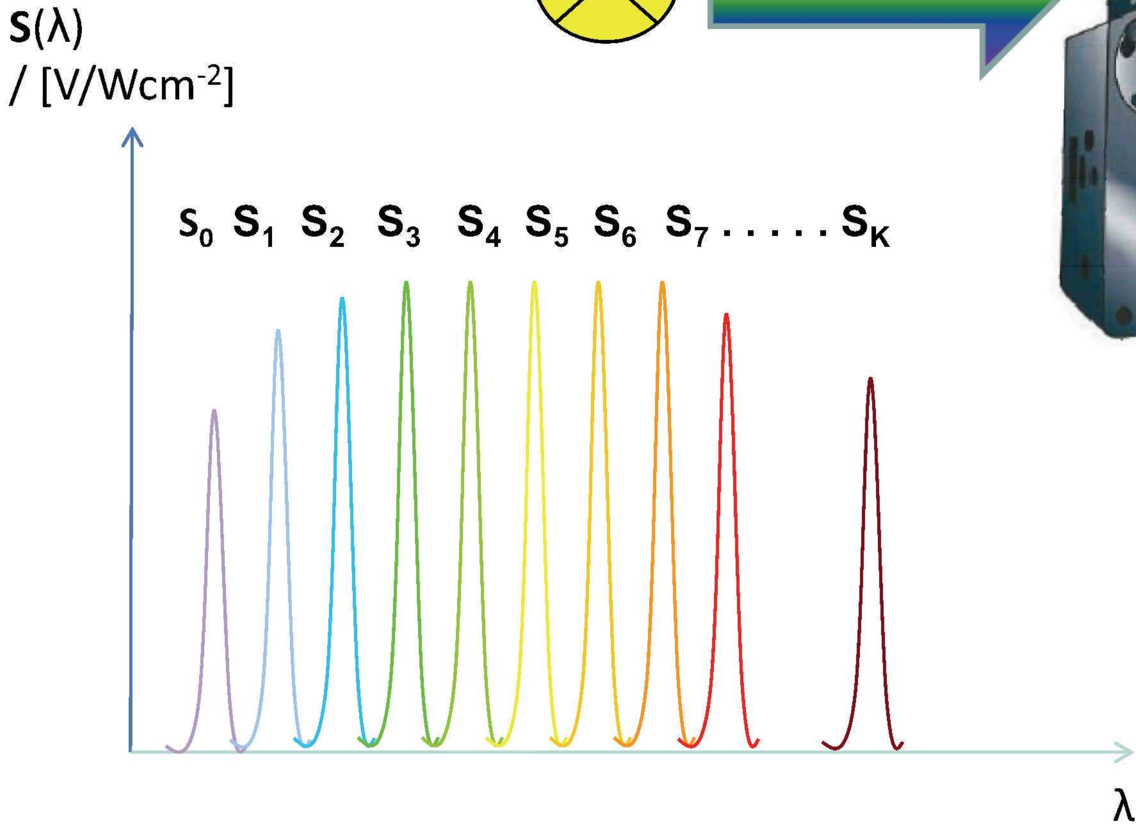


We describe the spectral properties of the object in terms of $\mathbf{U_{channel}}$. More channals can provide better describtion of **R**

Classical spectrometers

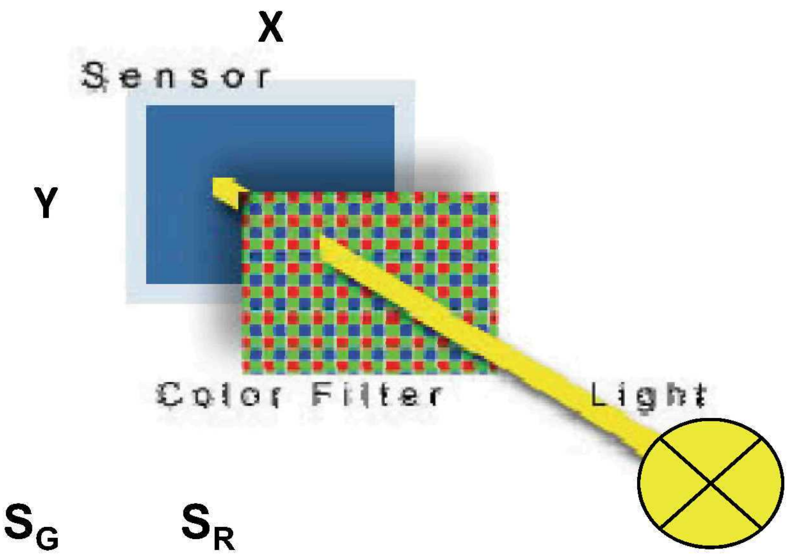
Sources of spectral data

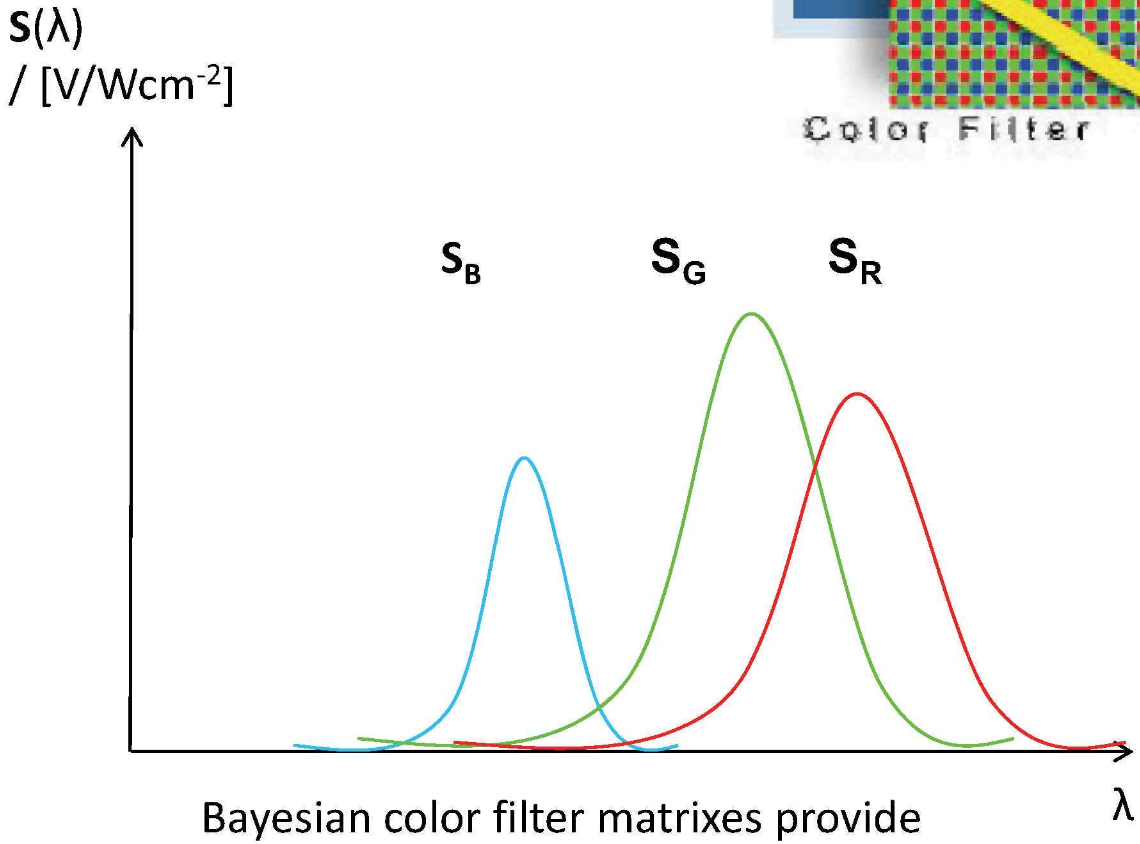




Filtering and/or dispersion of light provide thousands of spectral bands. We obtain a **1 times K** vector containing a spectrum from each measurement.

Color cameras

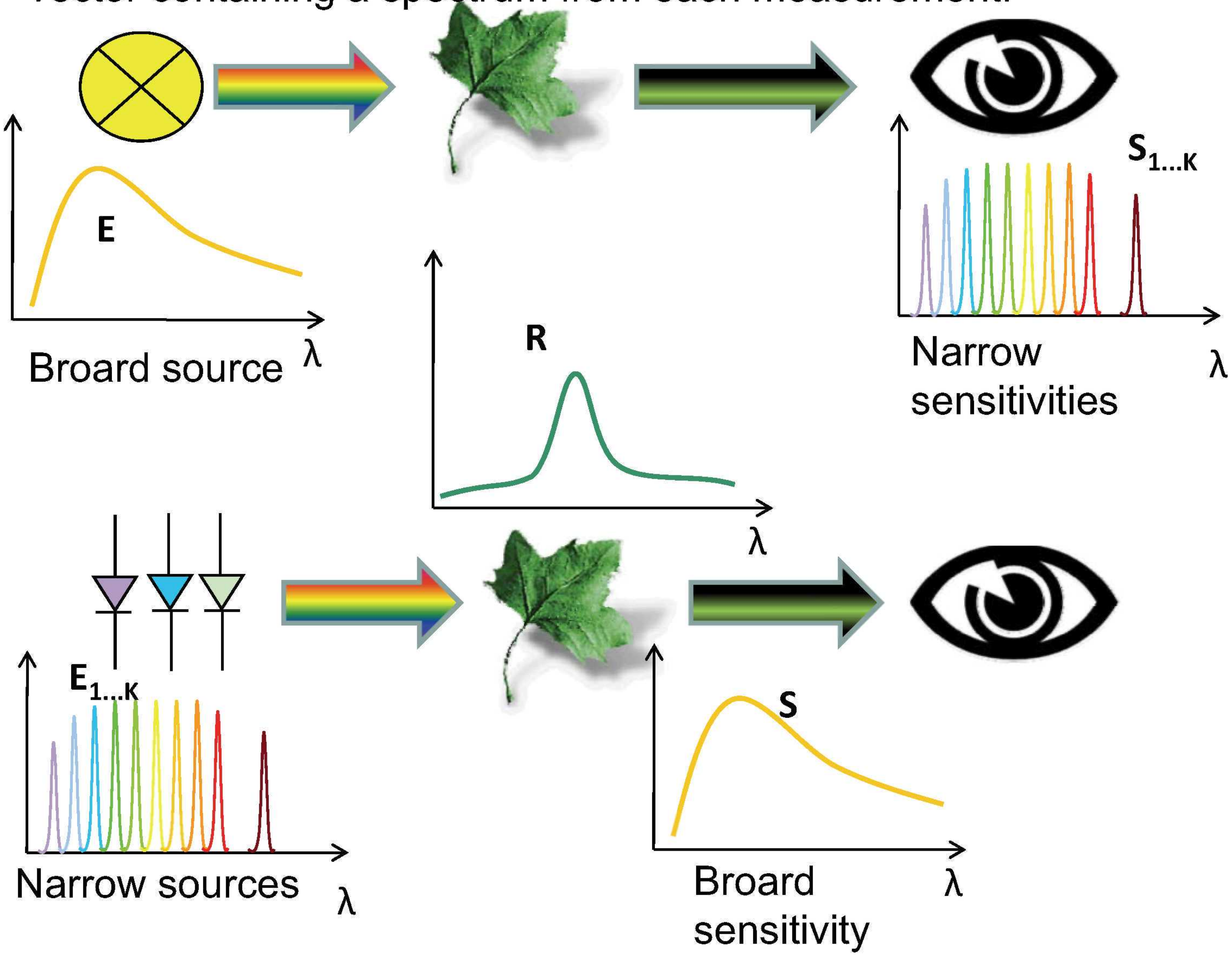




Bayesian color filter matrixes provide spectral bands in color pictures. We obtain **XY times 3** individual spectral measurements. Objects in the pictures is described spectroscopically in the **RGB** color space.

Assessment of spectral properties R

Spectral properties **R** might as well be assessed with various **E** rather than various **S.** We obtain a **1 times K** vector containing a spectrum from each measurement.

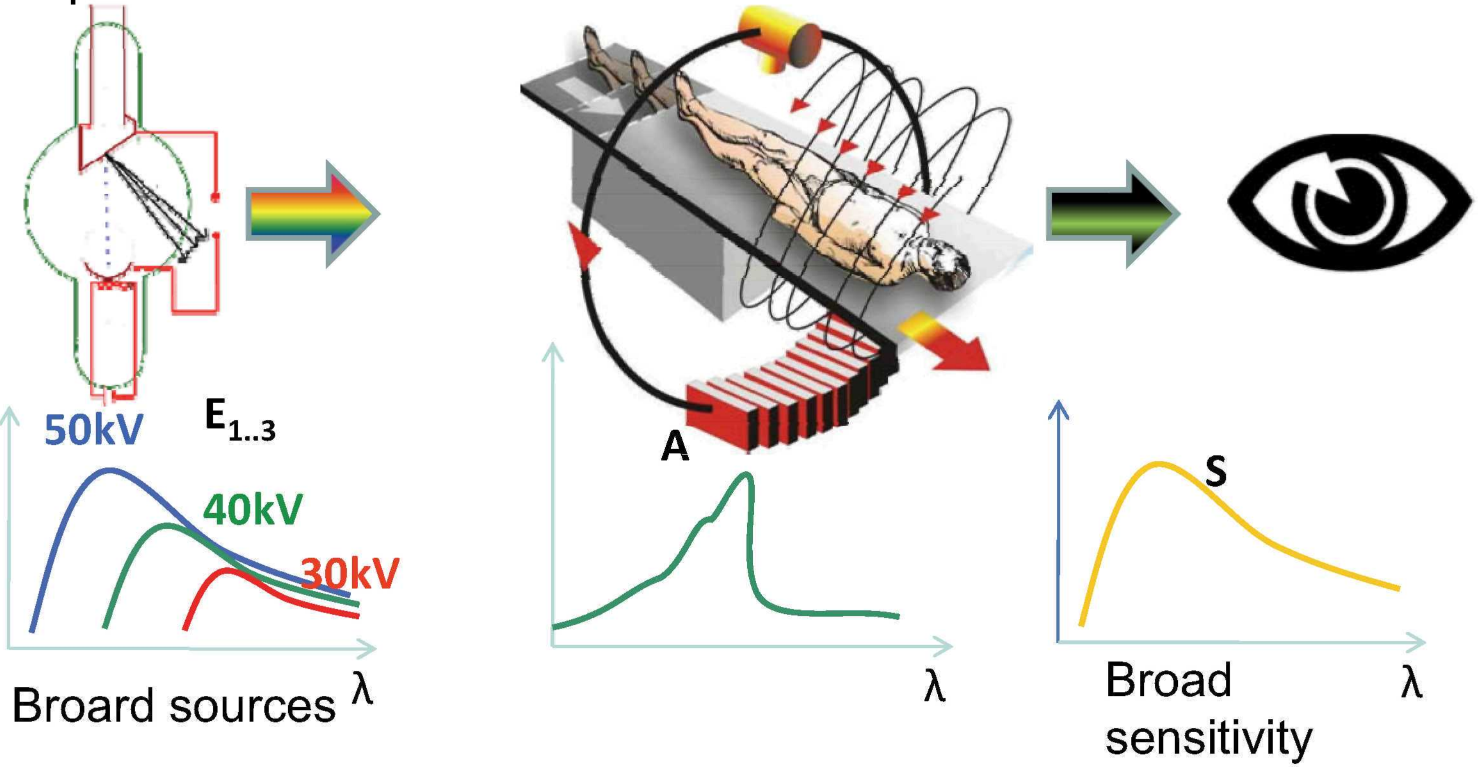


Sources of spectral data

Sources of spectral data

Example: Multispectral CT

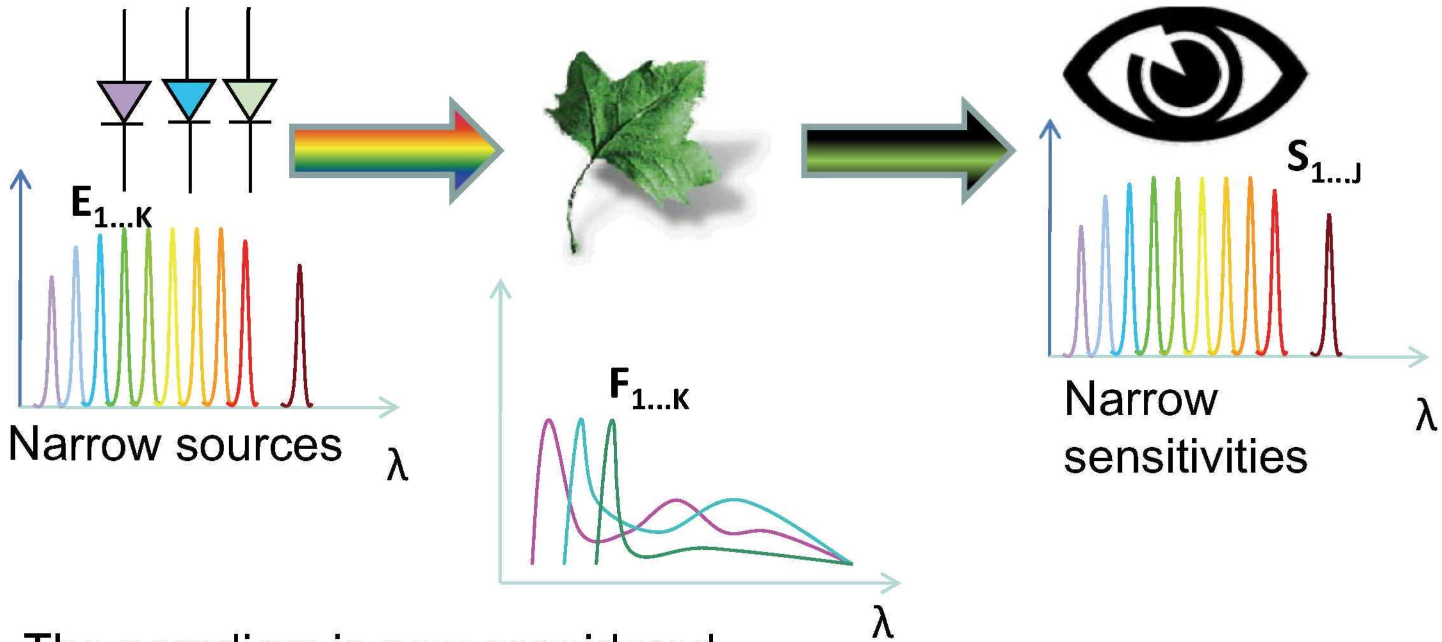
Different organs are distinguished by spectral properties accessed by changing X-ray tube voltage to produce different **E**. We obtain a **1 times 3** vector containing a spectrum from each measurement



Sources of spectral data

Example: Fluorescence

The 2D EEM spectral properties **F** can be assessed with various **E** and various **S.** We obtain a **K times J** matrix containing a excitation emission surface from each measurement. Data are **rearranged** into a **1 times KJ** vector and one dimension is discarded.



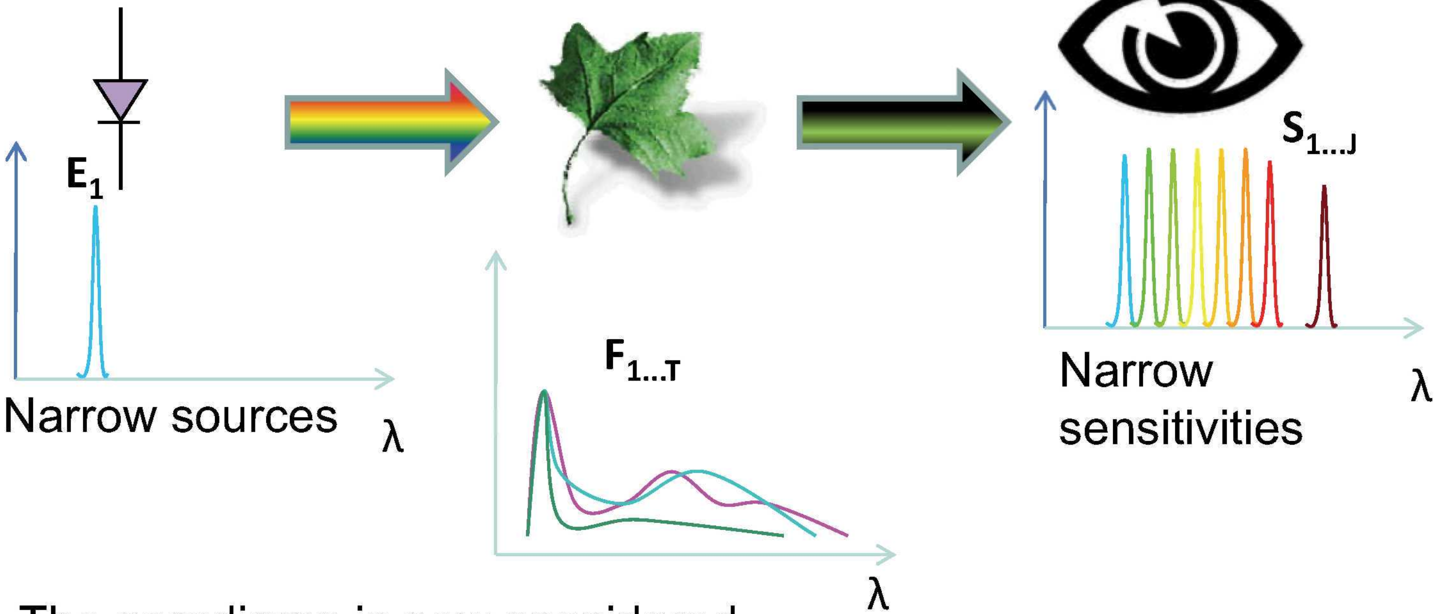
The paradigm is now considered

$$U_{channel(k,j)} = \int_0^\infty F_{(E_{k(\lambda)},\lambda)} S_{j(\lambda)} d\lambda$$

Sources of spectral data

Example: Fluorescence

Fluorescence properties **F** can be assessed at various bleaching times **t,** or at different sample temperature**T** to study photokinetics. Resulting light is studied at various **S.** We obtain a **K times J** matrix containing a excitation emission surface from each measurement. Data are **rearranged** into a **1 times KJ** vector and one dimension is discarded.



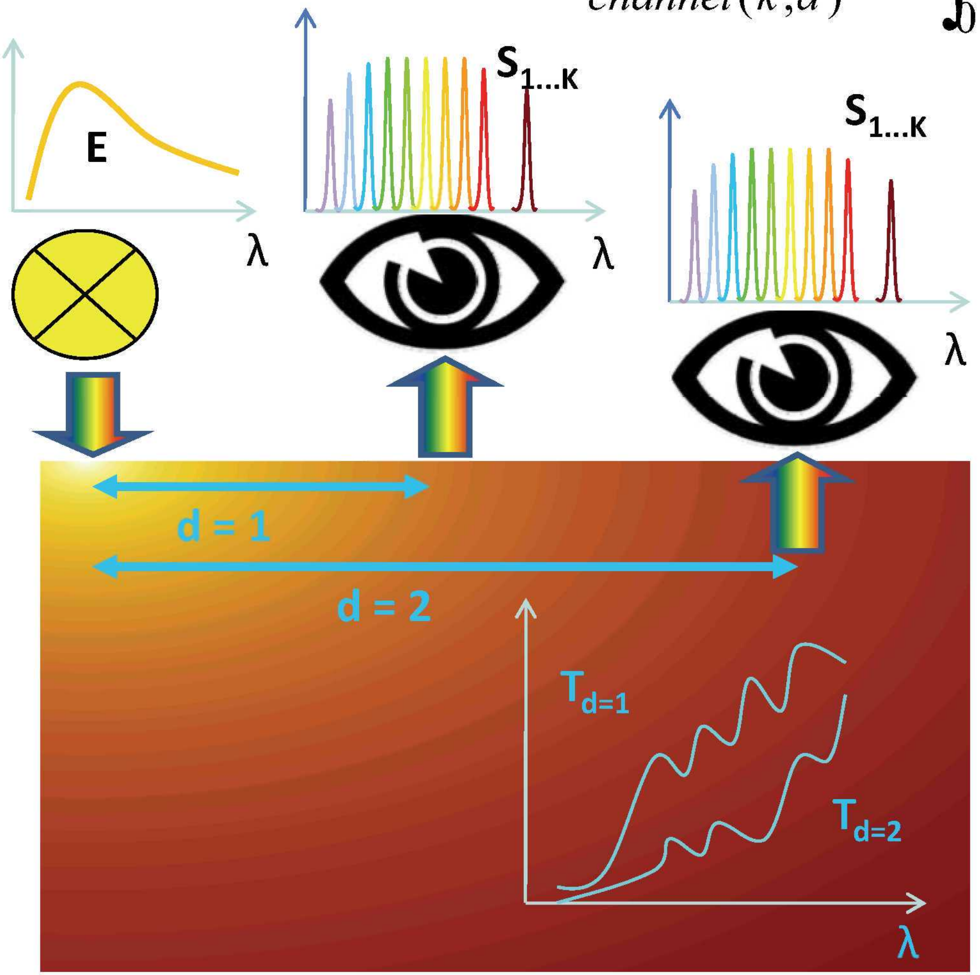
The paradigma is now considered

$$U_{channel(k,j)} = \int_0^{\infty} F_{(t,\lambda)} S_{j(\lambda)} d\lambda$$

Example: Scattering properties

Scattering properties can be accessed by tranmission spectras **T** from several path lengths **d**. A **K times 2** matrix is obtained and rearranged into a **1x2K** vector from each measurements.

$$U_{channel(k,d)} = \int_0^\infty E_{(\lambda)} T_{(d,\lambda)} S_{k(\lambda)} d\lambda$$

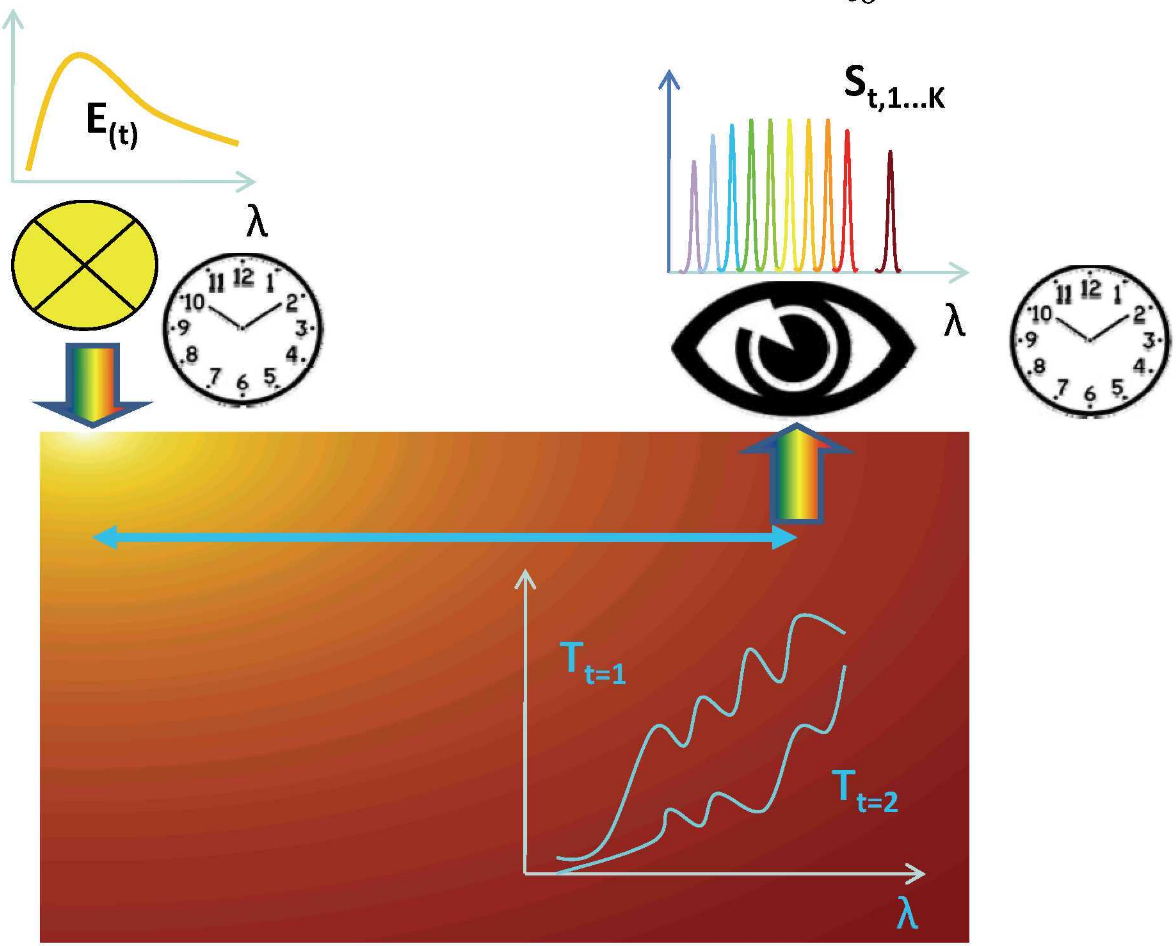


Sources of spectral data

Example: Time resolution

Scattering properties can be accessed by tranmission spectras **T** from several travel times **t**. A **K times T** matrix is obtained and rearranged into a **1xTK** vector from each measurement.

$$U_{channel} = \int_0^\infty E_{(\lambda,t_0)} T_{(\lambda,\Delta t)} S_{(\lambda,t_T)} d\lambda$$

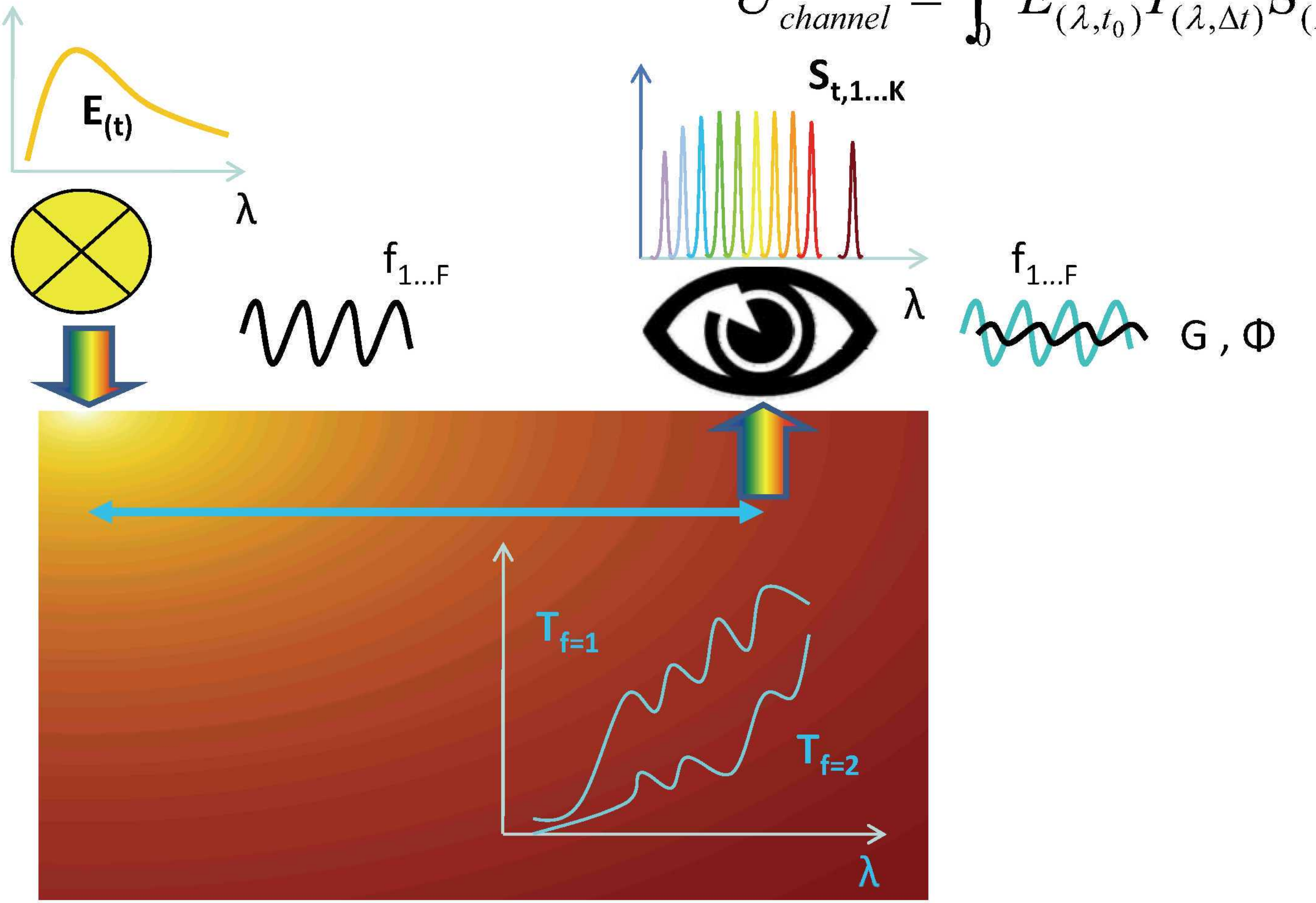


Sources of spectral data

Example: Time resolution

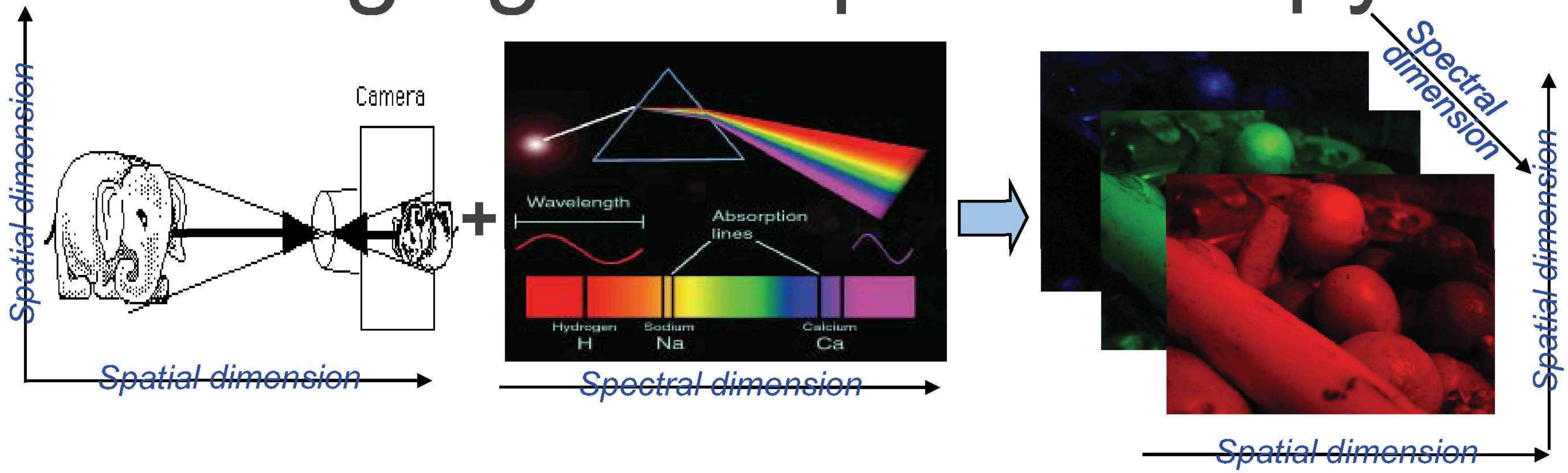
Scattering and fluorescence properties causing delays can be studied equally in time or frequency domain. Spectras **T** from several modulation frequencies **f are obtained**. A **K times F** matrix is obtained and rearranged into a **1xFK** vector from each measurement.

$$U_{channel} = \int_0^\infty E_{(\lambda,t_0)} T_{(\lambda,\Delta t)} S_{(\lambda,t_T)} d\lambda$$



Sources of spectral data

Imaging and spectroscopy



- **Imaging**. Obtaining a BW picture, a 2D matrix with spatial information of the sum of all wavelength
- **Spectroscopy**. Obtaining a 1D vector with all wavelengths in one point.
- **Multi spectral imaging**. Obtaining a 3D matrix with spectral information in each spatial point.

Example 1: A multi spectral CT scan would have three spatial dimensions and on spectral, equals 4D

Example 2: A real-time CT scan, will have above four dimension plus a temporal, equals 5D

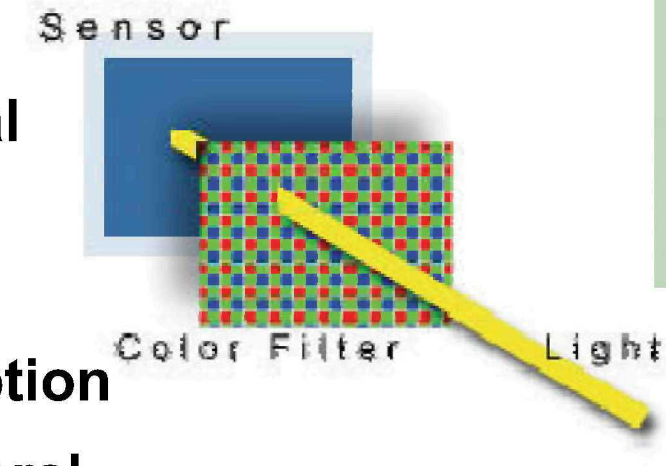
Trade off's

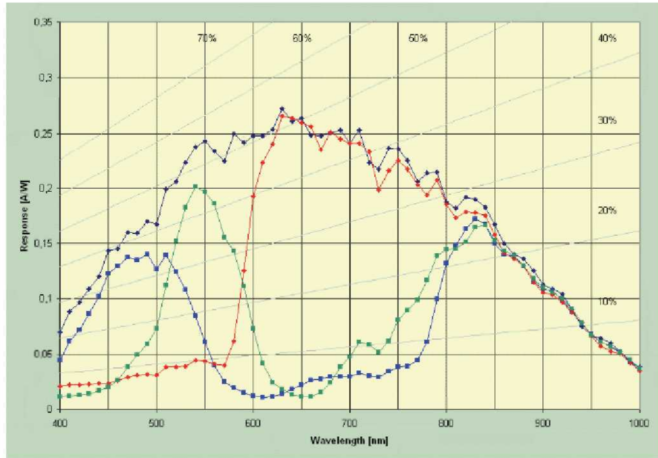
- In all following acquisition methods we will see a trade off between spatial, spectral and temporal resolution. We can not have it all with the same technique.

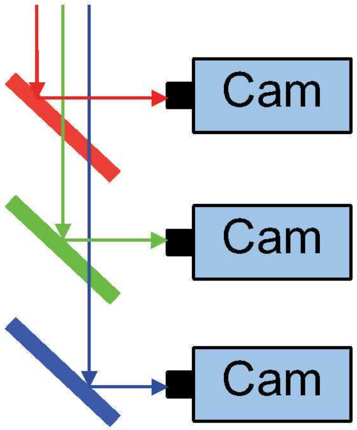
- During the lecture try to keep track on the dimensions and the quantization

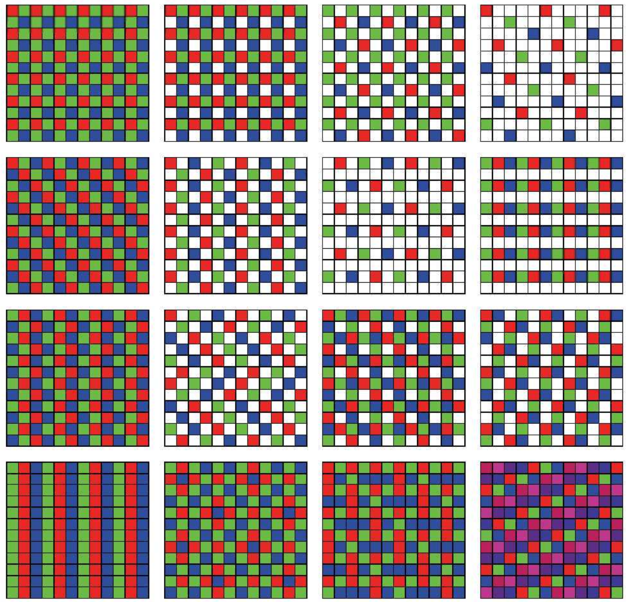
Color filters

- Digital color cameras
- Loss of spatial resolution
- Loss of light during absorption
- Loss of temporal resolution due to sensitivity
- Only few color channels



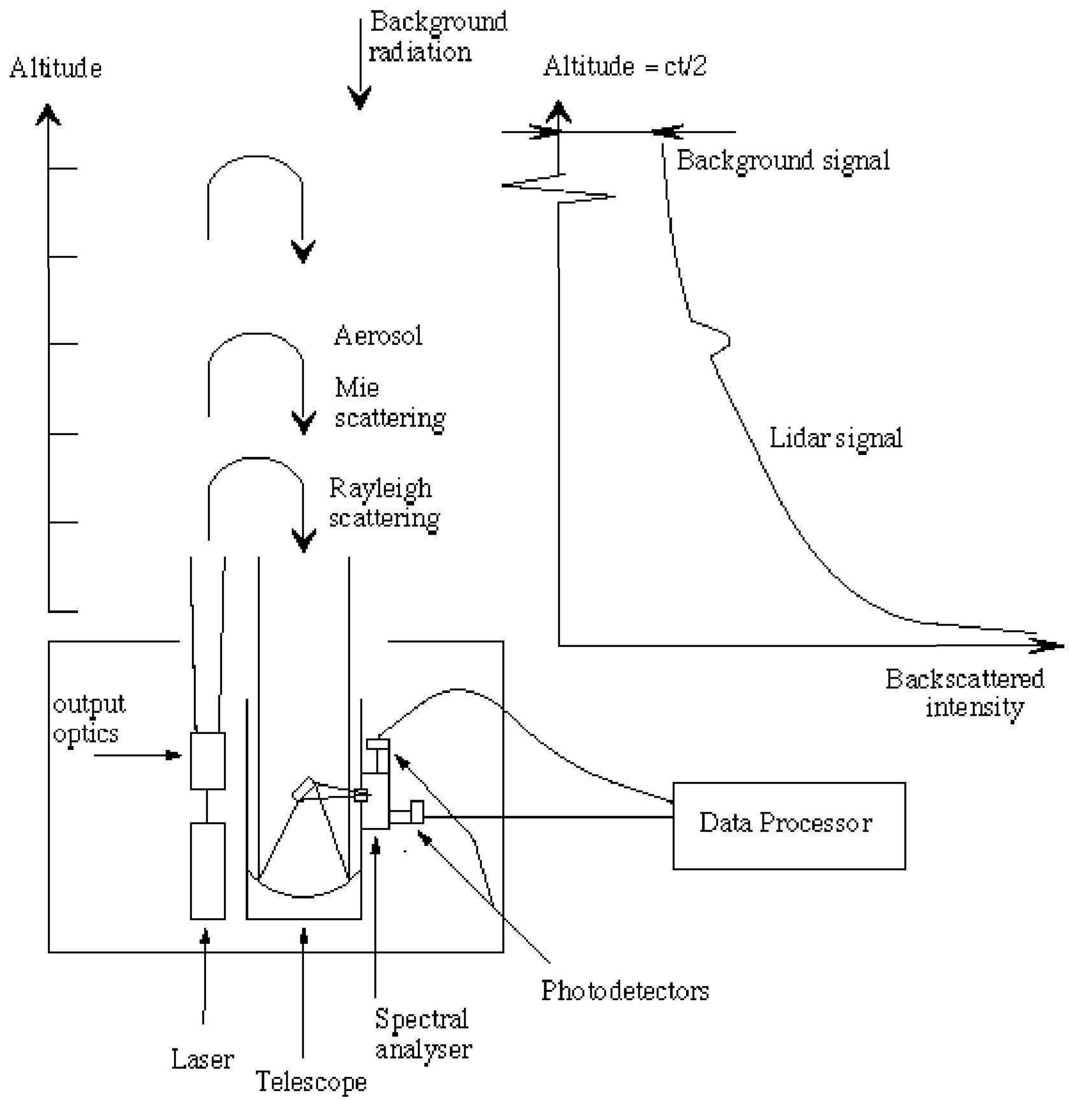






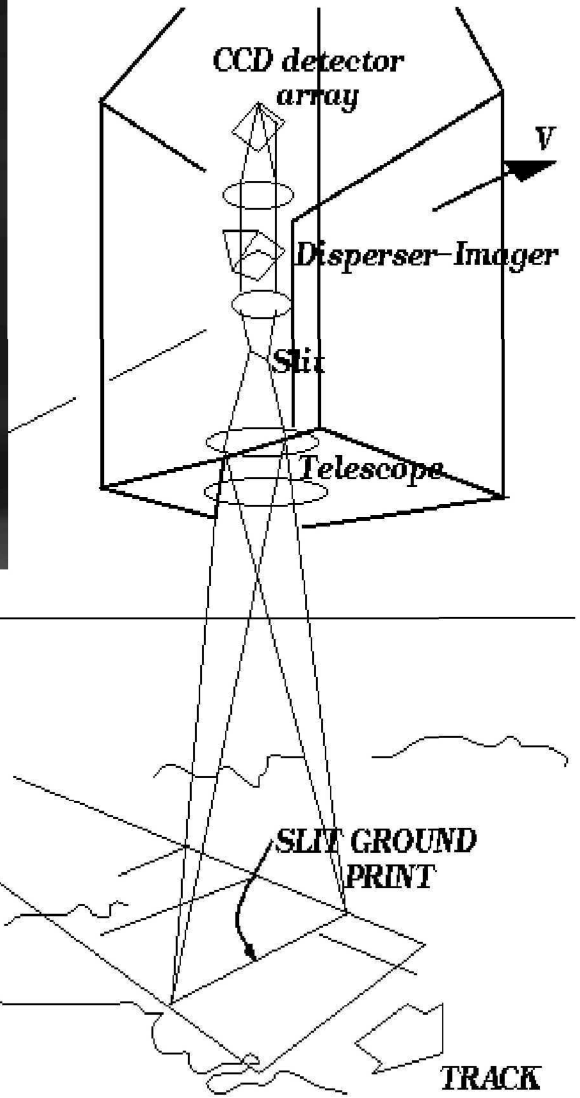
Scanning dot method

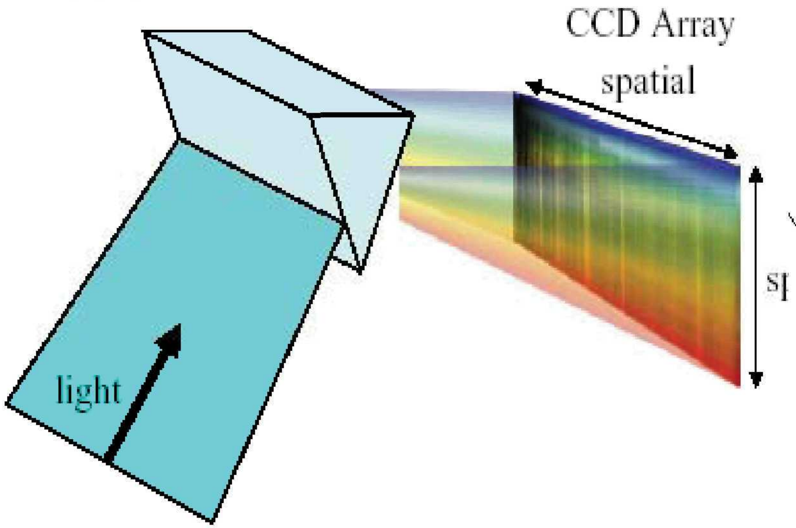
- Spectroscopy in one point at a time
- LIDAR
- Time axis and the travel of light
- Laser Induced Plasma Spectroscopy



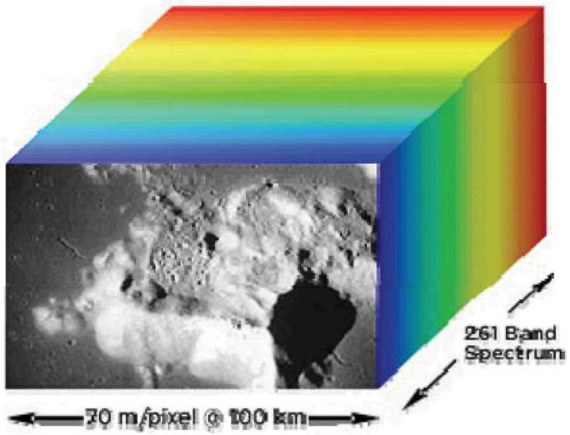
Push broom method

- CCD recieves one spatial dimension and one spectral
- Temporal dimension is used to reconstruct a 2D image
-Temporal resolution is poor
-Monochromatic stabilization

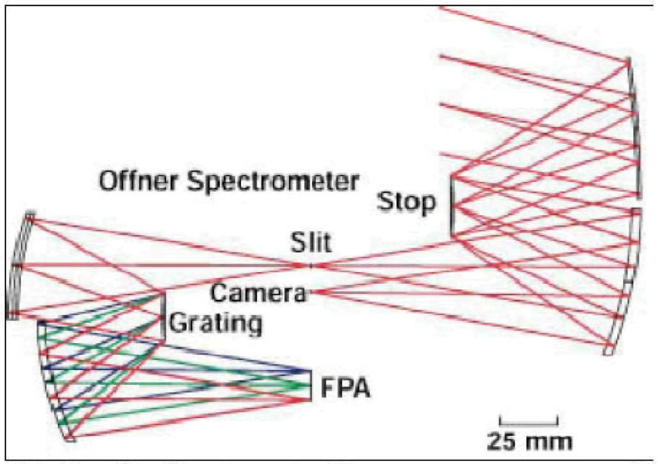




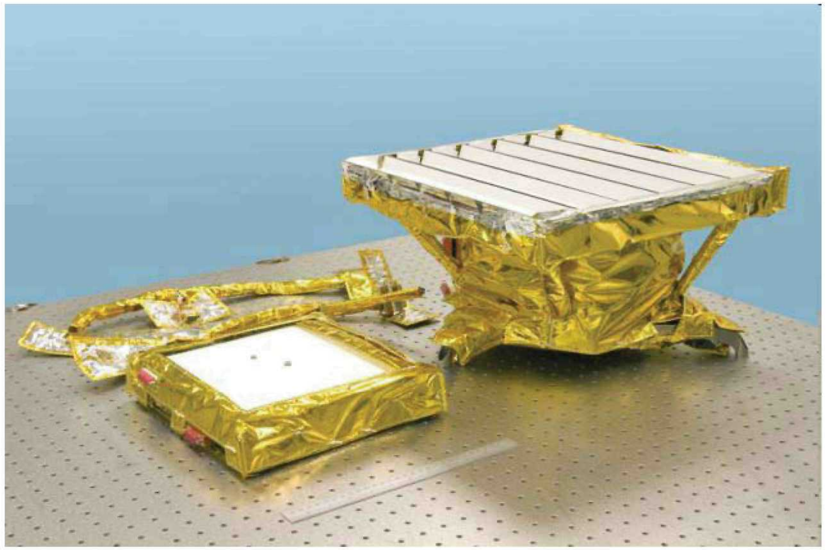
Push broom



NASA's M^3 imaging spectrometer
430-3000nm



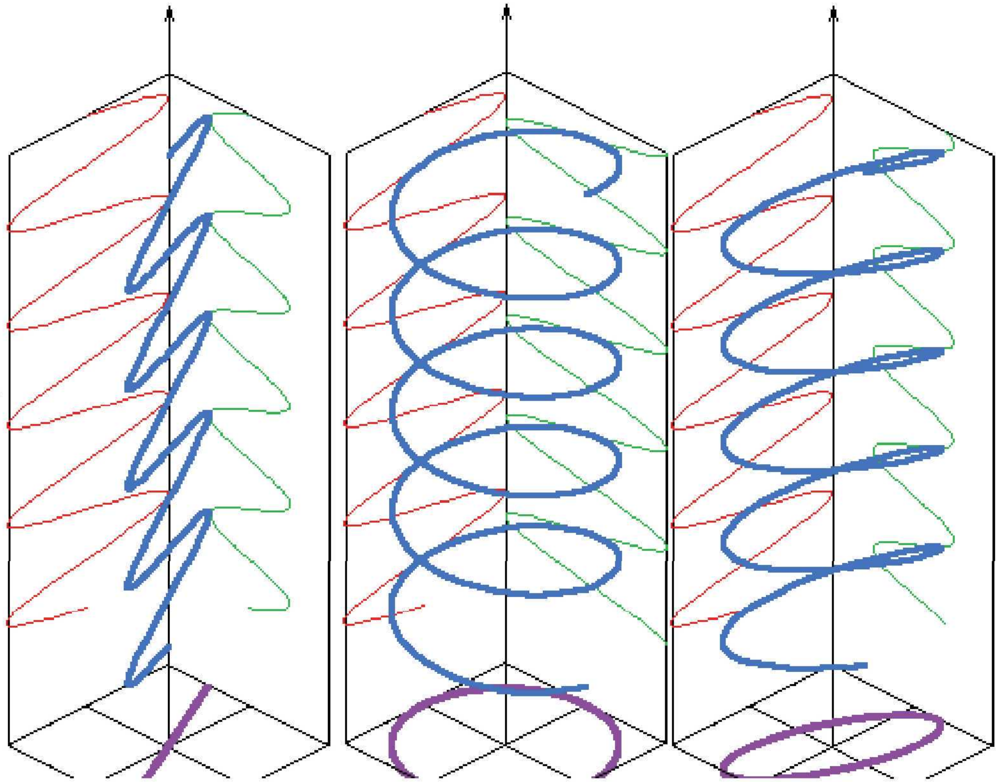
This **Imaging Spectrometer/Camera** uses a convex grating in a lateral Offner configuration to reduce distortion to less than 1 percent across a 12-mm spectrum with a 16-mm image.

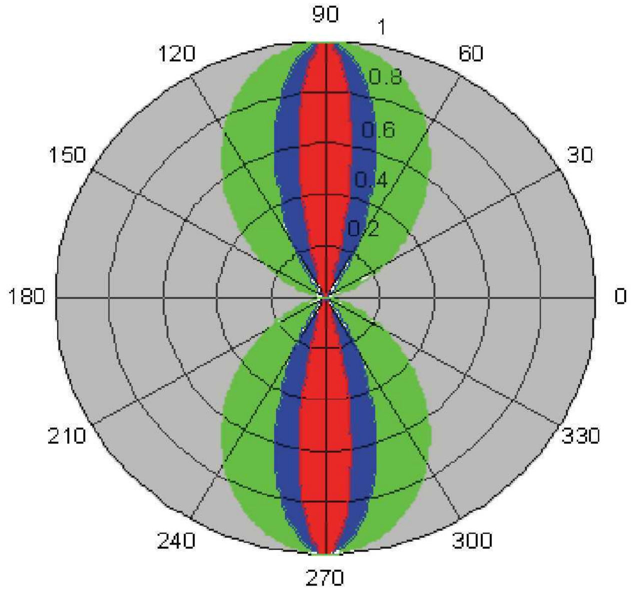


http://m3.jpl.nasa.gov/

Polarized light

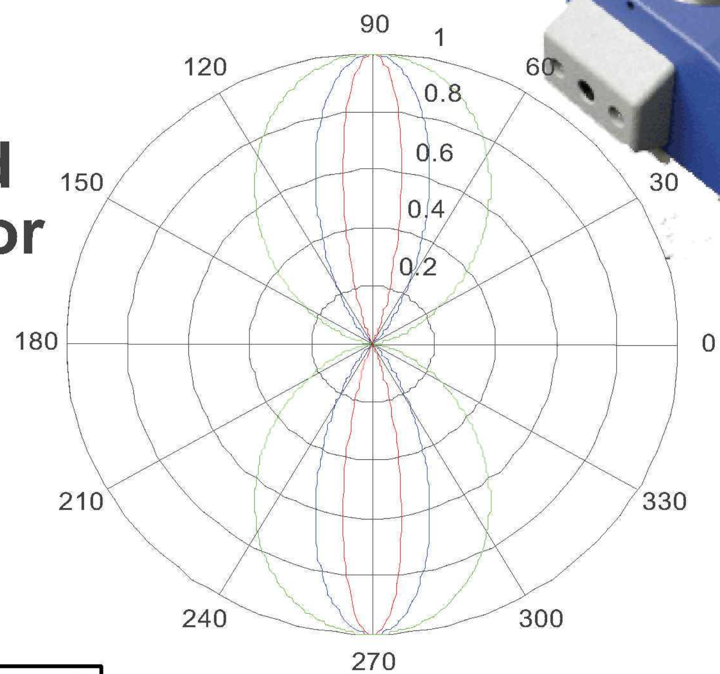
- The spatial orientation of the E-vector in respect to the propagation vector
- Polarized sources
- Mixed polarizations
- Reflections and polarization
- Polarizing filters

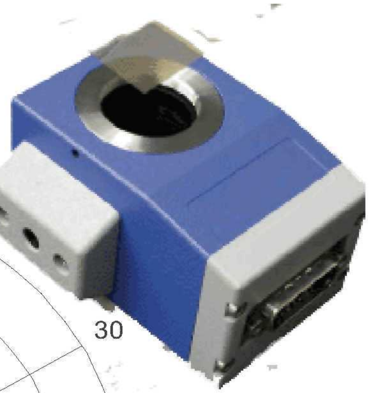


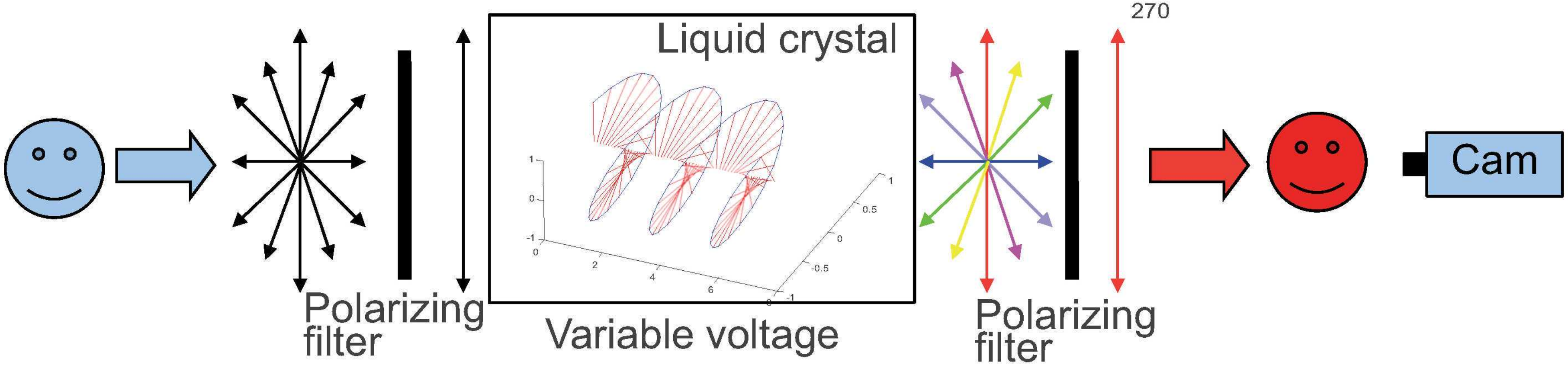


Tunable wavelength filters

- **Light is lost in the filters twice**
- **Narrow filters provide good spectral resolution, but poor light conditions and long exposure time**
- **Optimization for relevant acquisitions can be made**

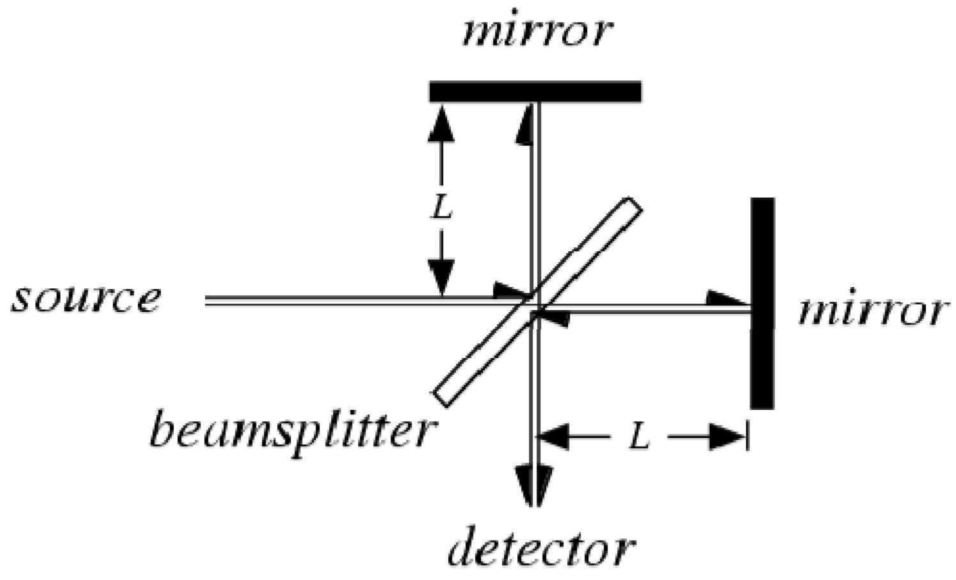


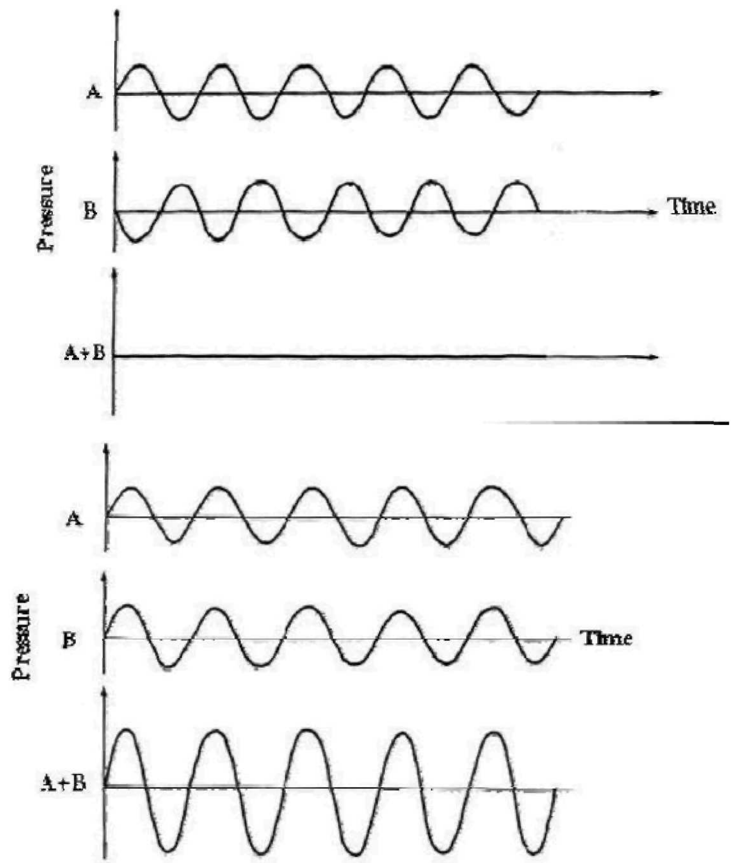




Interferometry

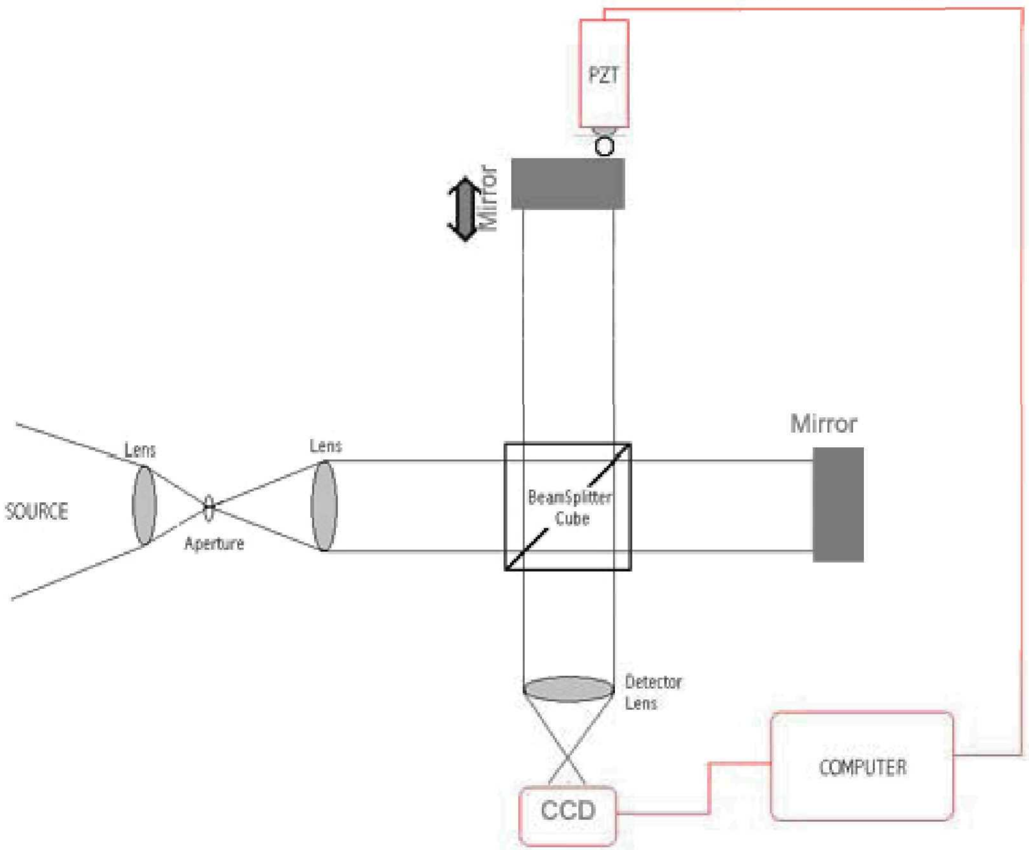
- Michelson Interferometer
- Particle wave dualism
- Constructive / destructive interference

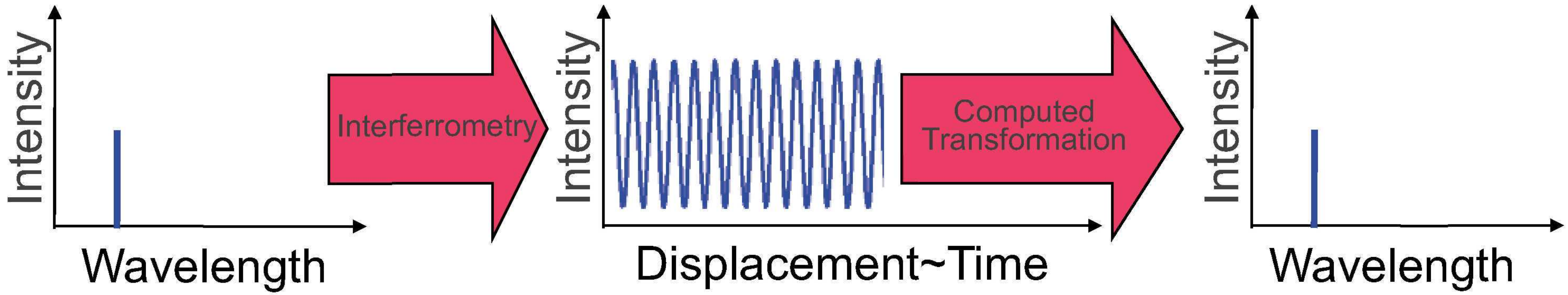




Interferometric Fourier imaging

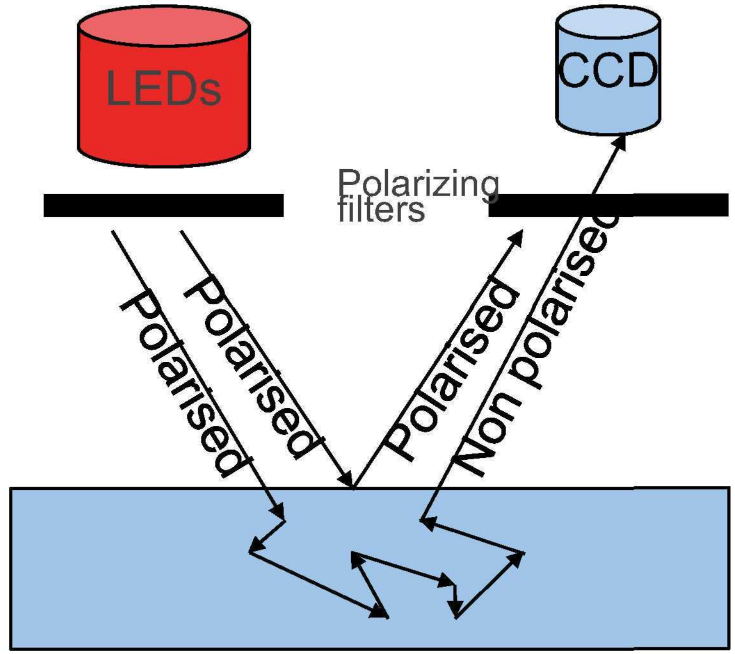
- **No light is wasted**
- **Temporal dimension is converted to spectral dimension**





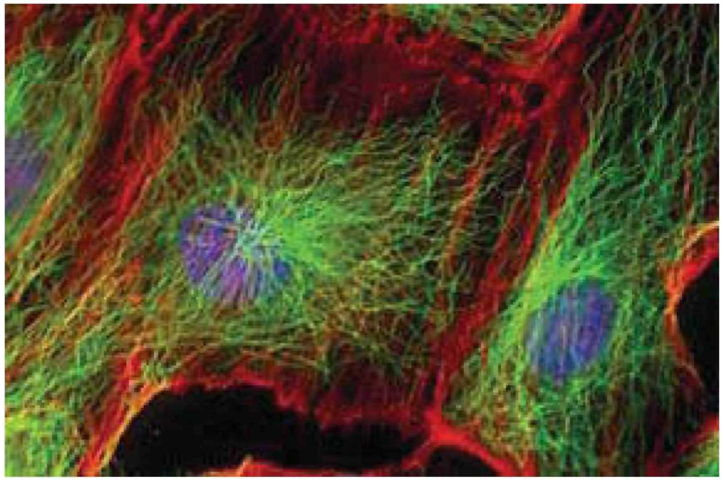
Absorption imaging

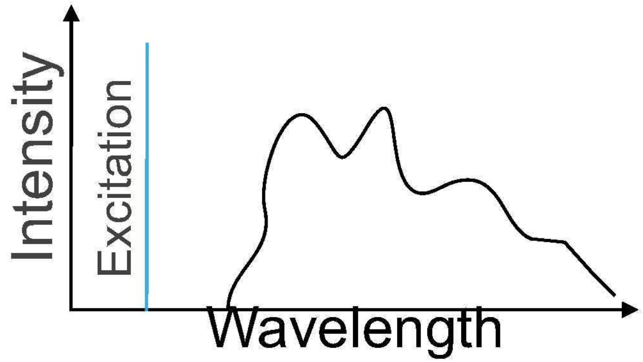
- Monochromatic light sources
- Avoiding direct reflection
- Inexpensive technique
- Elimination of distance and shadows

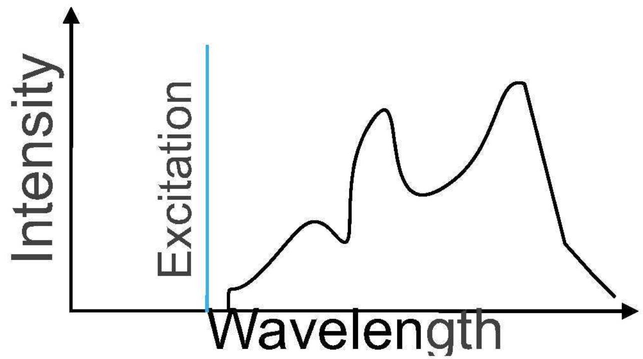


Fluorescence imaging

- Additional spectral profiles
- Sensitizers
- Designed emission spectres
- Life time imaging

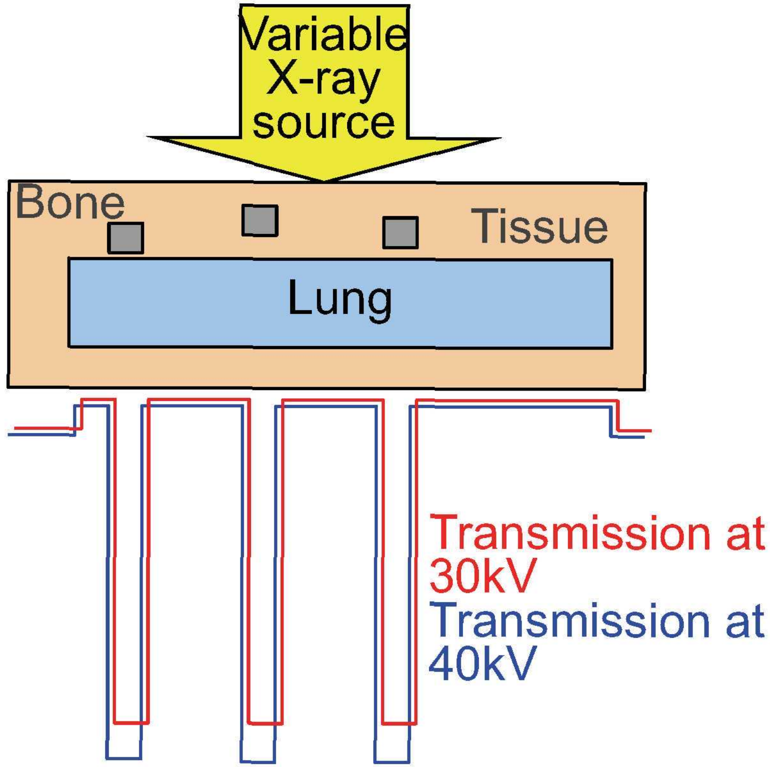






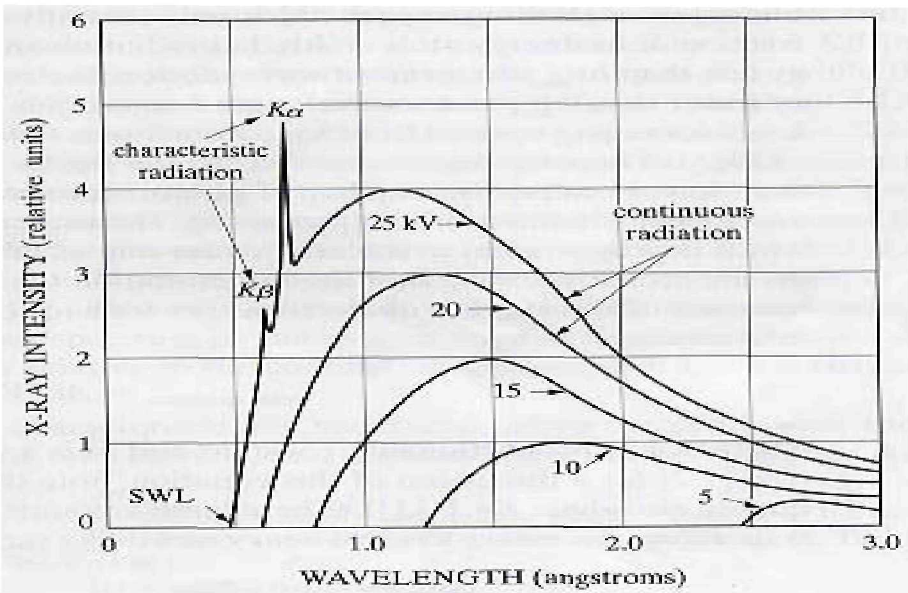
Multi spectral X-ray

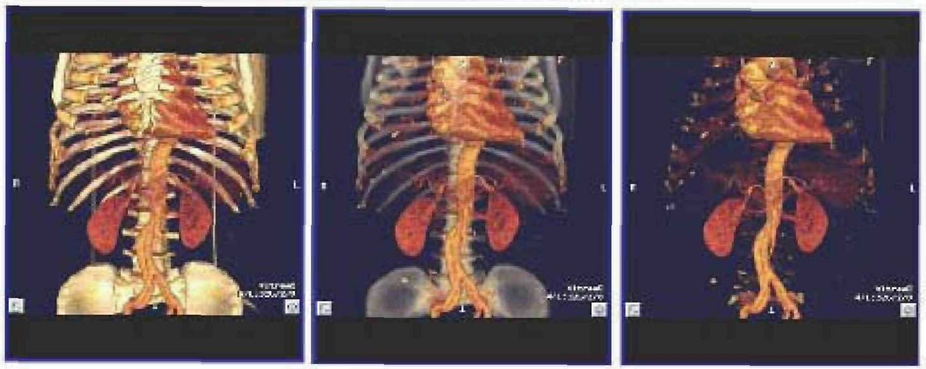
- **By changing the tube voltage, the ”colour” of the X-rays is changed.**
- **Different tissues, bone, liquids and gasses have different absorptions spectres**

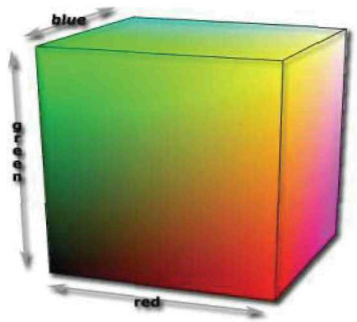


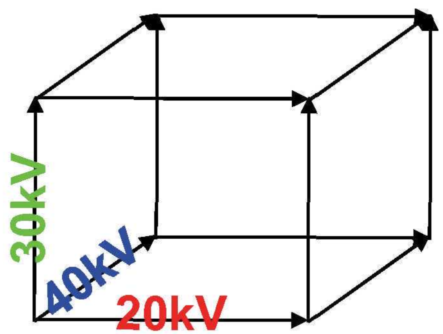
$$T_{30} - \beta * T_{40} = T_{contrast}$$

Contrast function for lungs



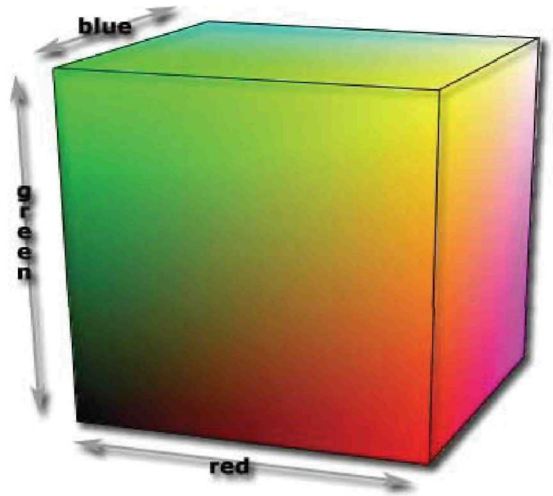


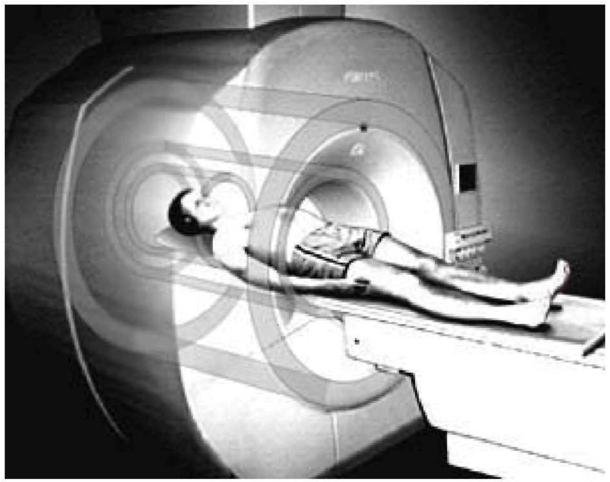


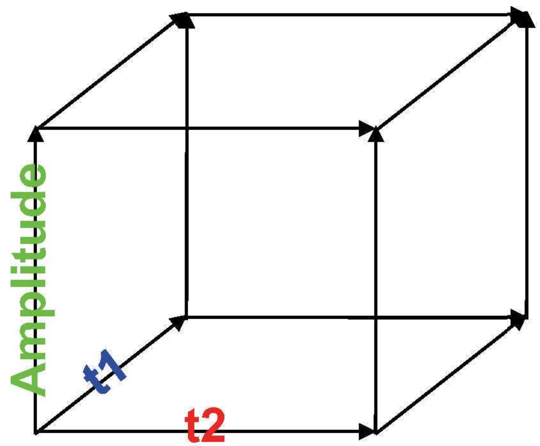


MRI and colors

- **RF and Nuclear magnetic resonance**
- **Obtains voxels in three spatial dimensions**
- **Obtain three values from each voxel: Amp, t1, t2**







Third-World multispectral imaging?

Examples of components pricing for multispectral imaging:

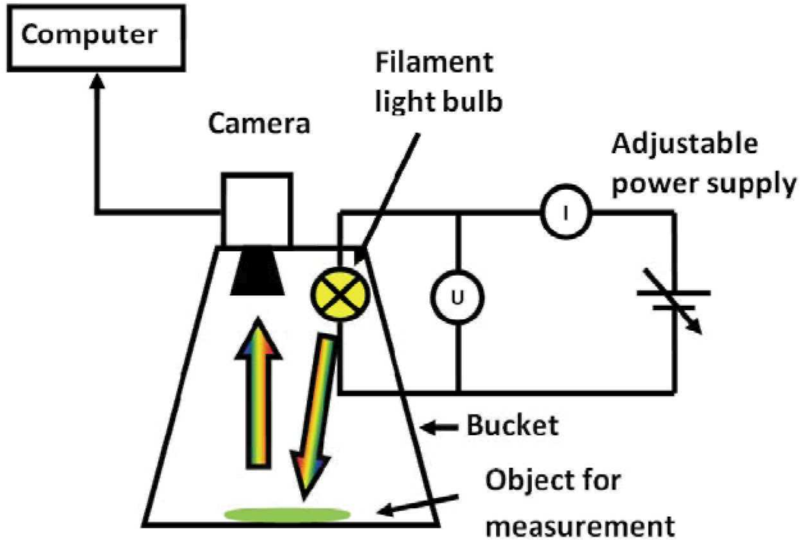
- Multispectral satellite 100 000 000$
- PET/CT scanner 2 000 000$
- Imaging Fourier transform spectrometer 500 000$
- Commercial push broom 200 000$
- IR Focal Plane Array (FPA) imager 100 000$
- Tunable wavelength filter 50 000$
- Optical table 10 000$
- Scientific imager 5 000$
- Optical scanning stage 3 000$
- Fiber spectrometer 1 000$
- Diffraction grating 500$
- Industrial CMOS imager / Commercial RGB camera 200$
- LEGO, LabView microcontroller, steppers, encoders, sensors 200$
- Interference filter 60$
- Absorption filter 20$
- Polarization filter 10$
- Light Emitting Diode, LED 1$
- Google earth, multispectral satellite data 0$!!!

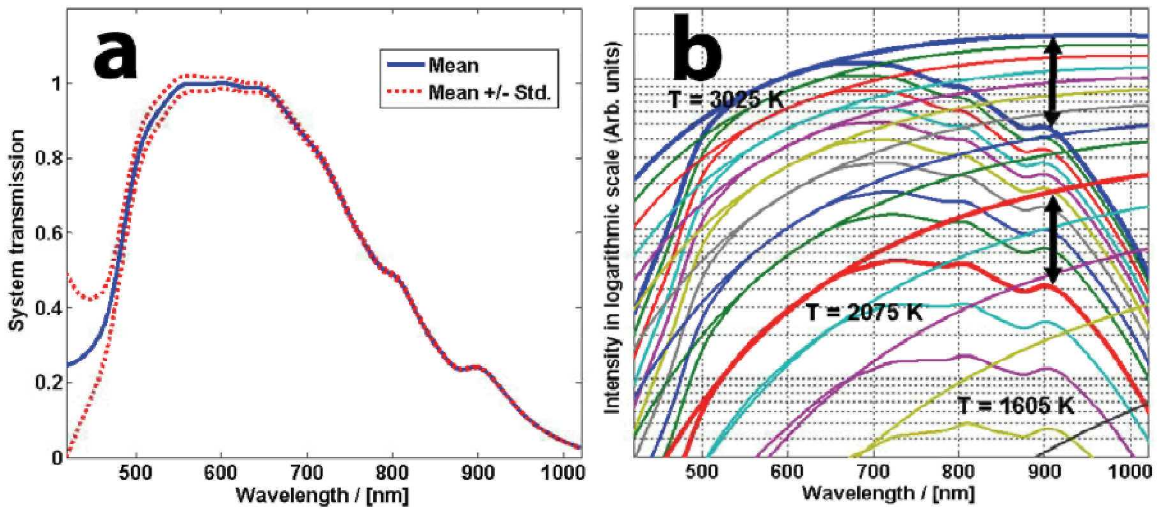
Limit of realism

Wien Shift Imaging

- Simulation of multispectral X-ray

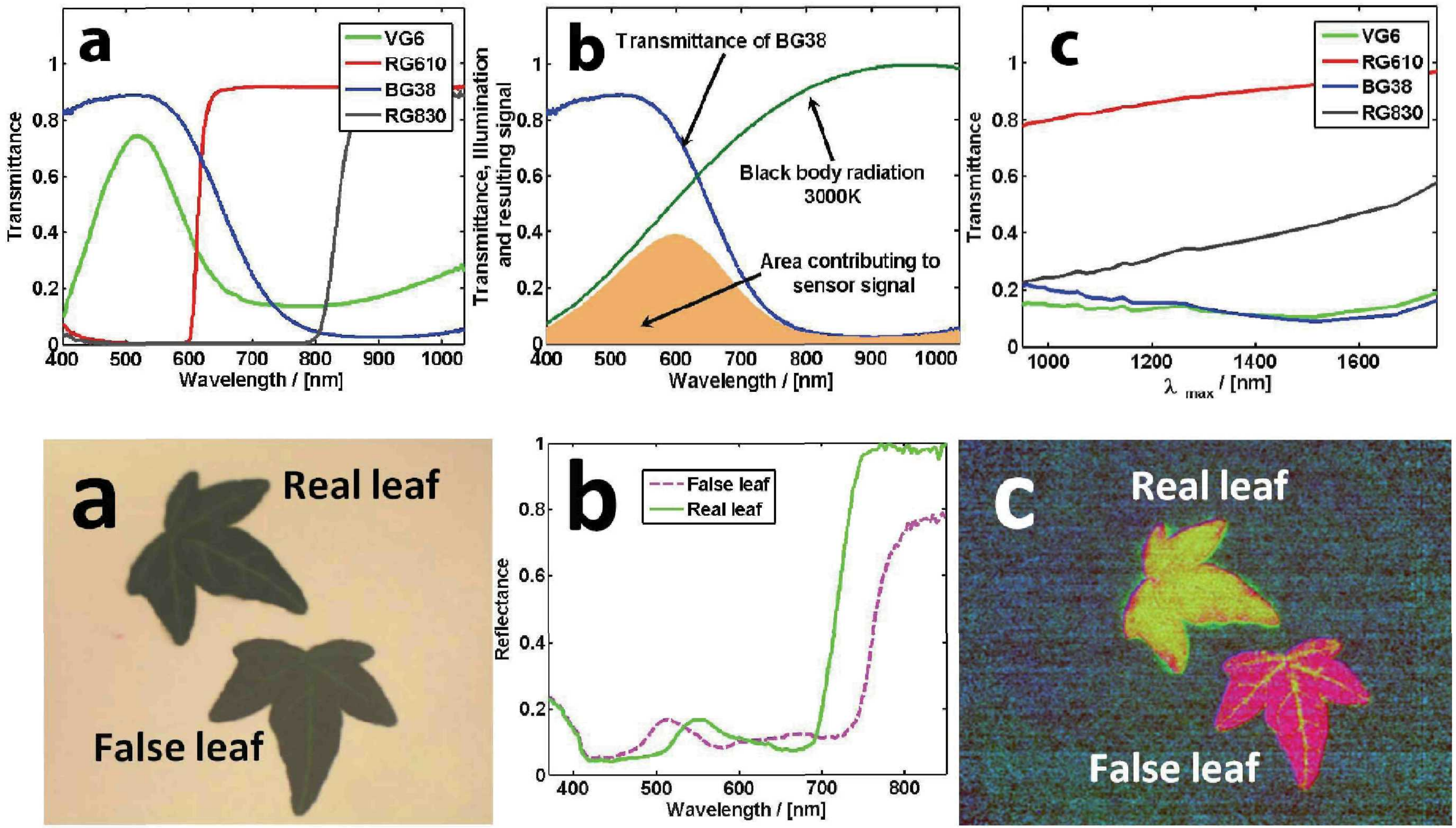
- Educational exercise





M. Brydegaard 2009

Wien Shift Imaging



LED based system for imaging transmission spectroscopy

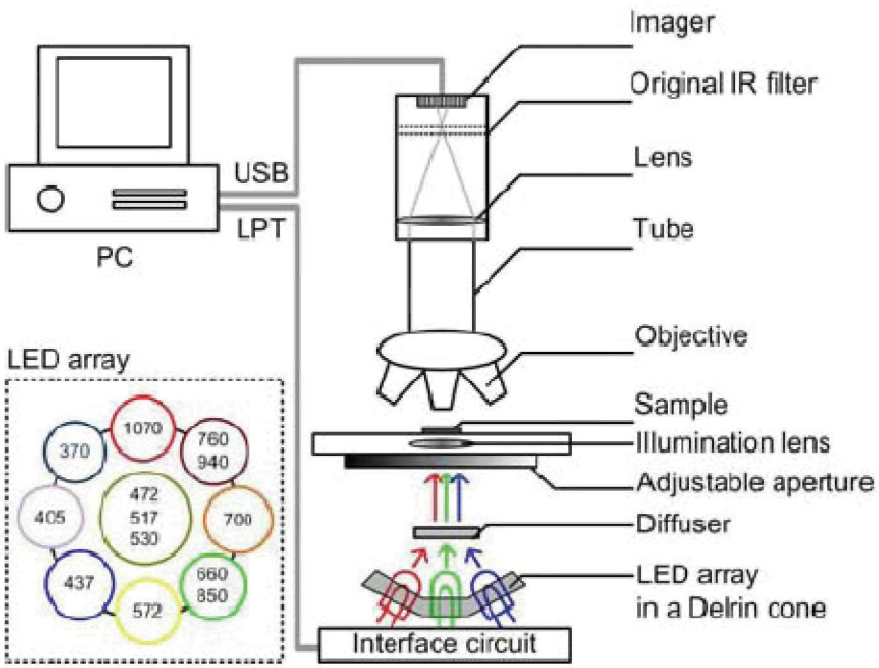


Fig. 1. Arrangement for multi-spectral transmission microscopy employing multiple LED illumination.

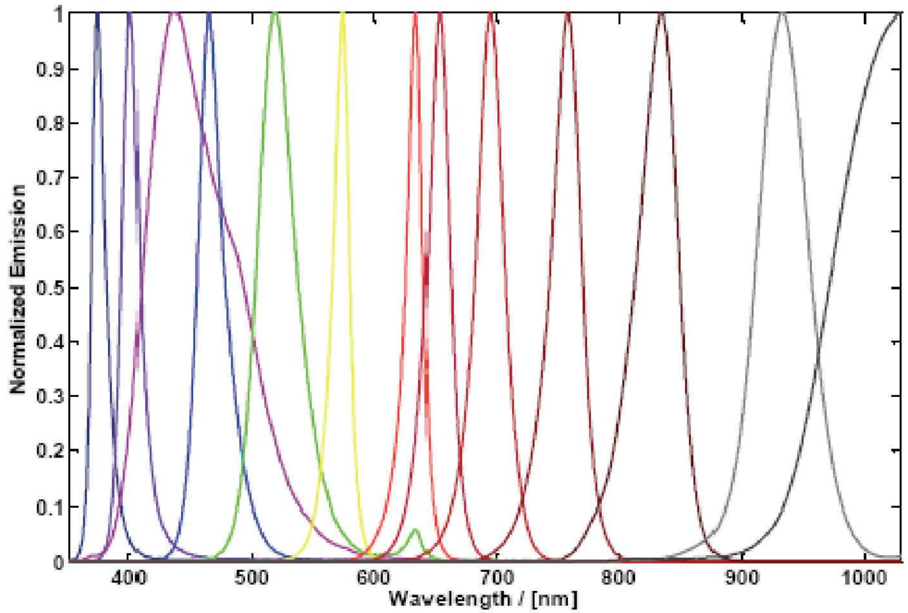


Fig. 2. Normalized spectral emissions of the different LED sources used in the multi-spectral microscope.

The 200$ microscope!

Spectral domain – Hair measurements

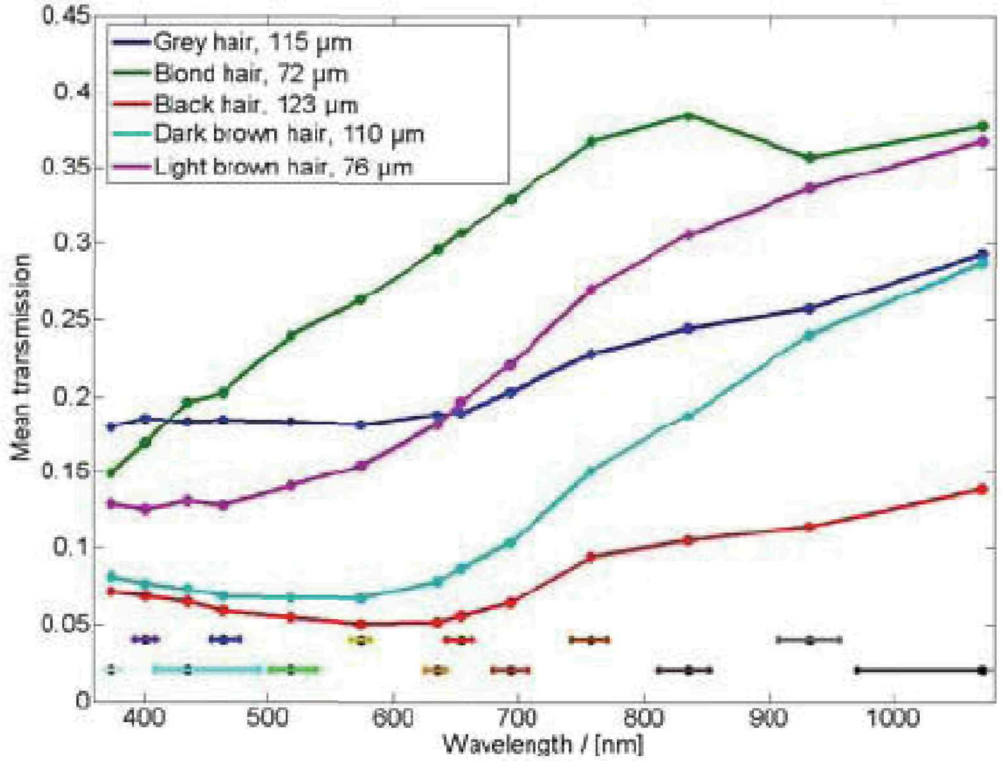


Fig. 5. Transmission spectra of hair strands from five individuals. The symbols in the lower part indicate λ_{max} and FWHM for each spectral band.

Applications:
Forensic science
Zoology
Cosmetics

Spatial and spectral domain

- Spatial – spectral studies
- Assessment of refraction index
- Polarization phenomena

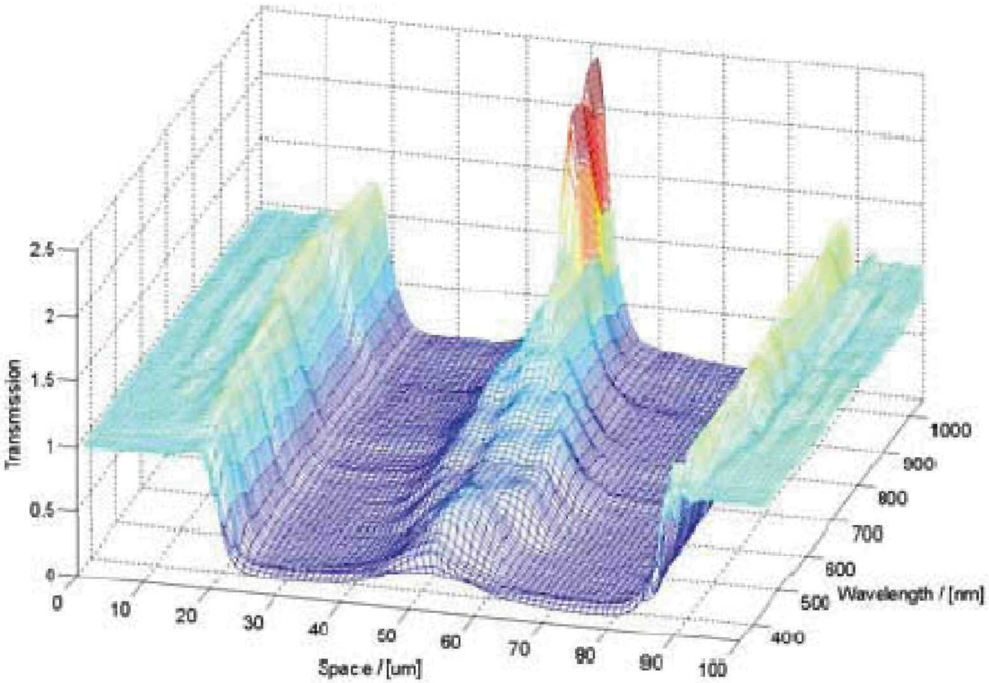
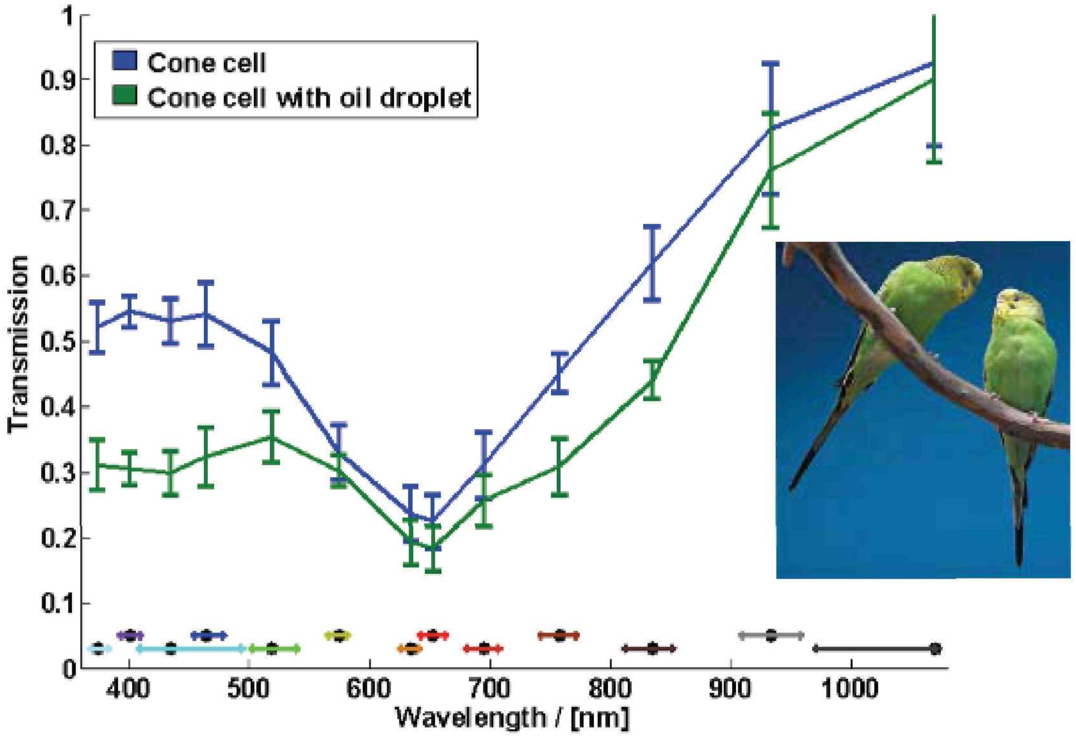
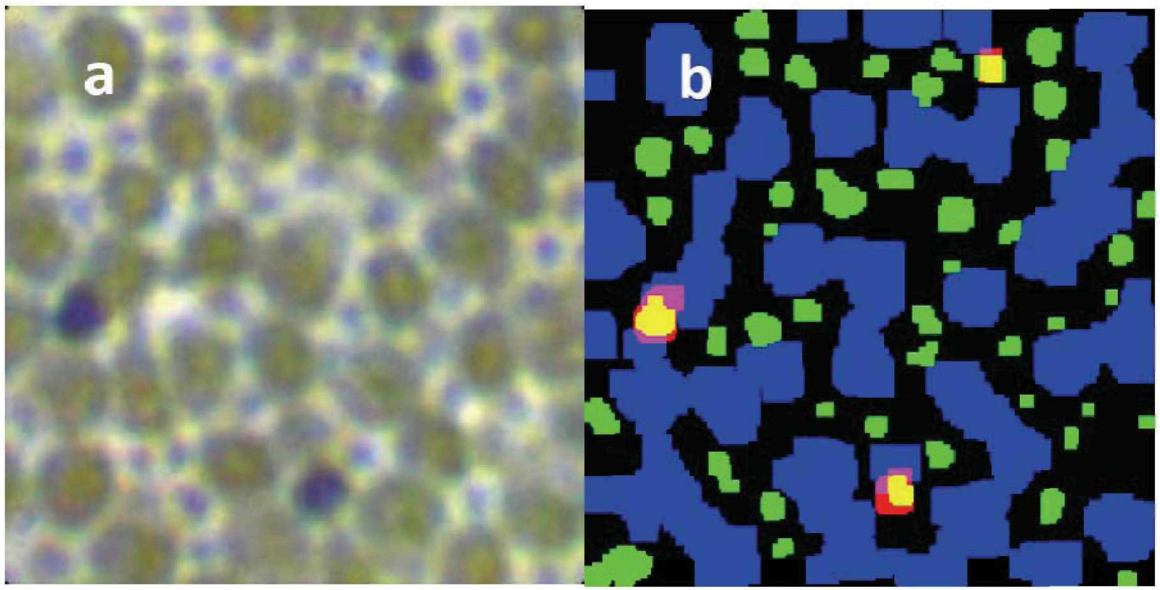


Fig. 6. Transmission cross section for a blond hair, represented in terms of one spatial and one spectral dimension.

Object identification

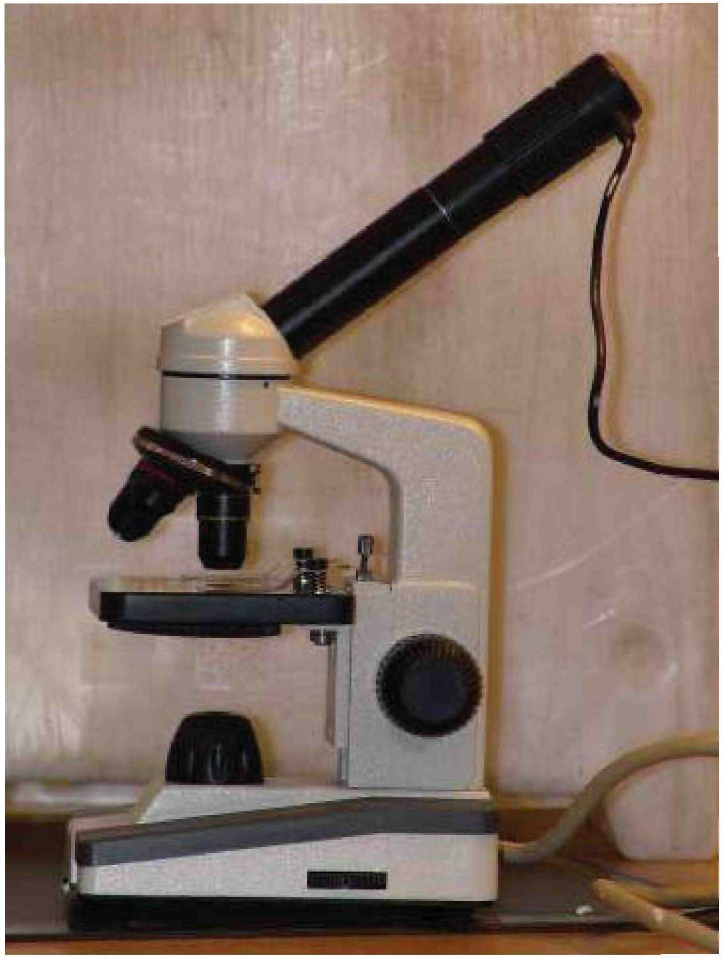
- Spatial stretching by land mark technique
- Information compression
- Multivariate modeling
- Morphological binary operations: erosion and dilation
- False color representation





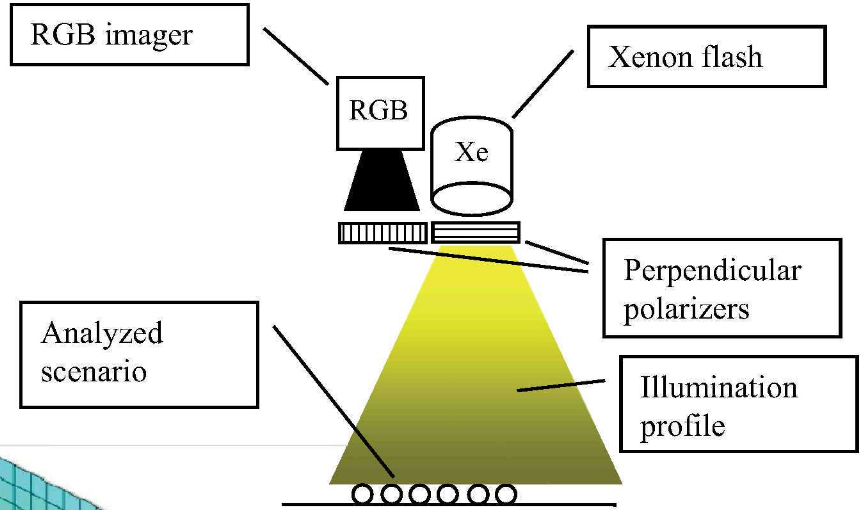
Possible further development

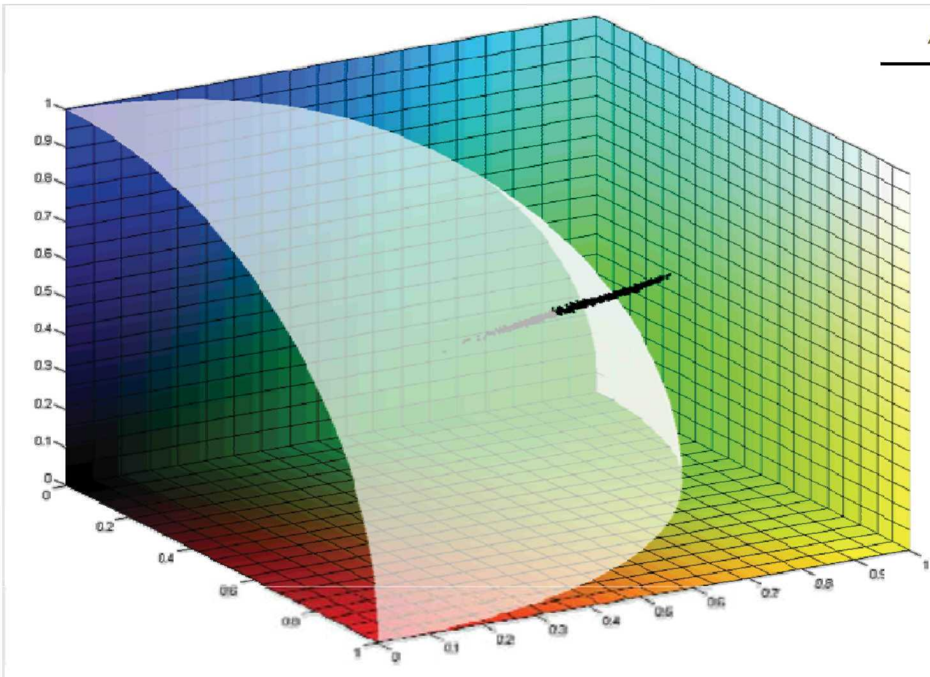
- Synchronization and Flashing
- Improved stray light rejection
- Reflecting objective, automated focus or spatial disconsolation
- Imaging fluorescence spectroscopy
  - Emission excitation acquisition
  - Life time imaging
- Angular discrimination
  - Back scattering geometry
  - Dark field
  - Assessment of μ_a, μ_s, g
  - Multi spectral microscopic tomography
- Polarization studies



RGB imagers

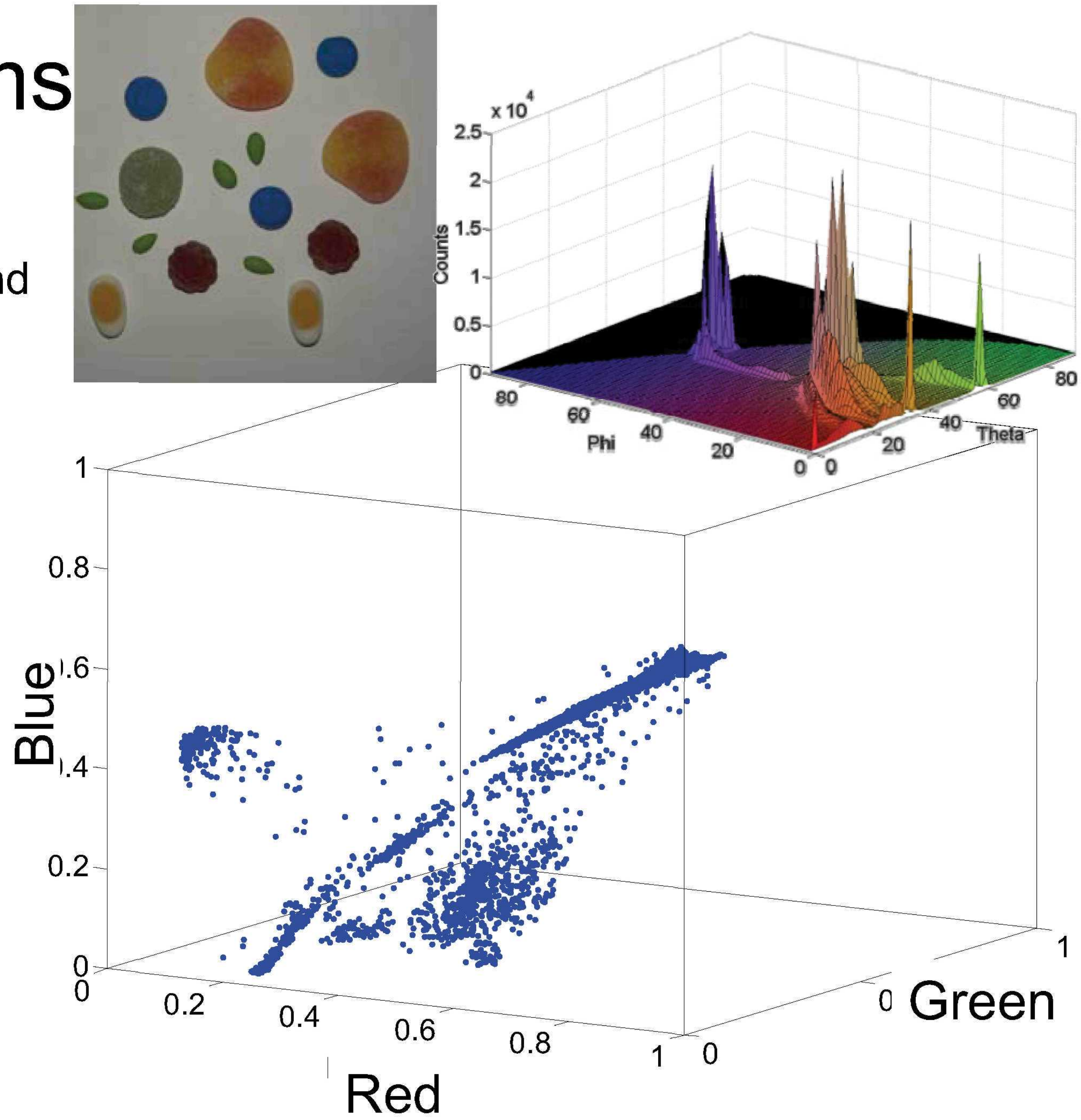
- Controlled illumination
- Ensuring multiple scattering





2D histograms

- Histograms, spectra and images
- Distribution decomposition

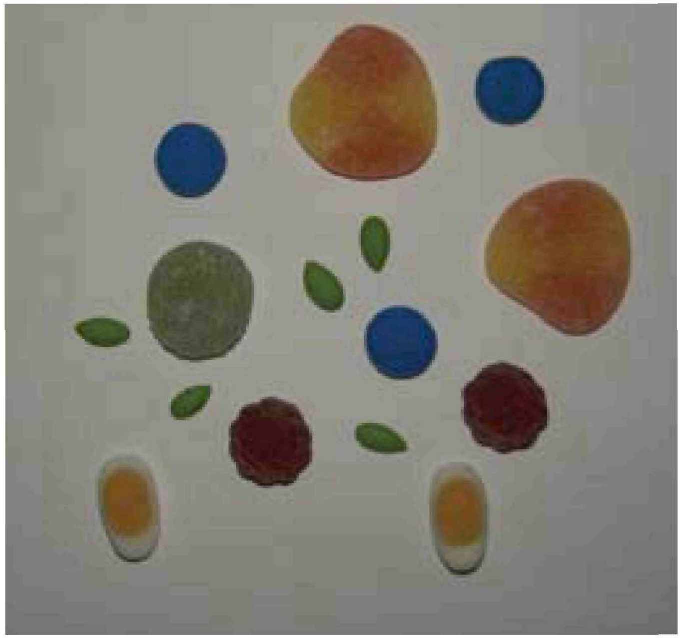


Candy counting, how many?

$Y=F(U)$

$Y=k_0+k_1*u_1+\ldots k_8*u_8$

Residuals and quality



Type	Peaches	Green gum	Red gum	Blue tablet	Green seed	Eggs	Orange tablet
Reality	2	1	2	3	5	2	0
Estimated	1.9707	1.0408	1.8816	2.7244	5.3621	2.0872	0.0358

Multi-spectral Imaging Basics Part 3: Color spaces, Data handling and Contrast functions

Abdus Salam International Center for Theoretical Physics, 2009

Mikkel Brydegaard

Division of Atomic Physics

Lund University, Sweden

#2

Descretization

Subject to discretization:	Light intensity	Light energy	Space	Time
Domain:	Dynamical -	Spectral -	Spatial -	Temporal -
Discretized by:	Bits	Spectral bands	Pixels / Voxels	Frames
Resolution:	Dynamic -	Spectral -	Spatial -	Temporal -
Res. limited by:	Signal to noise ratio / photons	Channel / illumination bandwidth	Point spread function	Exposure time / flash envelope
Range:	Dynamic -	Spectral -	Field of view	Recording time

Table 1: Comparison terms associated with discretization along various domains.

Summary: Spectral sources

Sources of spectral data

- There are infinetily many ways to get a finite amount of rather incomprehensible spectral data, containing information arising from a finite amount of physical phenomena.
- Regardsless whelther we access values for nuclear spin, x-ray absorption, fluorescence, reflectance, scattering coefficients or whatever phenomena we end up with the vector **U** filled with various spectral properties from each measurement **n**.

$$U_n = \begin{bmatrix} u_1 & u_1 & \dots & u_K \end{bmatrix}$$

The goal

Question:

How to interprete **U** to get the information we are interested in?

How to find spectral properties of interest?

How to discard irrelevant data in spectral properties?

Question:

How to determine **F** when:

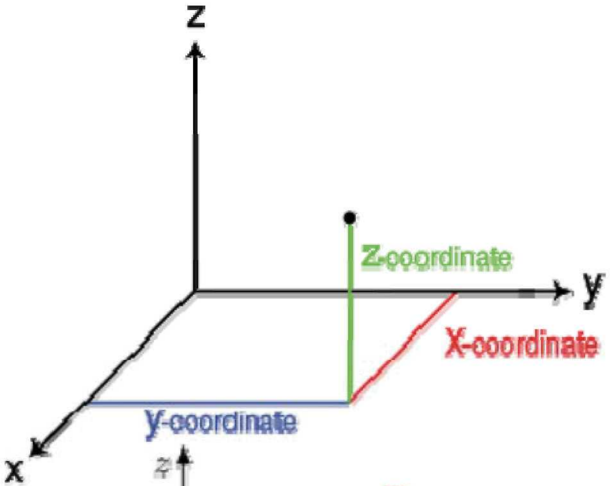
Y = F(U),

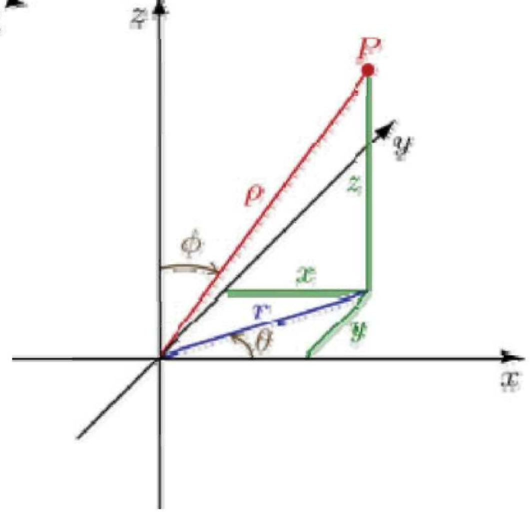
where **Y** is the answer to question we are interested in, and **U** is a list of spectral properties?

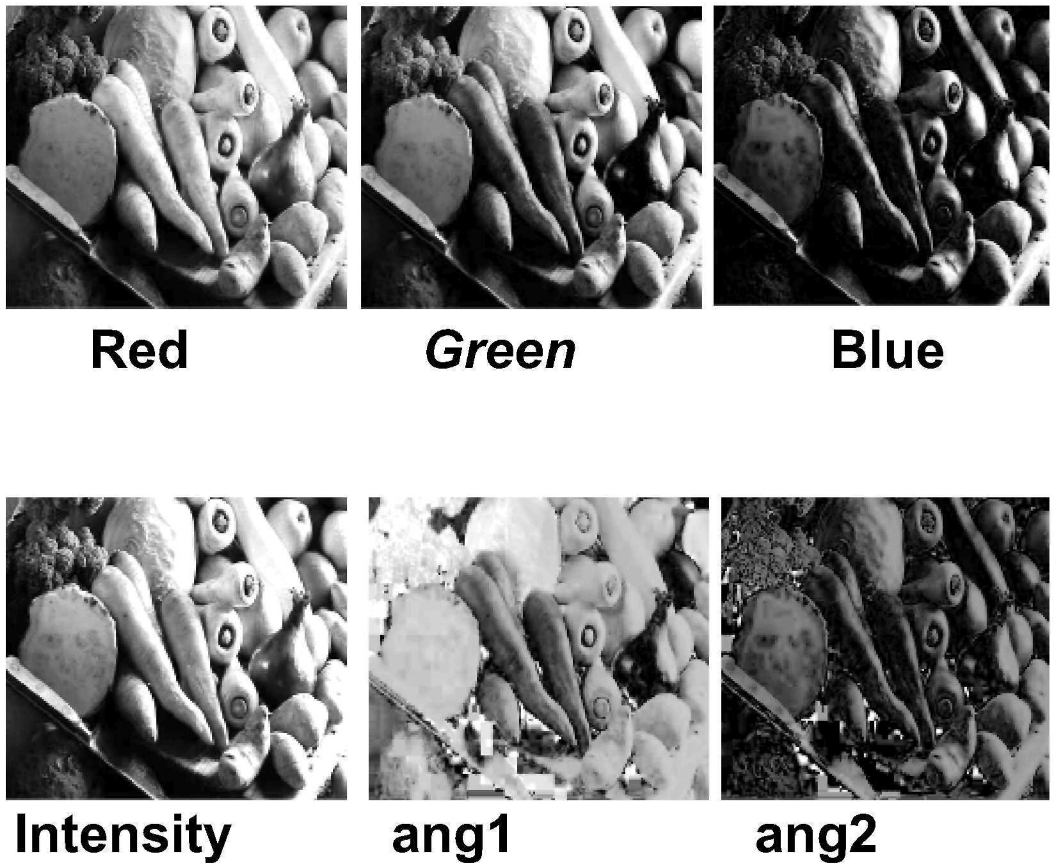
We will refer to **F** as the spectral model or contrast function.

Colour spaces

- **Cartesian coordinates [dim1 dim2 dim3 ... dimN]**
- **Spheric coordinates [abs ang1 ang2 ... angN]**
- **Conical**
- **Others**
  - ***What are the units in either of the colour spaces?***

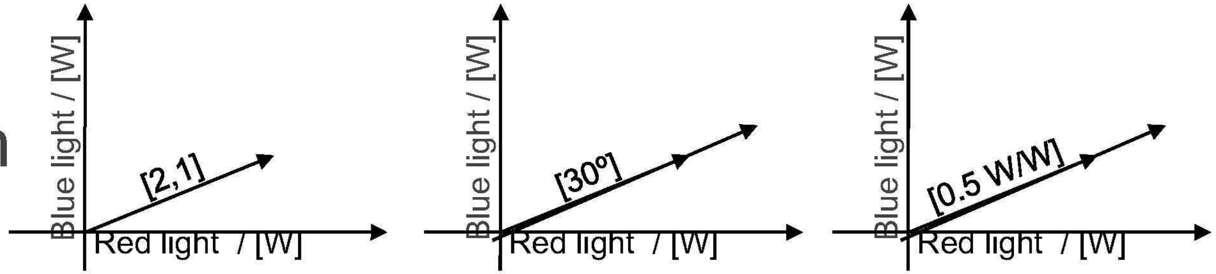




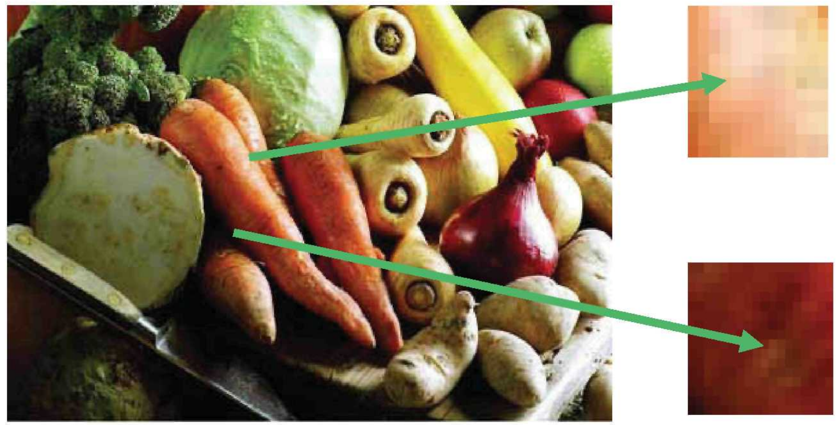


Unit-less functions

- Unitless functions cancels effects of variating illumination and shadows suposing that the illumination is ”white”

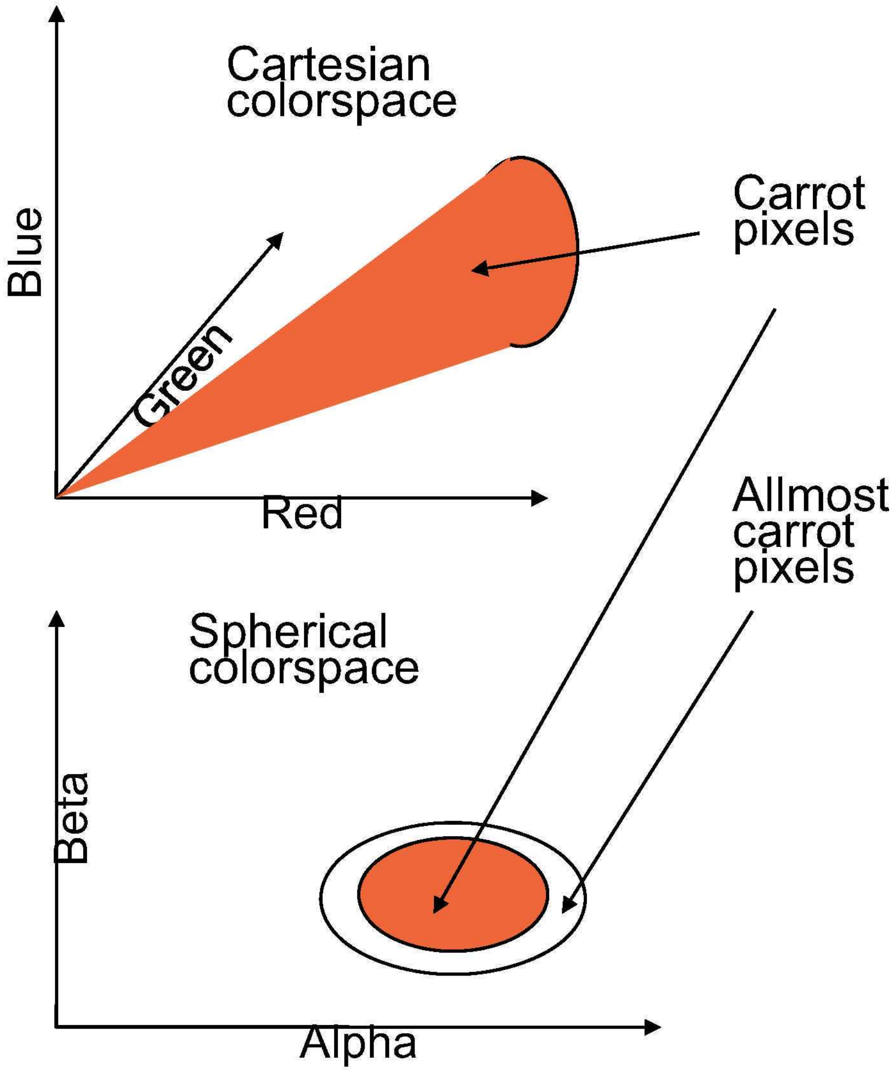


- Units can be cancelled by rational functions or trigonometrical functions

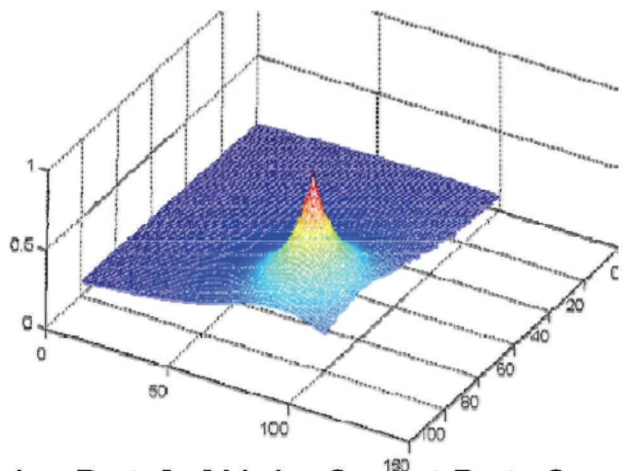


- When units are cancelled brightness information is lost

Dividing colour spaces



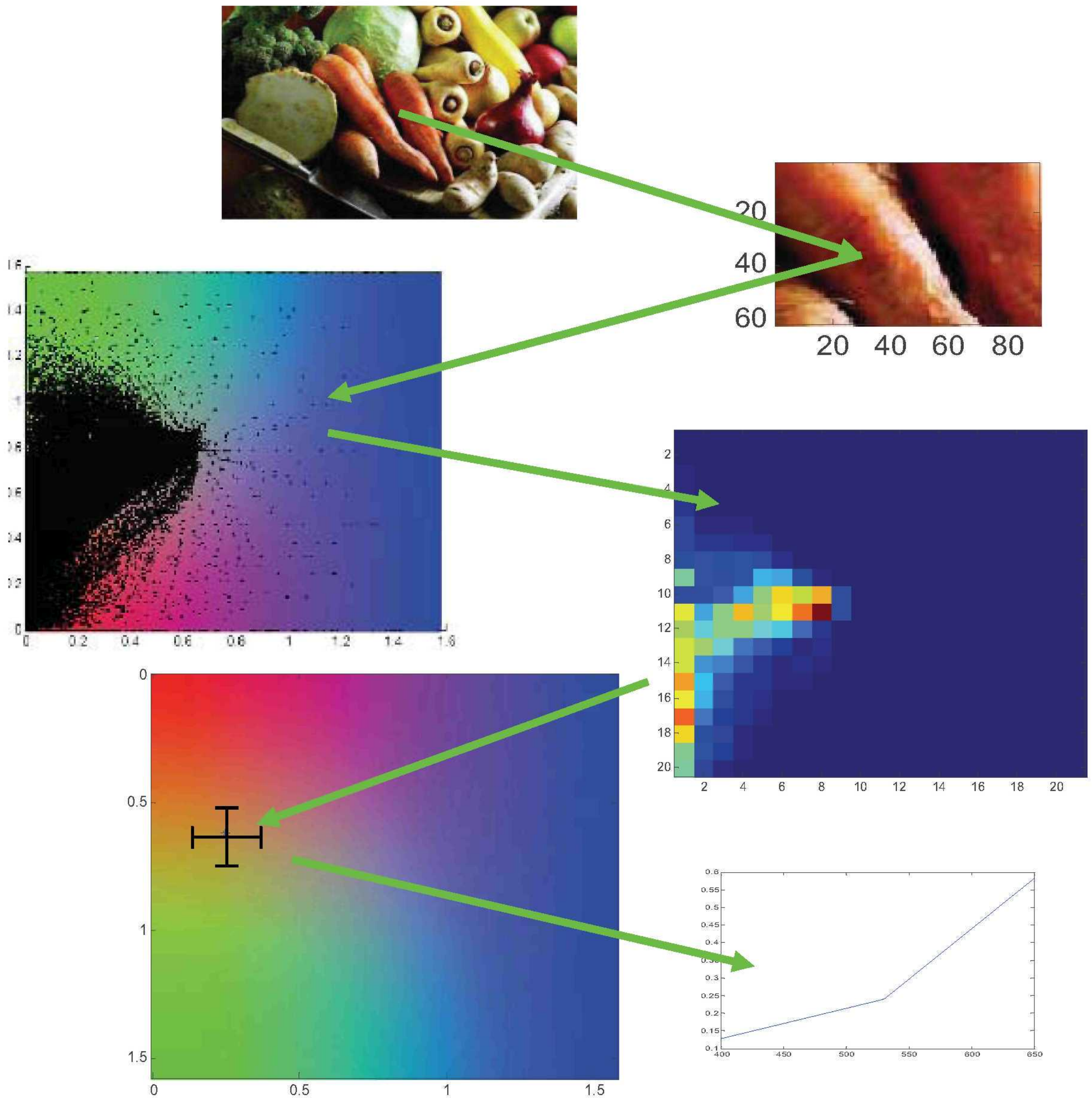
- **During the creation of a contrast function one divide the N-dim space into regions**
- **If irrelevant information is discharged the formulation will be easier**
- **Gradual functions determines to which grade a pixel is a carrot**
- **Example:**



isCarrot=1/(1+norm([Alpha Beta]- [AlphaCarrot BetaCarrot]))

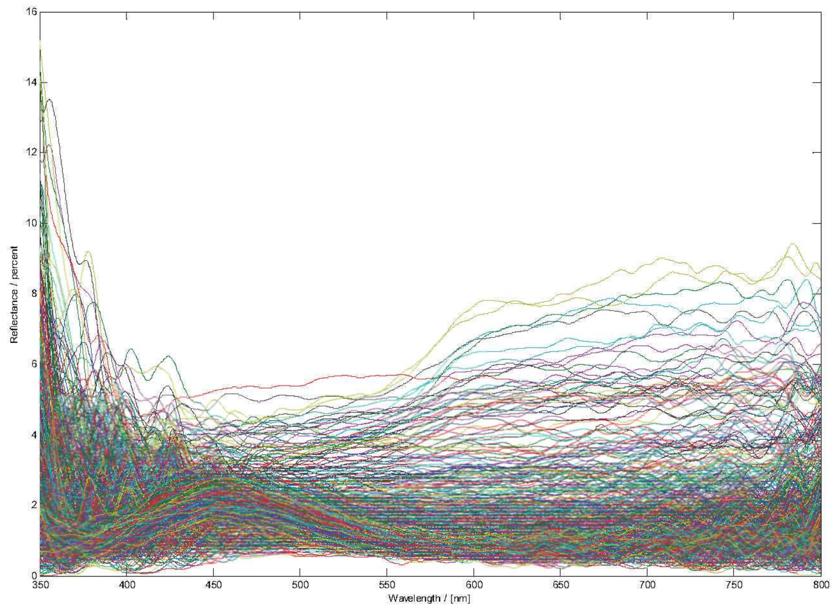
Identification in colour spaces

- 3 colour channels: → 2 unit-less angles in a spherical colour space
- The pixels of the carrots: become locations in the color plane
- The distribution has a shape, mean-vector, and variations.
- Imagine how the distribution would look like with more color channels



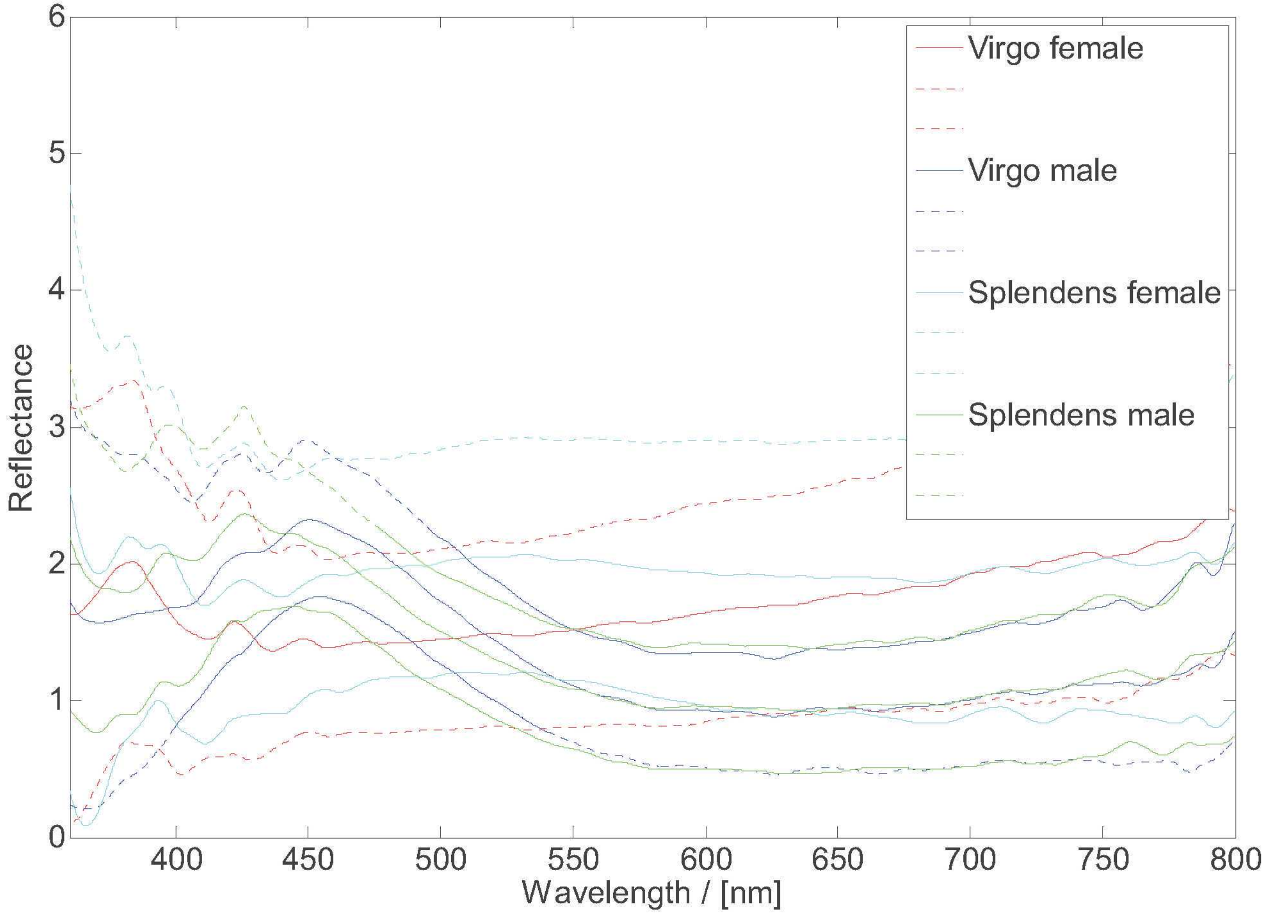
Interpretation of multivariate data

- Reflectance spectroscopy on damselflies, with LIDAR applications
- Two species, two sexes
- Reflectance and transmittance
- 2327 wavelength bands
- 699 individuals
- Who is who?



Mean values

- Your left foot is standing in liquid nitrogen.
- Your right foot is standing in boiling oil.
- Do you feel ok in average?



1. Reducing redundancy: Separating scrap from information

Arranging groups of spectral data:

Correct answer from a professional, e.g. IsMale

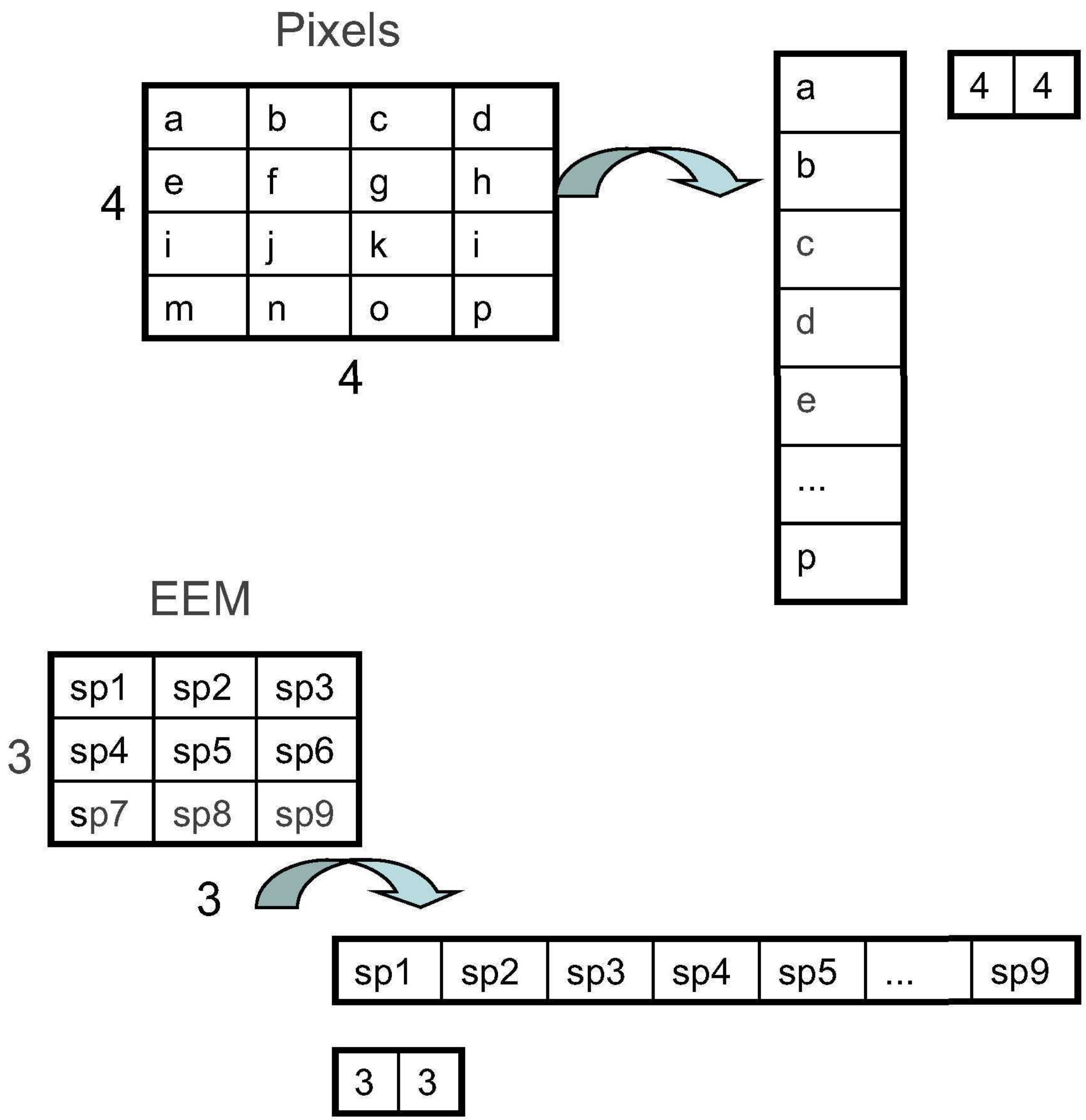
Measured parameters, e.g. R_λ

$$Y = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \ldots \\ 0 \end{bmatrix} \qquad M = \begin{bmatrix} m_{11} & m_{12} & \ldots & m_{1K} \\ m_{21} & m_{21} & \ldots & m_{2K} \\ \ldots & \ldots & \ldots & \ldots \\ m_{N1} & m_{N2} & \ldots & m_{NK} \end{bmatrix}$$

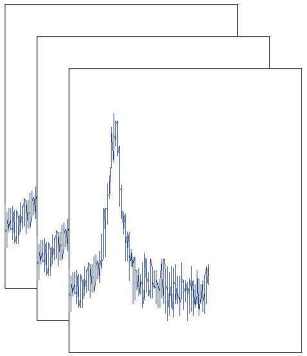
Measurements, individuals, pixels or voxels

Rearranging data

- If individual measurements have are arranged in more than one dimension, e.g. pictures, we can temporally discard the dimensions and restore them after analysis.
- If measured parameters have a higher dimensionality than one, e.g. EEM spectroscopy, we can do the same.
- We can always transform to a 2D matrix representation.
- Remember original dimension
- Scale the variance when merging
- Matlab: reshape



Principal Component Analysis (PCA)
Singular Value Decomposition (SVD)



Set of spectra, M

SVD / PCA

New base functions, or *principal components* (PCs, V)

Eigen values σ in Σ related to residuals and importancy of each PC

Linear coefficients or *loadings,* U, reduced representation of M

$$M = U \Sigma V^*$$

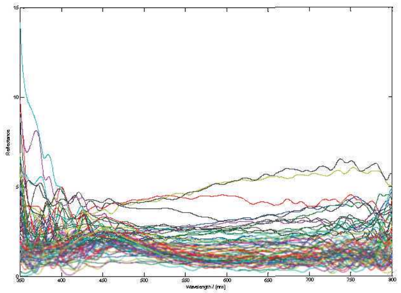
Matlab: svd / princomp

A new representation of M

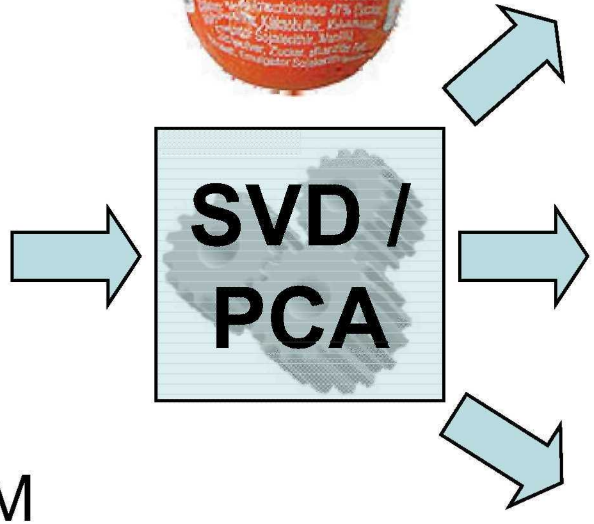
$$m_n = c_{n,1}\sigma_1 PC1 + c_{n,2}\sigma_2 PC2 + c_{n,3}\sigma_3 PC3 + \ldots$$

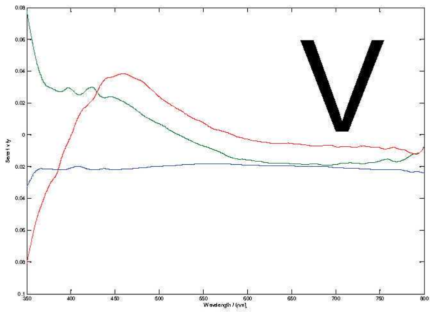
- $c_{n,1}$: Coefficient describing the amount of ingredient PC1 in individual n
- $PC1$: 1st base function, or principal component
- σ_1: Eigenvalue, importance of component 1
- m_n: Spectrum for individual n
- $+ \ldots$: Truncation of representation

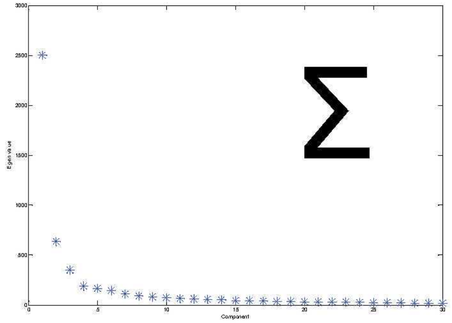
Principal Component Analysis (PCA)
Singular Value Decomposition (SVD)

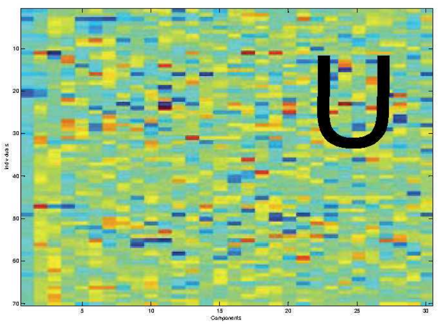


Set of spectra, M





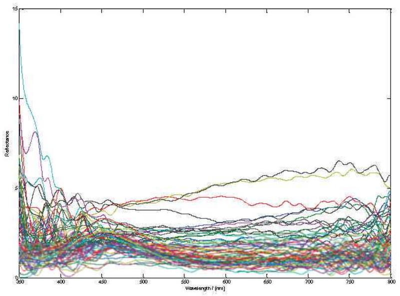




$$M = U \Sigma V^*$$

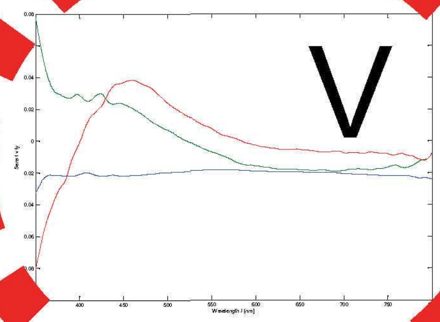
Matlab: svd / princomp

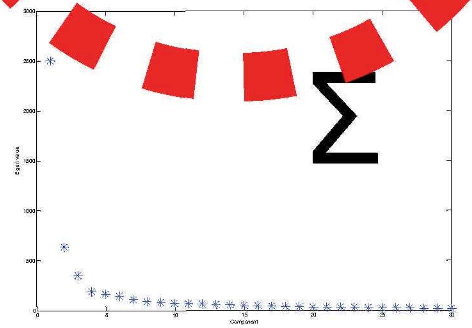
Principal Component Analysis (PCA)
Singular Value Decomposition (SVD)

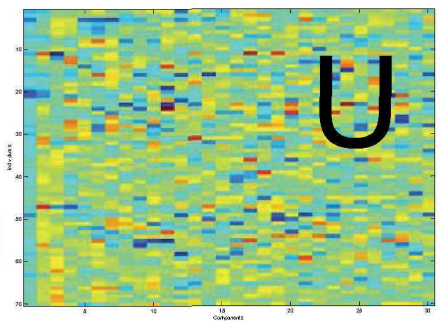


Set of spectra, M

SVD / PCA





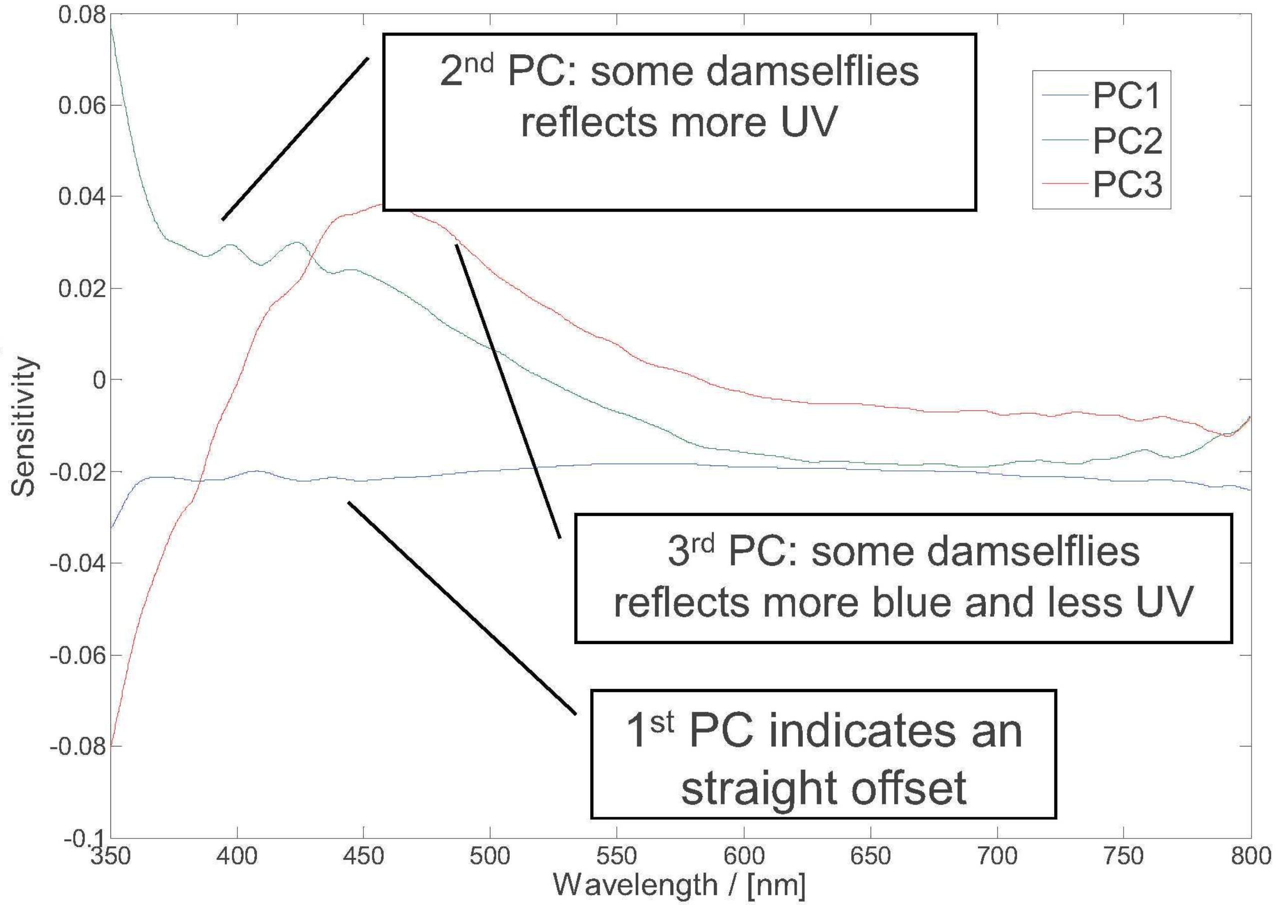


$$M = U \Sigma V^*$$

Matlab: svd / princomp

V, the new base spectra

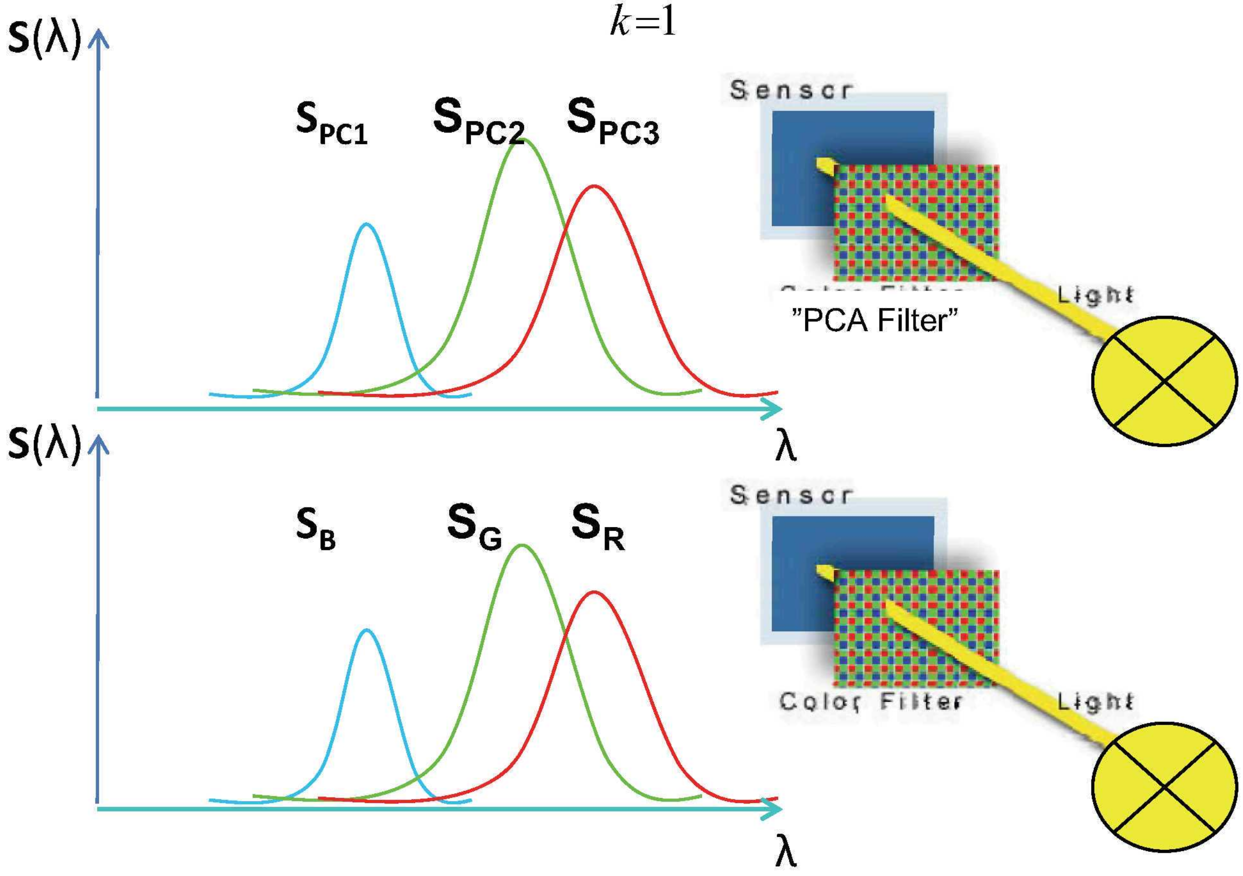
- These new base functions are generated from the variance in the set of spectra
- This information indicates were in the spectrum the samples varies
- A perfect instrument or LIDAR to classify damselflies, should have these sensitivity curves
- Note: Orthogonal base functions: PC1'*PC2=0



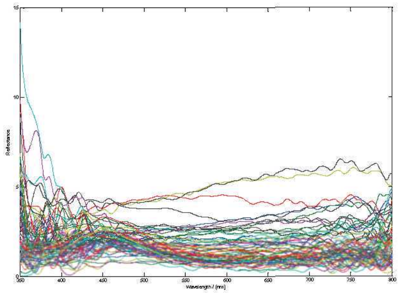
SVD/PCA - Sensitivity Analogy:

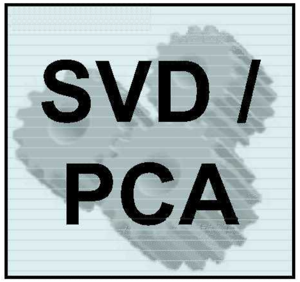
We realize that what fall into the loading c_1 is in fact R seen with the sensitivity PC1. Only that this time the sensitivity is optimized for the variance in the spectral data set.

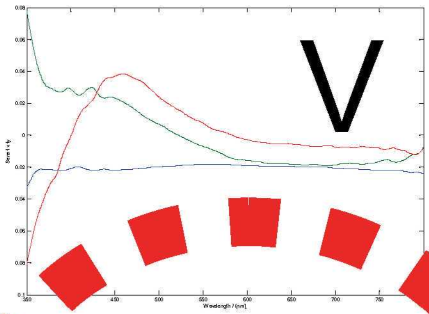
$$c_1 \sim \sum_{k=1}^{K} M_k PC1_k$$

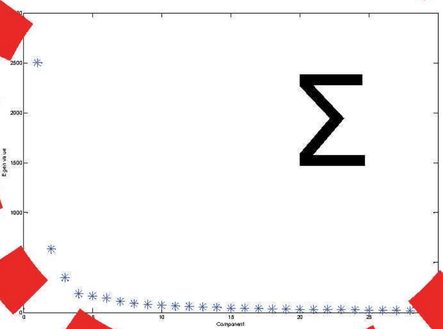


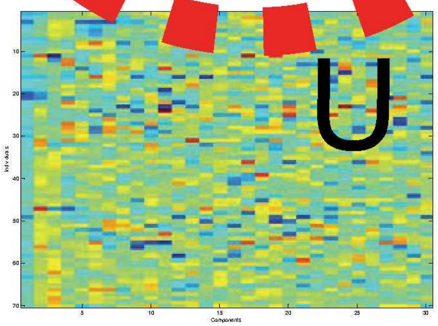
Principal Component Analysis (PCA)
Singular Value Decomposition (SVD)











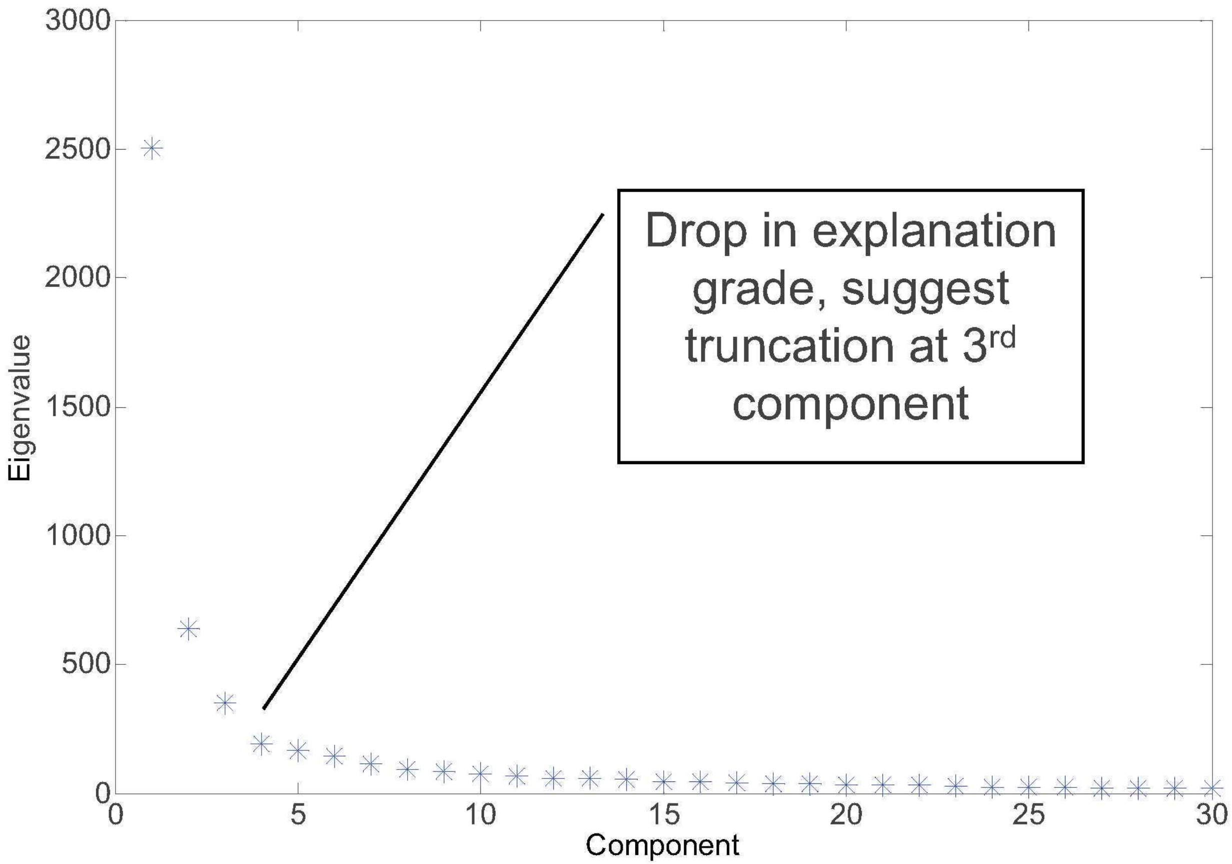
Set of spectra, M

$$M = U \Sigma V^*$$

Matlab: svd / princomp

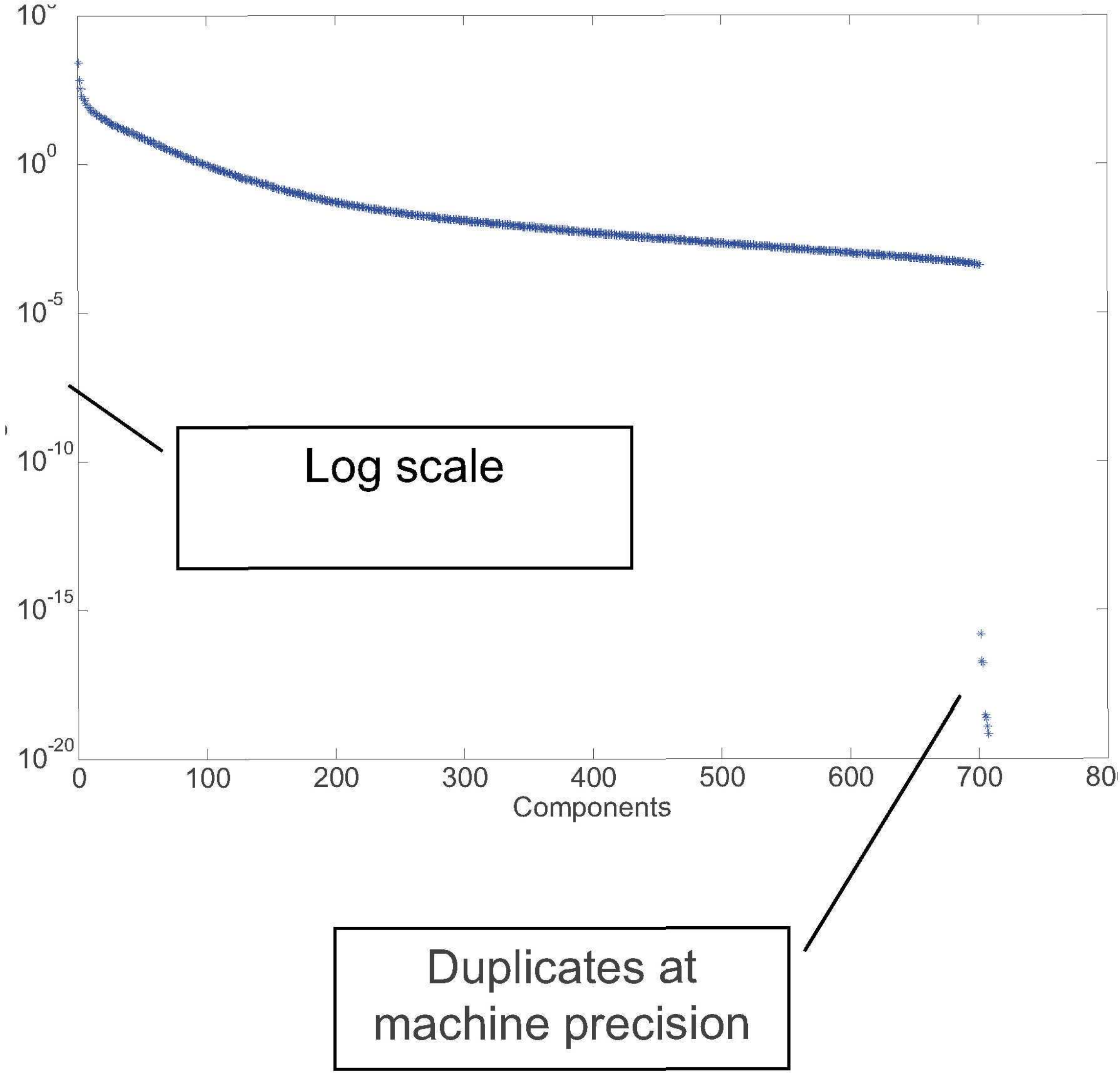
Σ, the Eigen values

- $\sigma_1 > \sigma_2 > \sigma_3 > ....$
- The importance of each component is given by the Eigen values
- The number of independent spectral components are seen as a drop
- The truncation of the representation should be based on the Eigen values
- The sum of remaining Eigen values is the residual.
- Related to Akaike information criterion

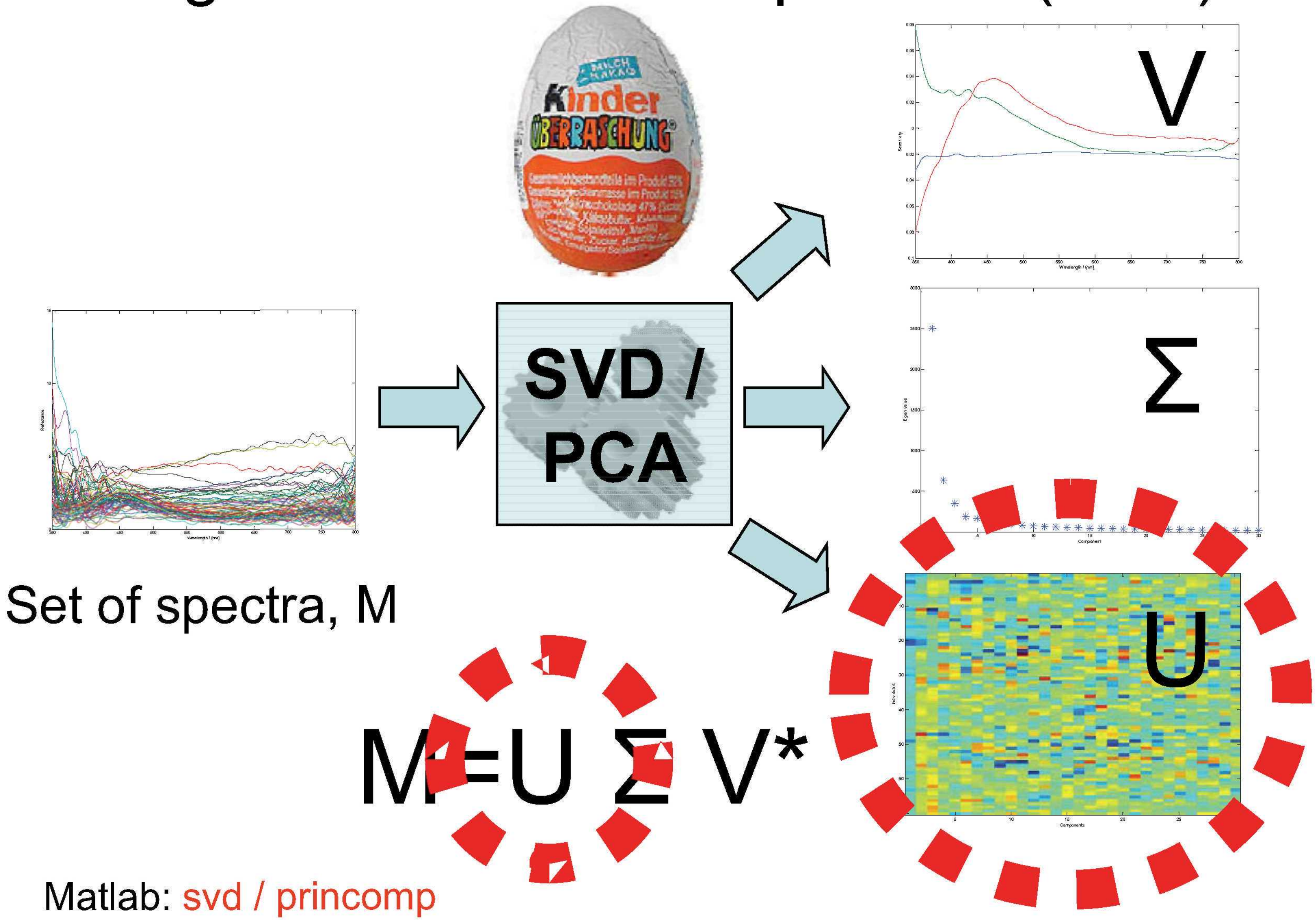


Reveal dublicate data

- How can you you know if you have two equal spectra in a set of 699 spectra?
- How can noise end the career of an untrue scientist



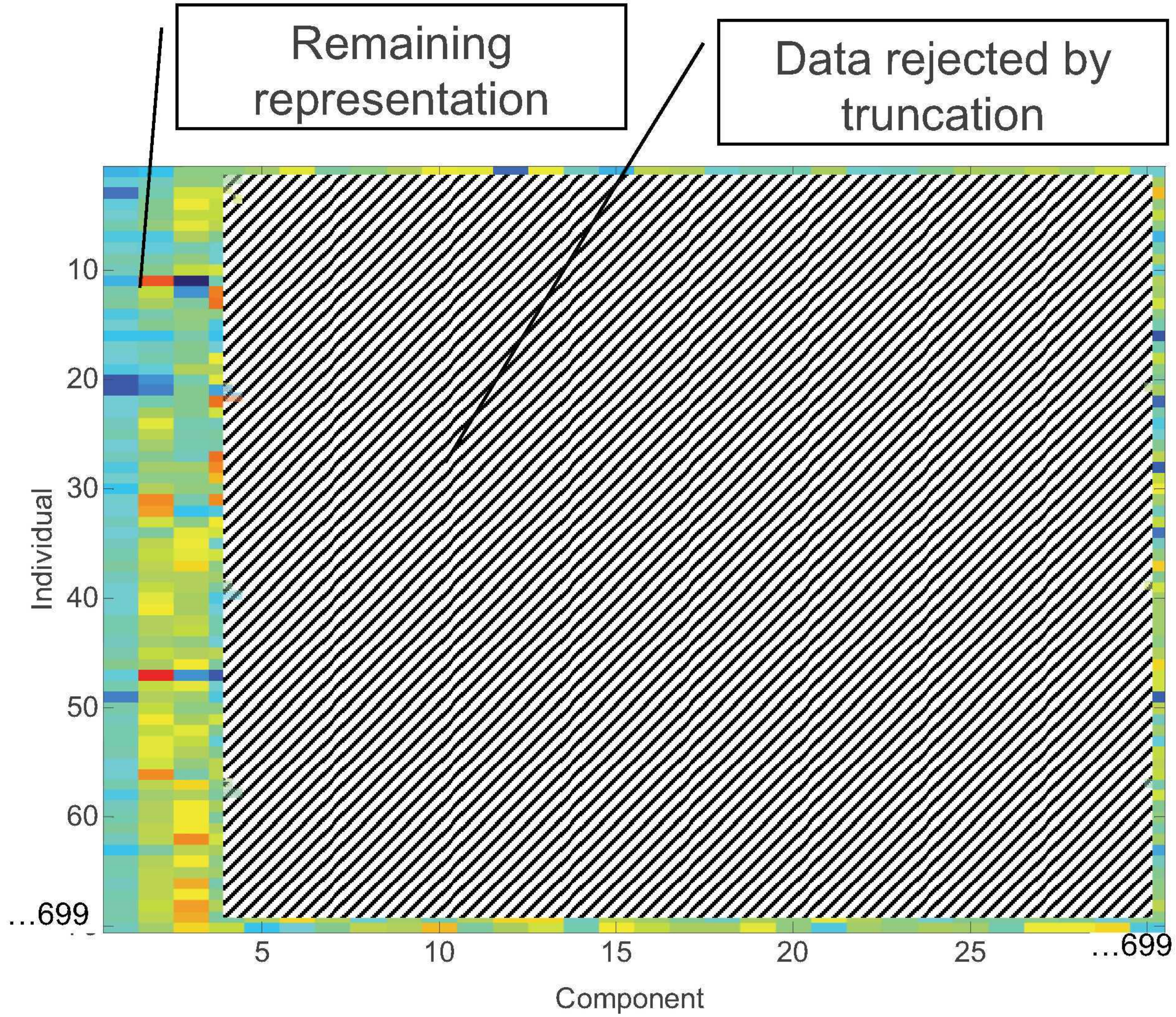
Principal Component Analysis (PCA)
Singular Value Decomposition (SVD)



Matlab: svd / princomp

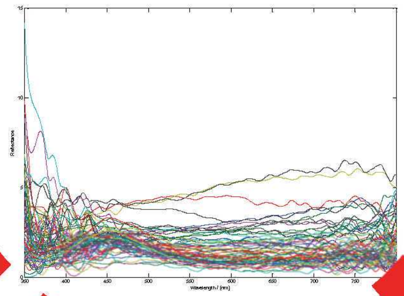
U, the reduced representation

- The spectrum of each individual is now represented by three coefficients
- The information survives, the redundancy and noise dies
- We can reconstruct M from a truncated U, Σ and V



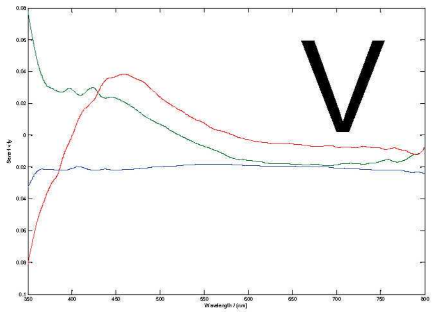
Principal Component Analysis (PCA)
Singular Value Decomposition (SVD)

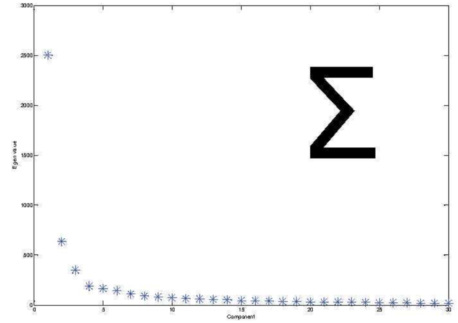
Last egg for today

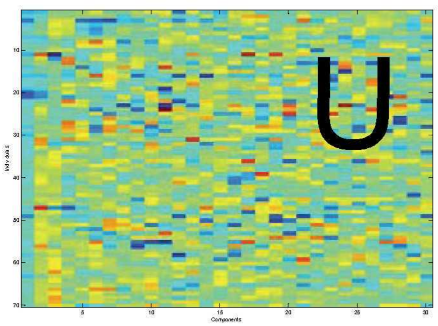


Set of spectra, M

SVD / PCA



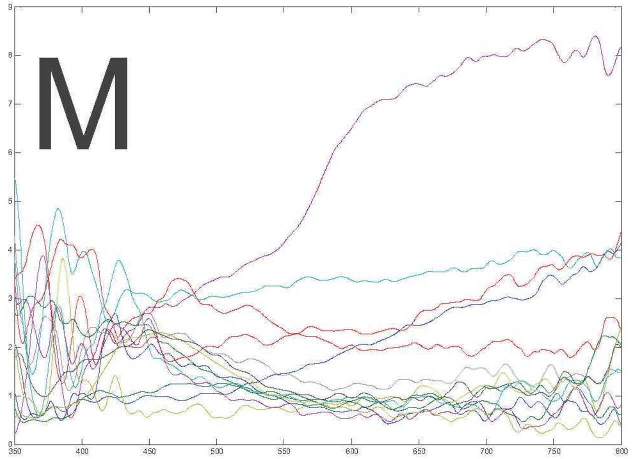


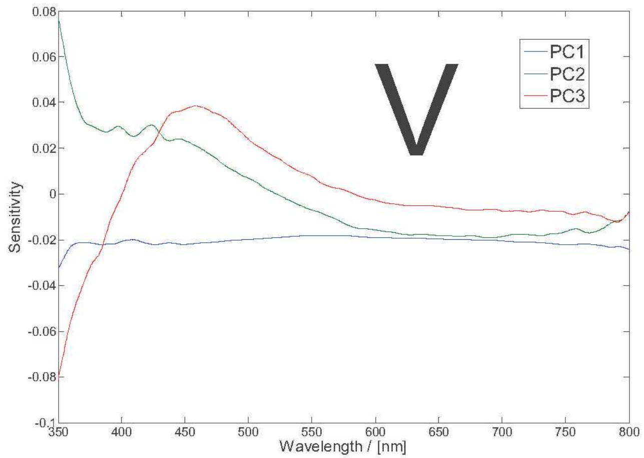


$$M = U \Sigma V^*$$

Matlab: svd / princomp

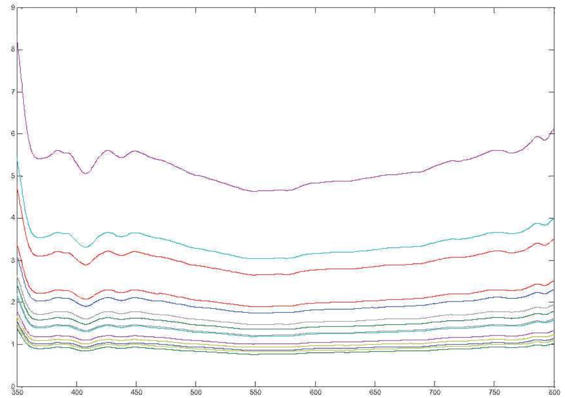
Reconstruction of M





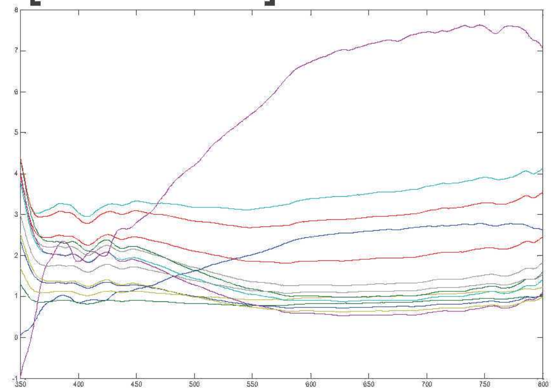
[PC1]

M_{red}



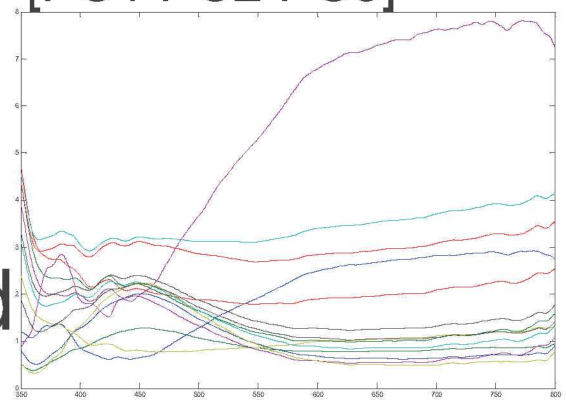
[PC1 PC2]

M_{red}



[PC1 PC2 PC3]

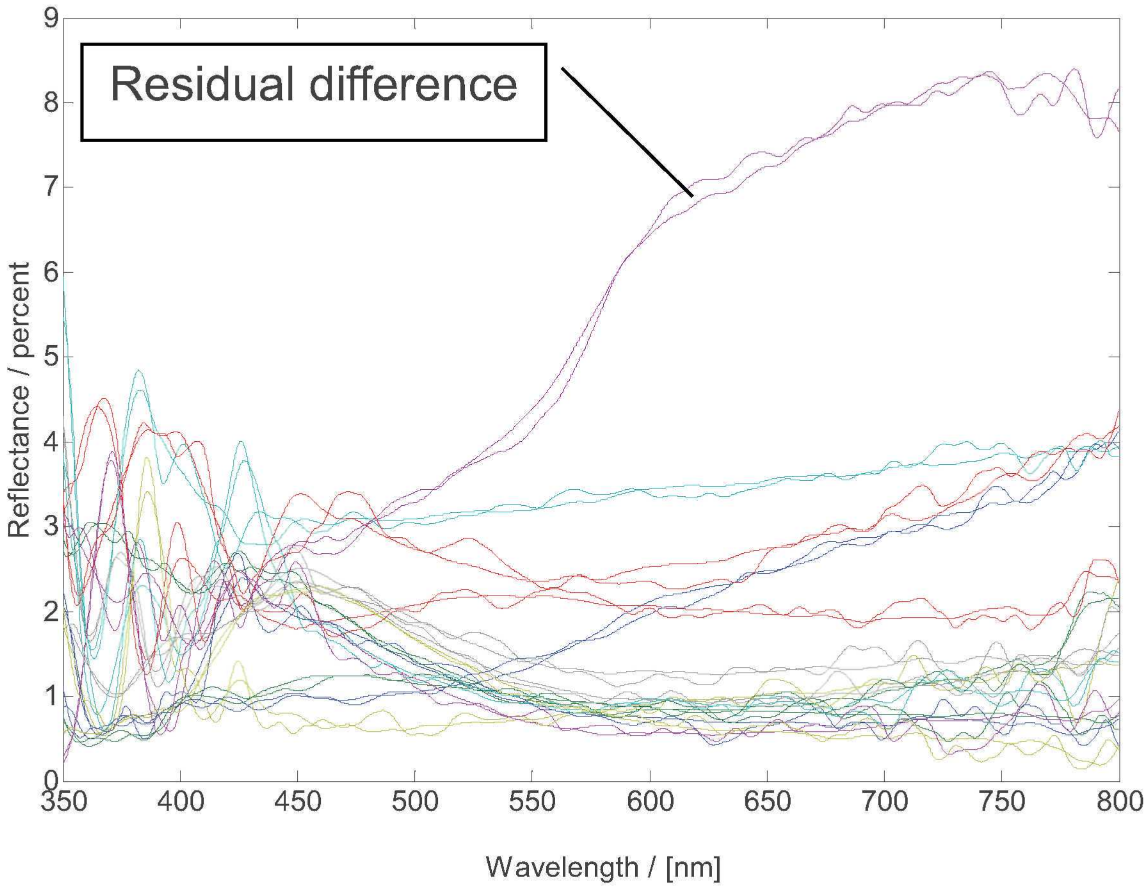
M_{red}



Residuals

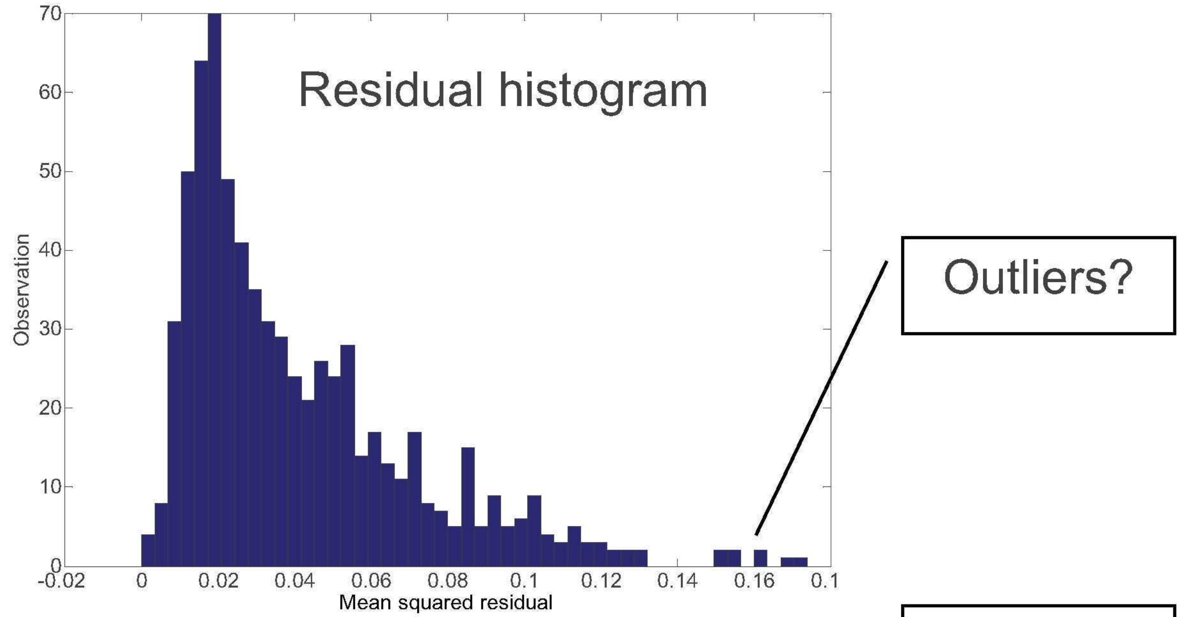
- Sum of squared residuals tells how well an individual n can be explained by the truncated representation

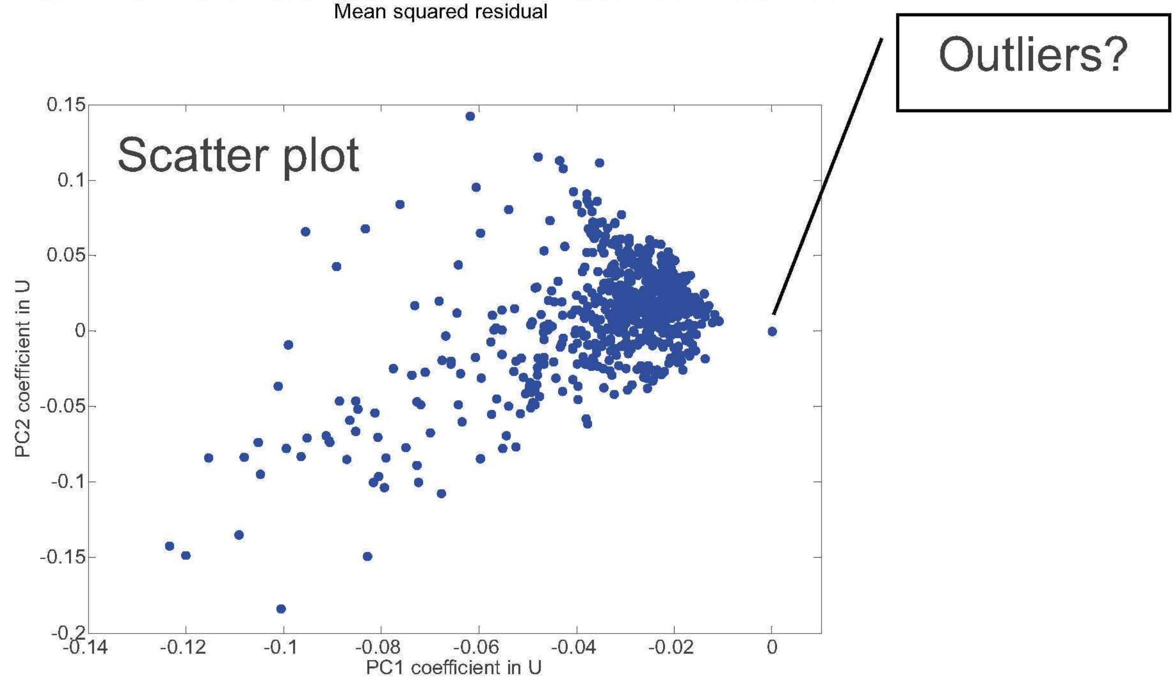
$$res_n = \sum_{k=1}^{K} \left(M - M_{red} \right)^2$$



Outliers

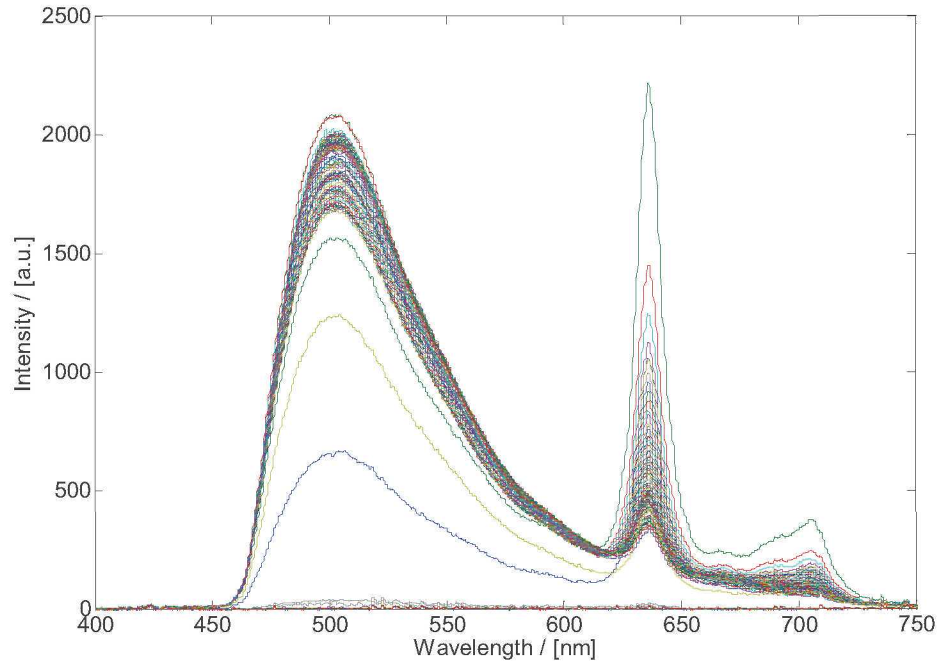
- Histogram of mean squared residuals
- Finding poorly compressed individuals
- Scatter plot
- Looking closer at extreme values in U
- Redo the SVD / PCA after exclusion!
- Matlab: find(res>0.14)

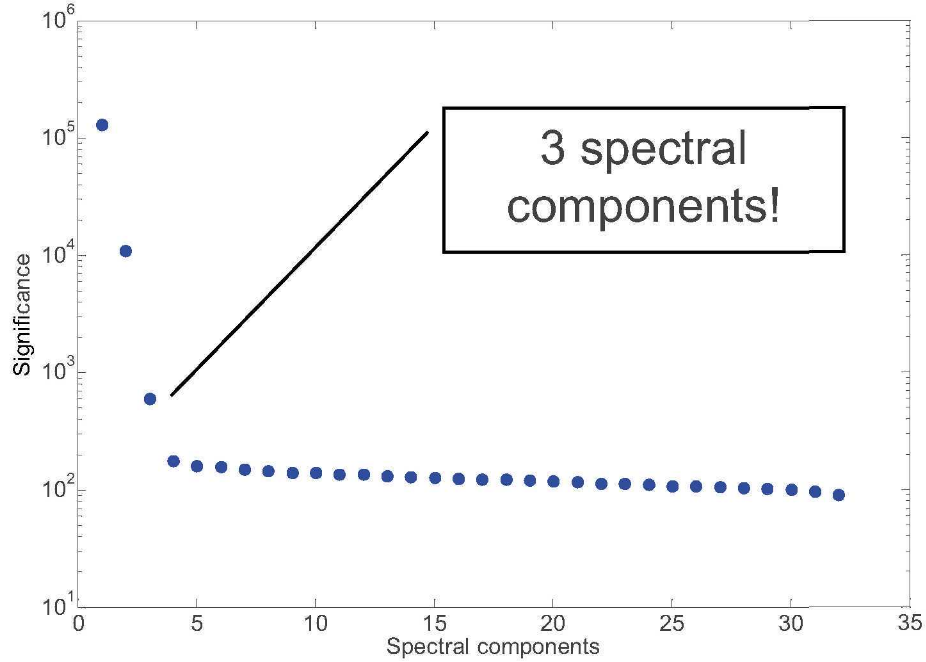




Bleaching

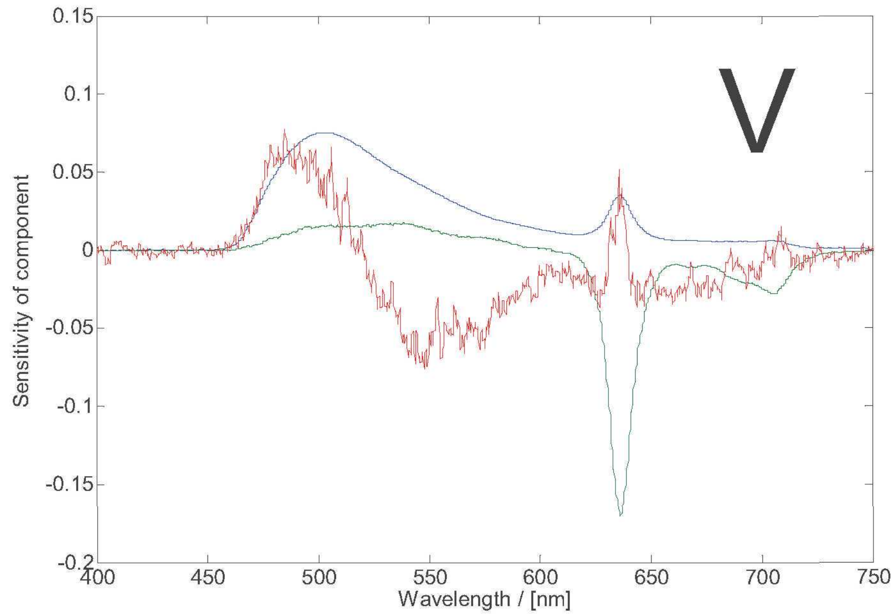
- 5mW bleaching of ALA in skin
- 80 time frames, 2048 wavelengths band
- SVD
- Eigen values: 3 molecules involved in the process
- What goes on?
- Can we predict the spectrum after 30s?





Bleaching

- PC not pure molecular spectra
- Complete summarizing
- Dynamic models can be fitted
- Dynamics can be studied for different intensities



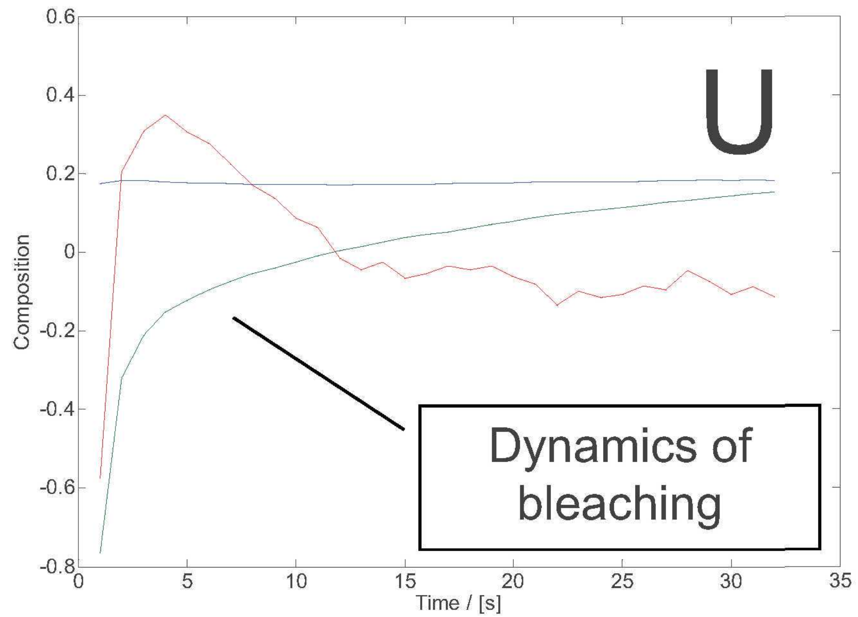
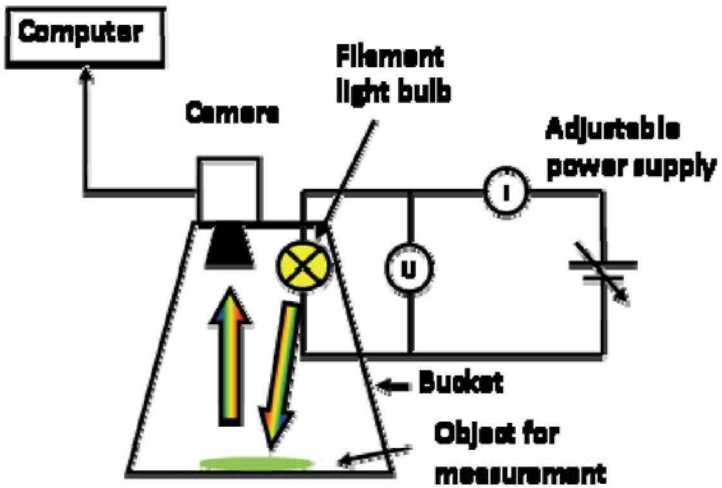
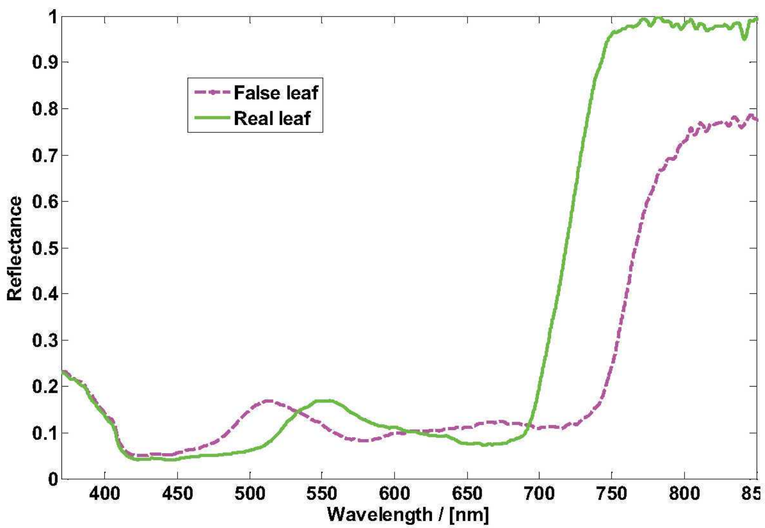


Image decomposition, Temp

17 temperatures, 10^6 pixels





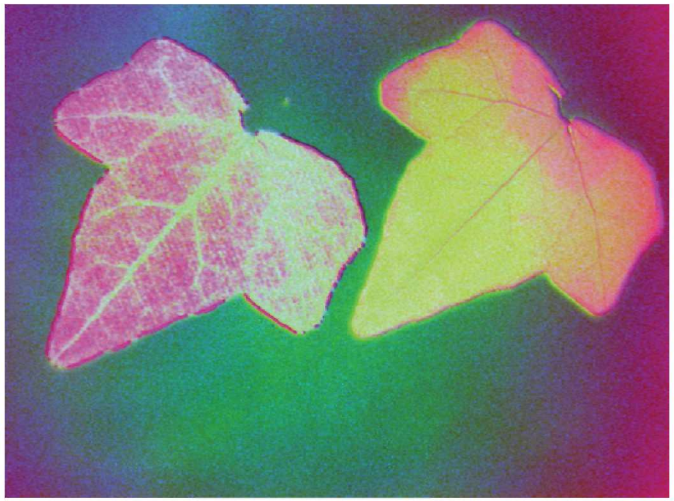
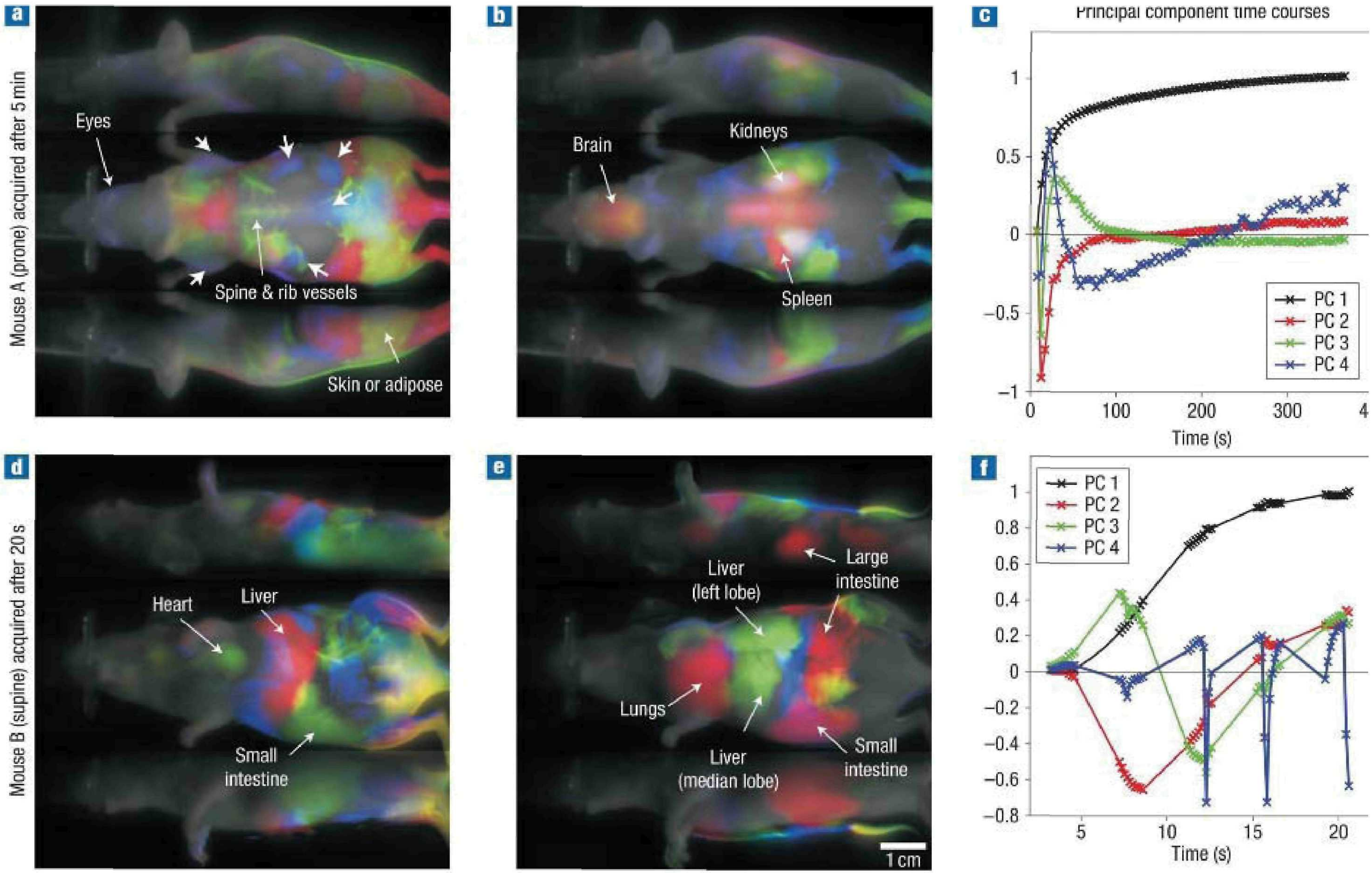


Image decomposition, Time



Nature, Hillman 2008

More than the sum of the parts

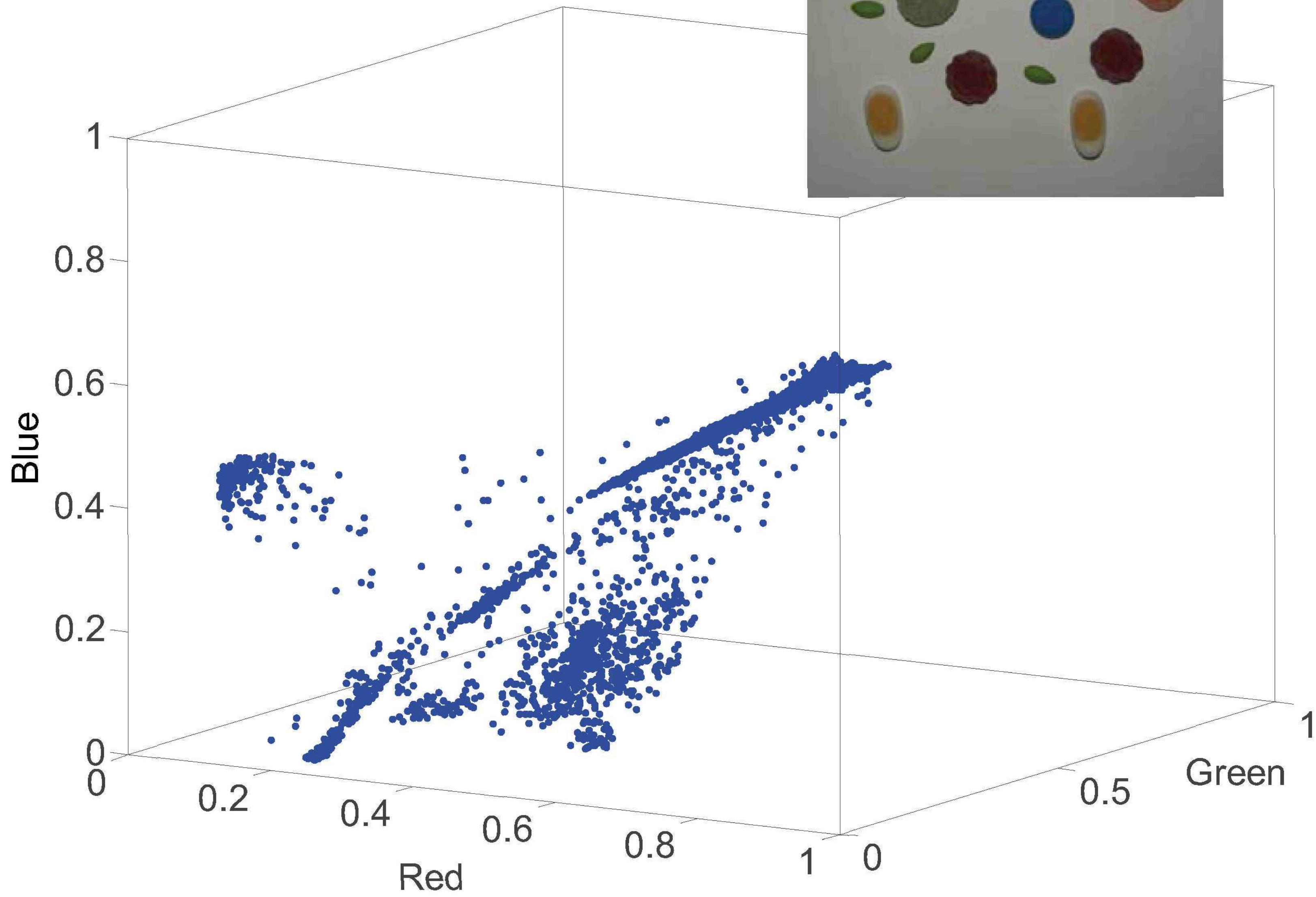
- In image and cytometry huge sample numbers N are acquired, and few parameters, K, are measured.
- E.g: 40Mpix RGB or 100.000 cells / s
- N>>K
- Decomposition of distributions
- Any difference from before?

Measured parameters, e.g. R_λ

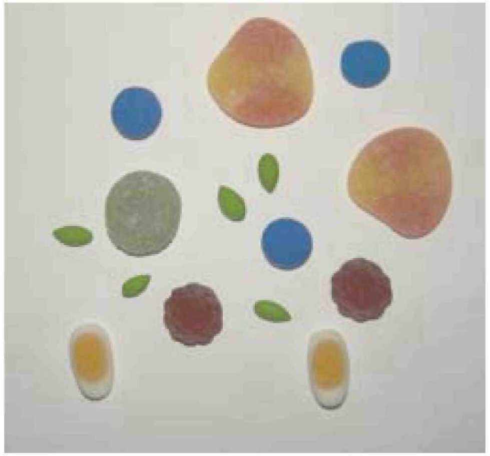
$$M = \begin{bmatrix} m_{11} & m_{12} & \dots & m_{1K} \\ m_{21} & m_{21} & \dots & m_{2K} \\ \dots & \dots & \dots & \dots \\ m_{N1} & m_{N2} & \dots & m_{NK} \end{bmatrix}$$

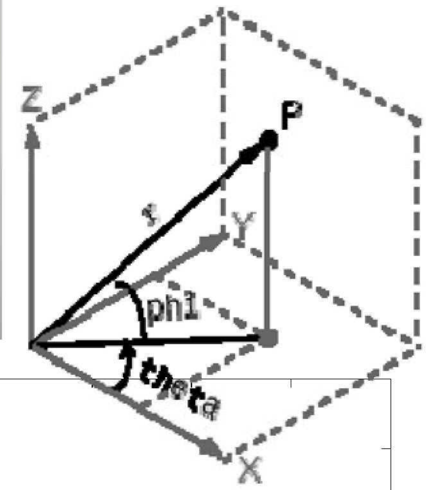
Measurements, individuals, pixels or voxels

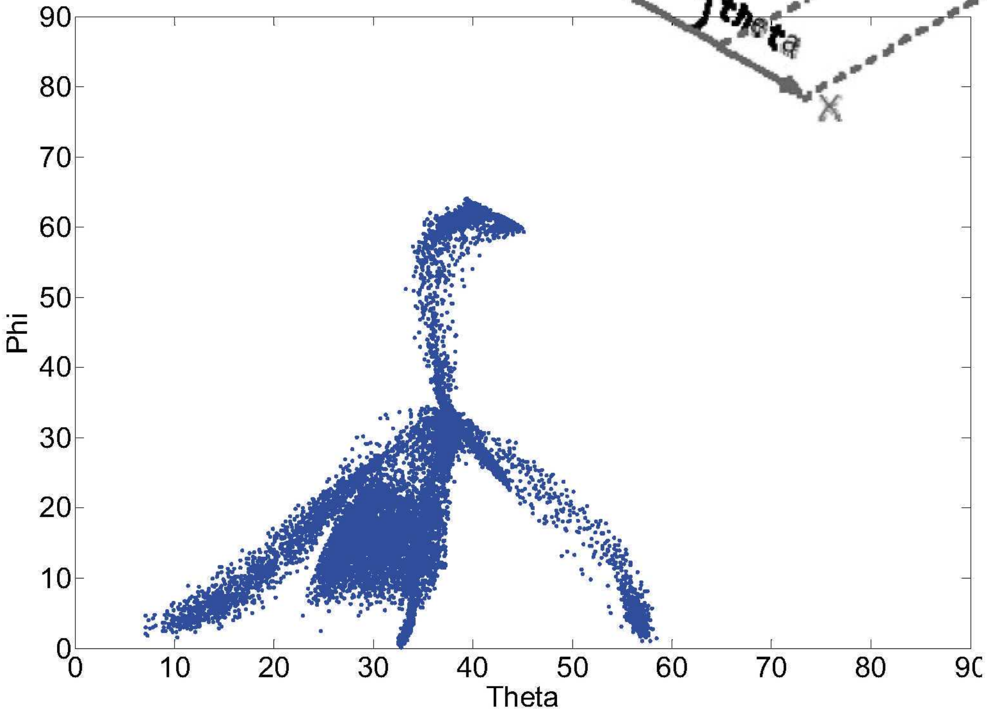
Color spaces and scatter plots



EEM and RGB decomposition



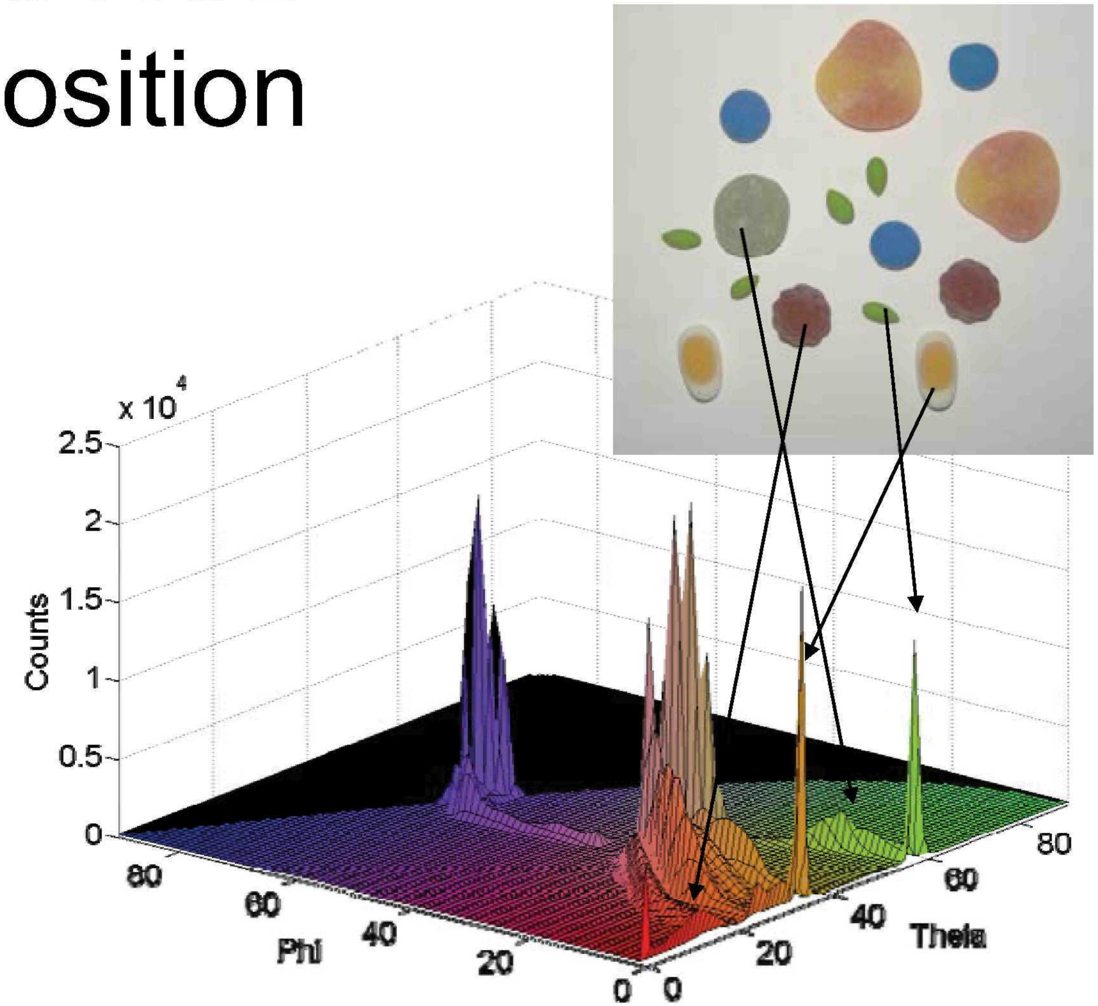




- Unit-less processing
- Concentration of observations?
- Multidimensional histograms

EEM and RGB decomposition

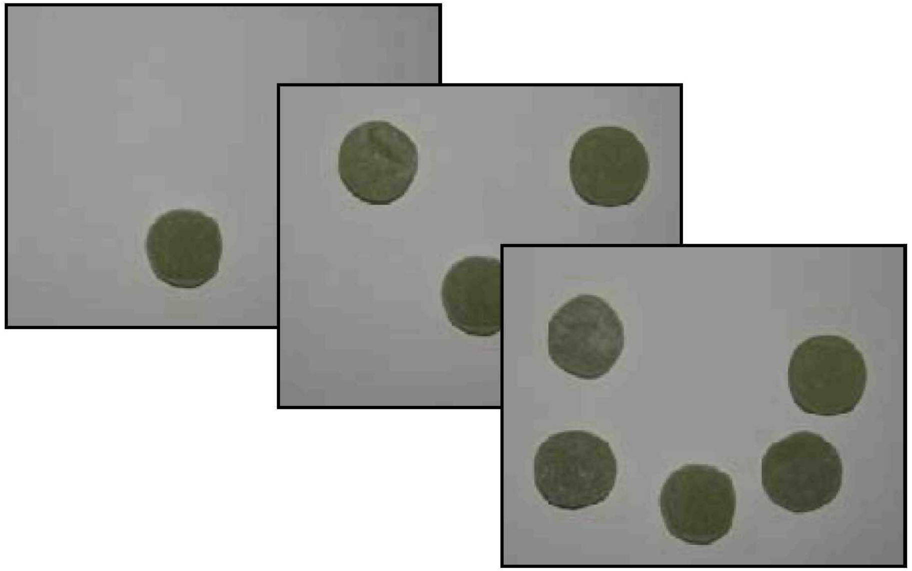
- Spectra, images or photon histograms?
- CCD, pixels and bins

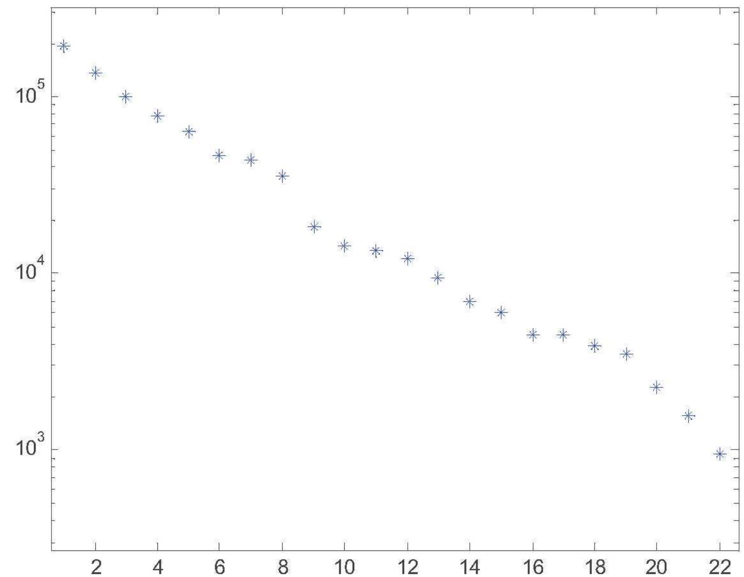


Standard addition

- 8 components detected

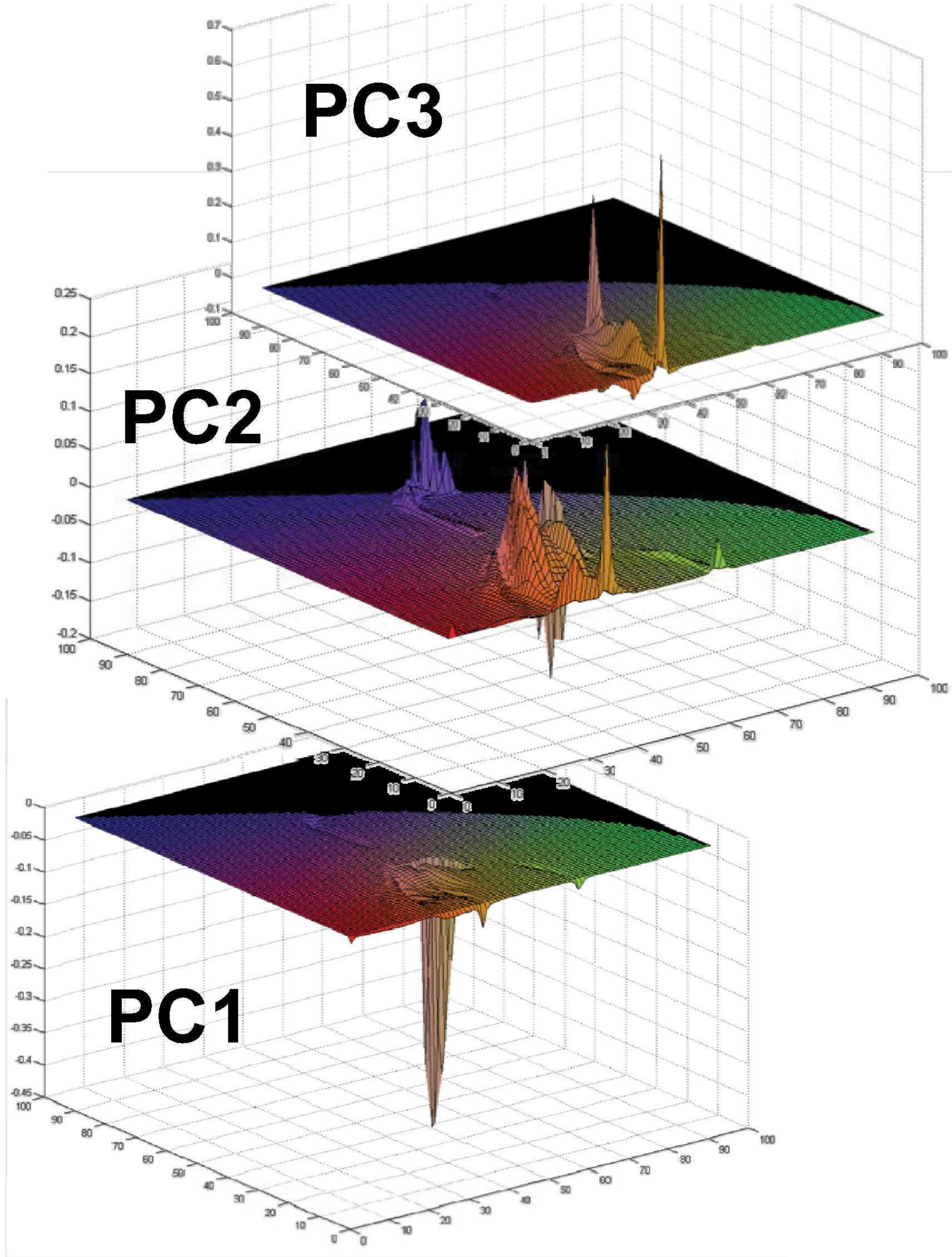
- 7 types of candy… + paper





2D principal components

- Orthogonal planes
- U reduced representation of M
- How many of each?



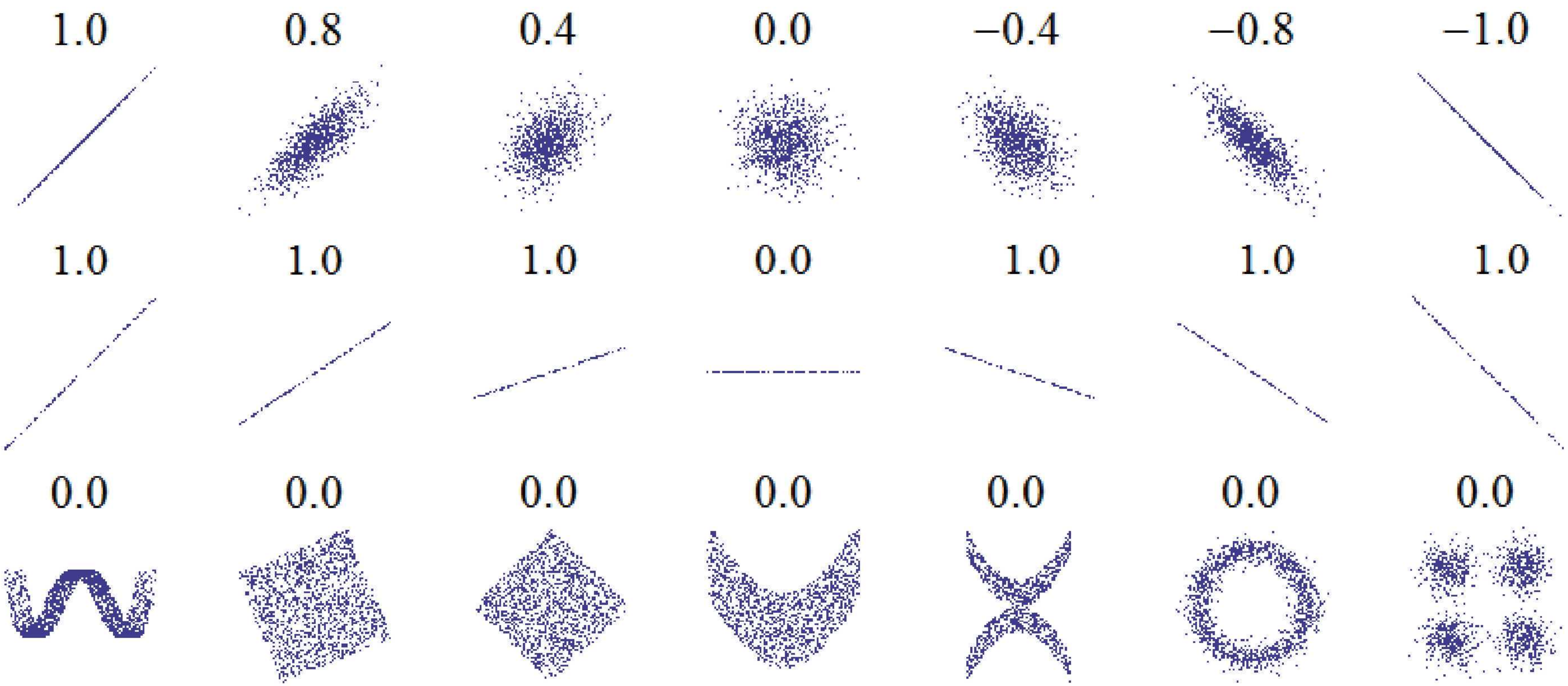
2. Modelling:
Can data answer my question?

Who is who?

How many candies?

Can we predict the spectrum after 30s?

2D correlations



…As far as we get in flat paper?

Contrast functions

- Is there a function F, so that **Y=F(M)** ?
- Better said: Is there a function F, so that **Y=F(U)?** Where U is a truncated representation of M.

Providing answers to the question:

Correct answer from a professional, e.g. IsMale

$$Y = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \dots \\ 0 \end{bmatrix}$$

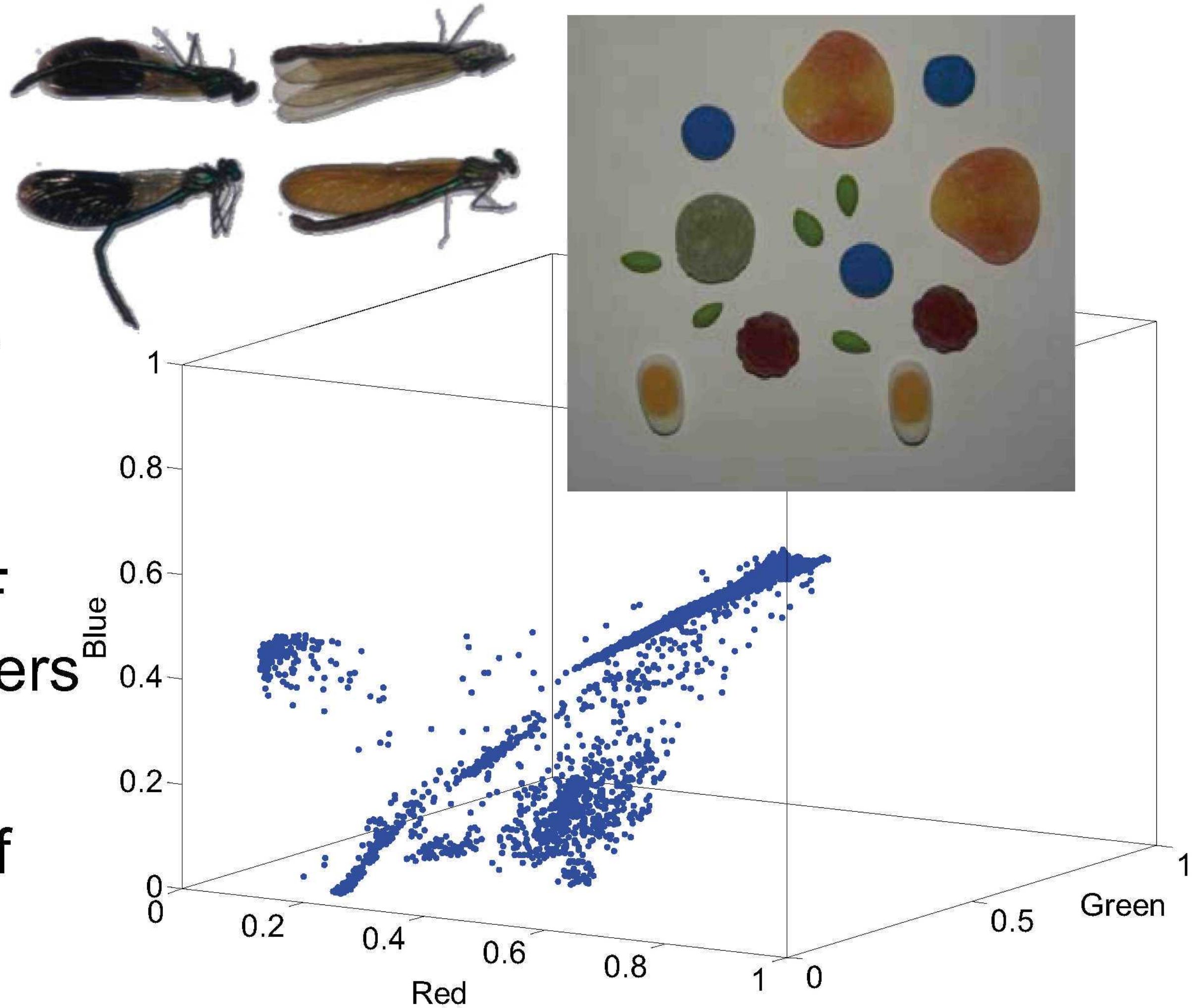
Reduced measured parameters, in terms of PCs

$$U = \begin{bmatrix} u_{11} & u_{12} & \dots & u_{1T} \\ u_{21} & u_{21} & \dots & u_{2T} \\ \dots & \dots & \dots & \dots \\ u_{N1} & u_{N2} & \dots & u_{NT} \end{bmatrix}$$

Measurements, individuals, pixels or voxels

Truncation

Decisive functions



- Y is a vector of true and false values

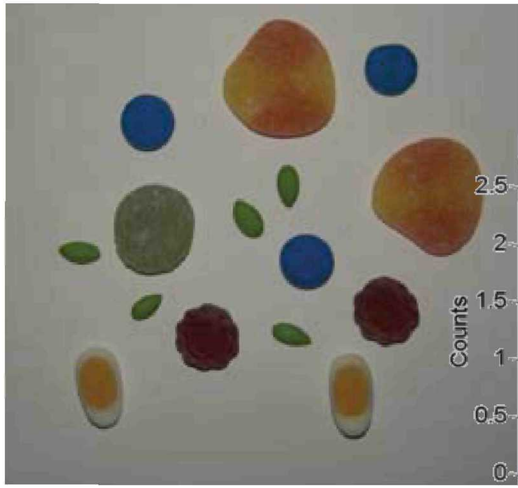
- Contour curves of F encircles data clusters

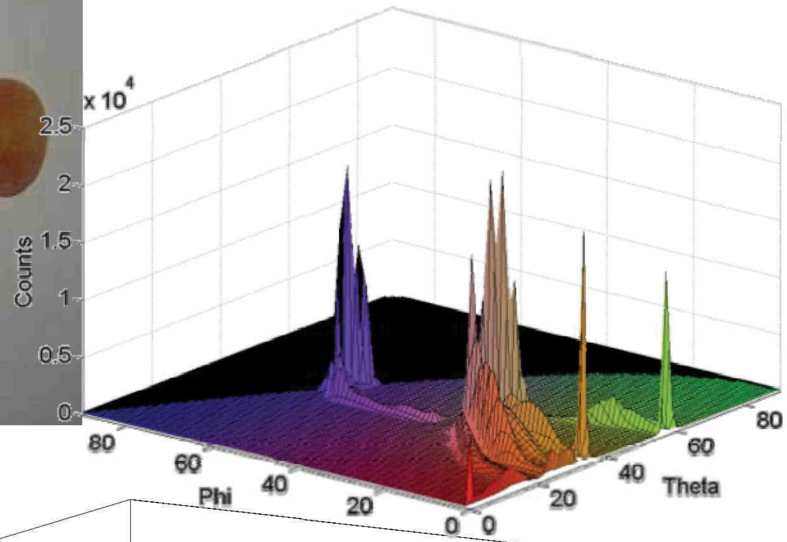
- Direct application of histograms

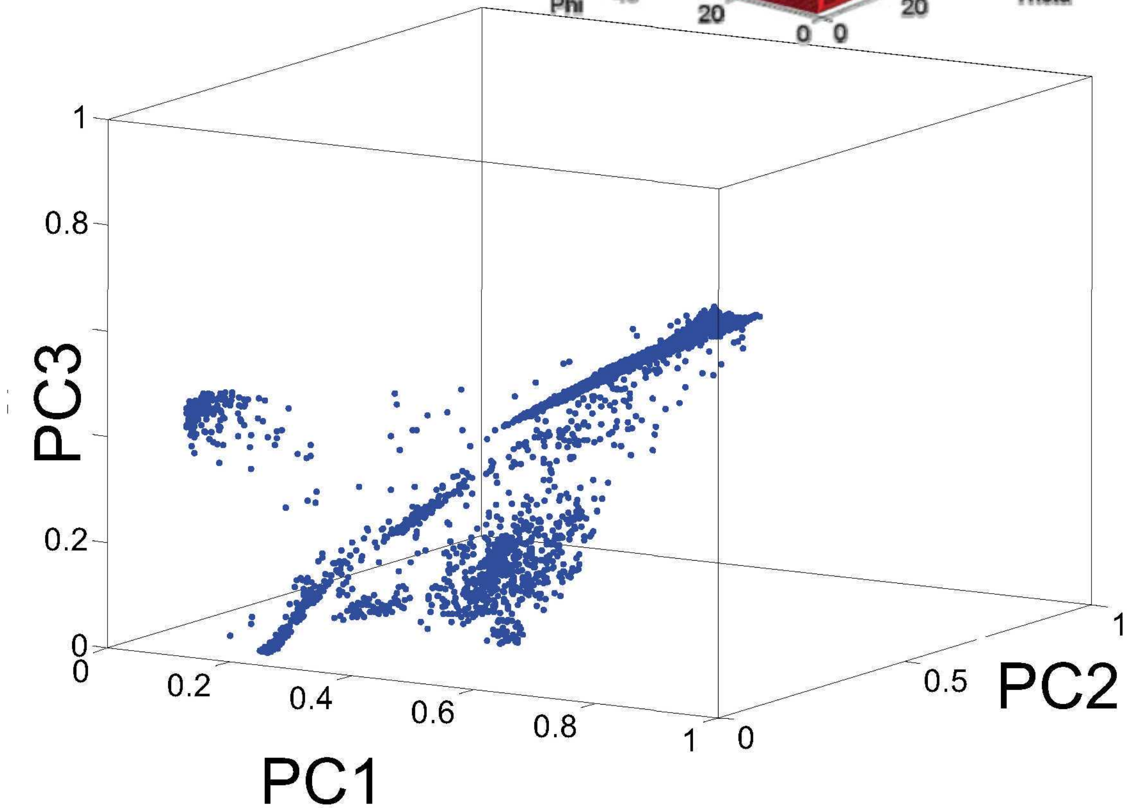
- Y=[0 0 0 1 0 1 1 0 …]’

Quantitative functions

- Y is a vector of analog values
- F maps data point to the correct answers
- Y=[2 eggs 0 eggs 5 eggs;
  5 seeds 2 seeds …]’







Models and matrix formulation

$$y = k_0 + k_1 * \sin(x) + k_2 * x^2$$

Provided answers

$$\begin{array}{ccc} 1 & \text{Sin}(x) & x^2 \end{array}$$

$$Y = \begin{bmatrix} 1 & 0.84 & 1 \\ 1 & 0.91 & 4 \\ 1 & 0.14 & 9 \end{bmatrix} * \begin{bmatrix} k_0 \\ k_1 \\ k_2 \end{bmatrix}$$

$$\Phi \qquad \theta$$

Unknown model coefficients

Reduced measured values

Matlab: θ=Φ\Y

Link functions

- Function space of Y
- Log, [0, ∞]
- Logit, [0, 1]
- E.g., Length, age, transmission, reflectance…

$$\log(Y) = \begin{bmatrix} 1 & 0.84 & 1 \\ 1 & 0.91 & 4 \\ 1 & 0.14 & 9 \end{bmatrix} * \begin{bmatrix} k_0 \\ k_1 \\ k_2 \end{bmatrix}$$

Link function

Matlab: θ=Φ\log(Y)

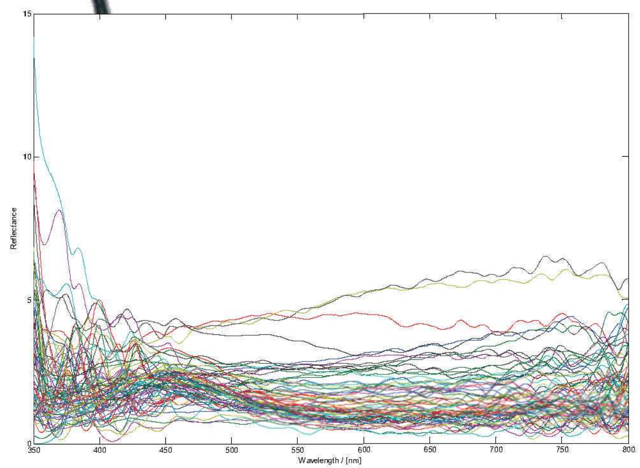
Fisher's discriminating function

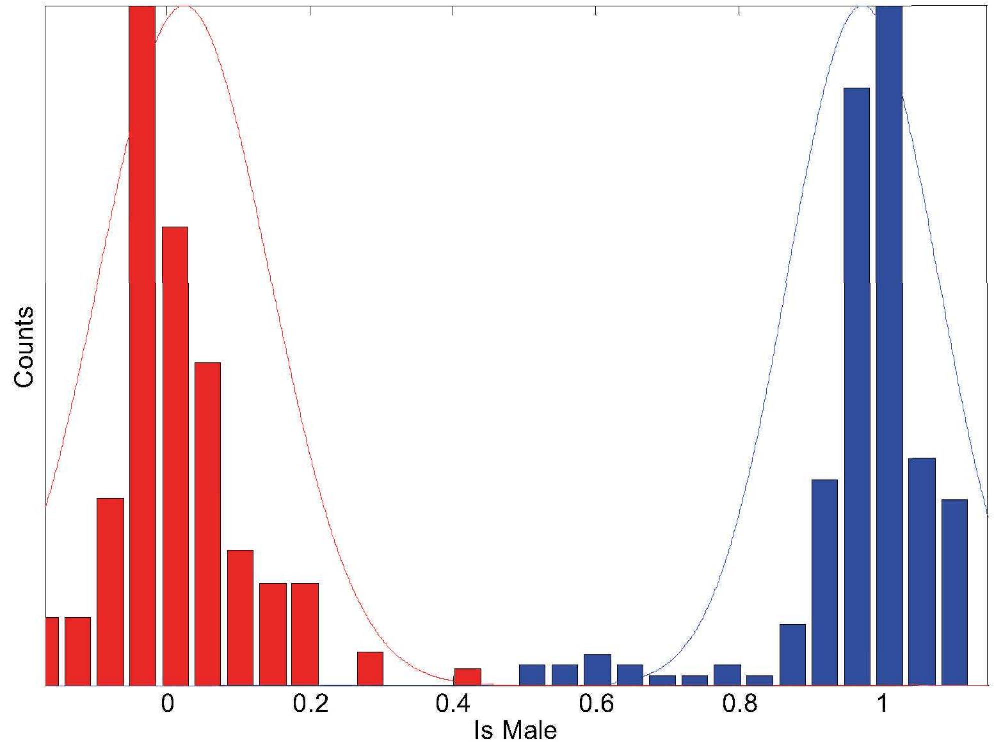
$y = k_0*1 + k_1*u_1 + k_2*u_2 + k_3*u_3 \ldots + k_T*u_T$

Offset

$$Y = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \cdots \\ 0 \end{bmatrix} \qquad Y = \begin{bmatrix} 1 & u_{11} & u_{1T} \\ 1 & u_{21} & u_{2T} \\ 1 & u_{N1} & u_{NT} \end{bmatrix} * \begin{bmatrix} k_0 \\ k_1 \\ k_T \end{bmatrix}$$

Who is who?





Females

Males

$$Q = \frac{|\mu_{F(A)} - \mu_{F(B)}|}{\mathrm{sqrt}(\delta_{F(A)} + \delta_{F(B)})}$$

F(R)

Y=k0+k1*u1+…k5*u5

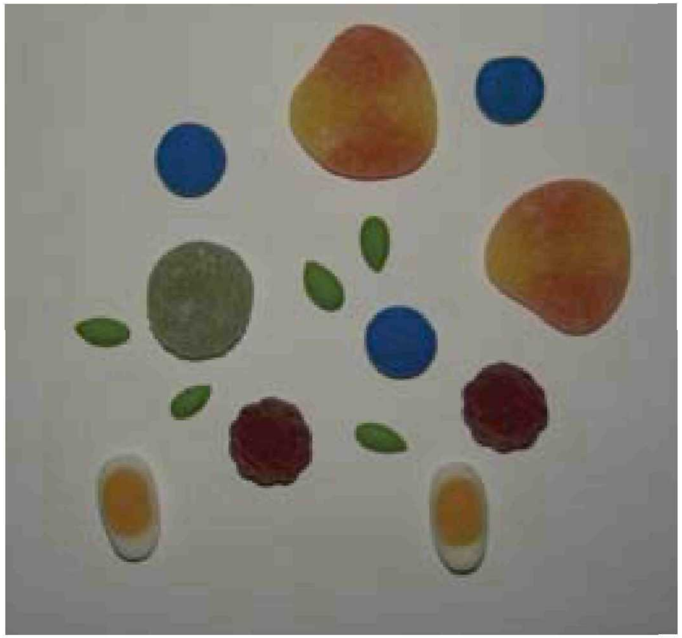
How many?

$Y=F(U)$

$Y=k_0+k_1*u_1+\ldots k_8*u_8$

Matlab: K=[ones U(:,1:8)]\Y

Residuals and quality



Type	Peaches	Green gum	Red gum	Blue tablet	Green seed	Eggs	Orange tablet
Reality	2	1	2	3	5	2	0
Estimated	1.9707	1.0408	1.8816	2.7244	5.3621	2.0872	0.0358

Polynomial non linear models

$$Y = F(\Phi) = \Phi\theta$$

- Convergence, Taylor theorem
- Inter-combinations, compare to 2D cosine transform

$$\begin{array}{c} pix_{1\ldots N} \\ \alpha^0 \\ \alpha^1 \\ \ldots \\ \alpha^m \end{array} \quad \begin{array}{c} \begin{array}{cccc} \beta^0 & \beta^1 & \ldots & \beta^m \end{array} \\ \begin{bmatrix} 1 & \beta & \ldots & \beta^m \\ \alpha & \beta\alpha & & \\ \ldots & & \ldots & \\ \alpha^m & & & \beta^m\alpha^m \end{bmatrix} \end{array}$$

$$\xrightarrow{\text{Re}\,shape}$$

$$\begin{array}{c} \\ pix_1 \\ pix_2 \\ \ldots \\ pix_N \end{array} \quad \begin{array}{c} \begin{array}{cccccc} \beta^0\alpha^0 & \beta^1\alpha^0 & \alpha^1\beta^0 & \beta^1\alpha^1 & \ldots & \beta^m\alpha^m \end{array} \\ \begin{bmatrix} 1 & 0 & 0.8 & 0 & \ldots & 0 \\ 1 & 0.9 & 0.1 & 0.09 & \ldots & 0.0081 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & 0.1 & 0.1 & 0.01 & \ldots & 0.00001 \end{bmatrix} \end{array} = \Phi$$

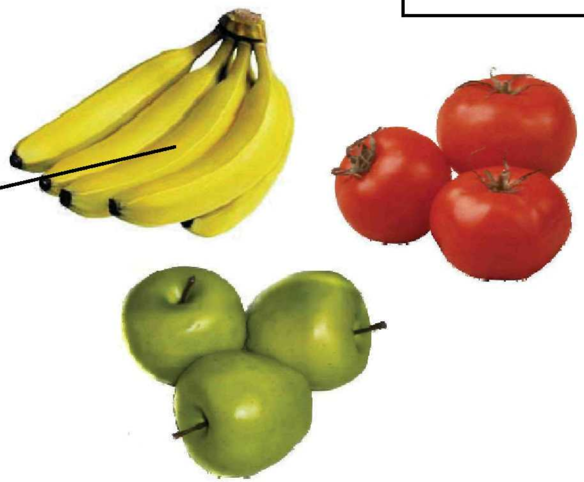
Fuzzy interpretations of polynomial functions

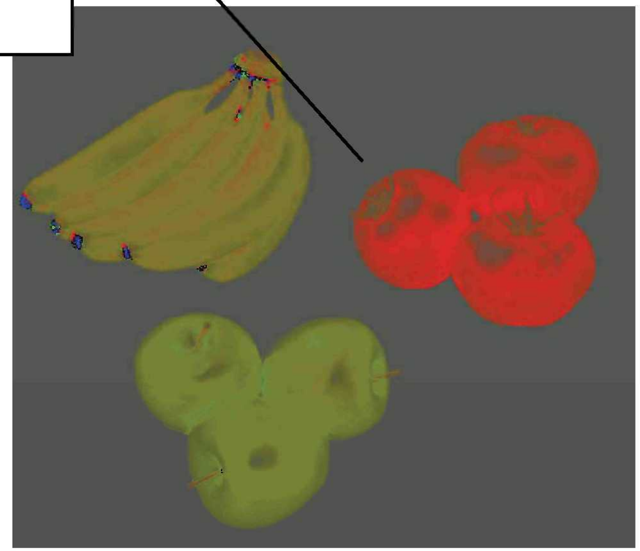
- Fuzzy logic:
  - A, B, C $\in$ [0…1]
  - C=A and B $\rightarrow$ C=AB
  - C=A or B $\rightarrow$ C=A+B-AB
  - C=not A $\rightarrow$ C=1-A
- Statistical analog
- Turtle eggs, depths and/or temperature
- Bananas, tomatoes and apples

Decisive banana contrast function:

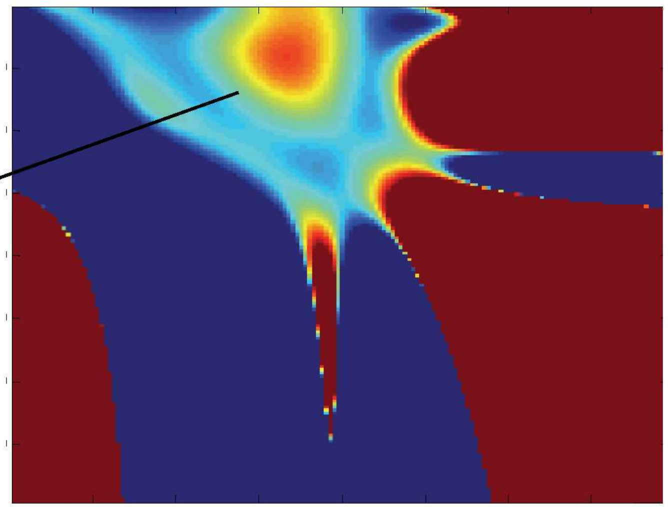
Unit-less processing

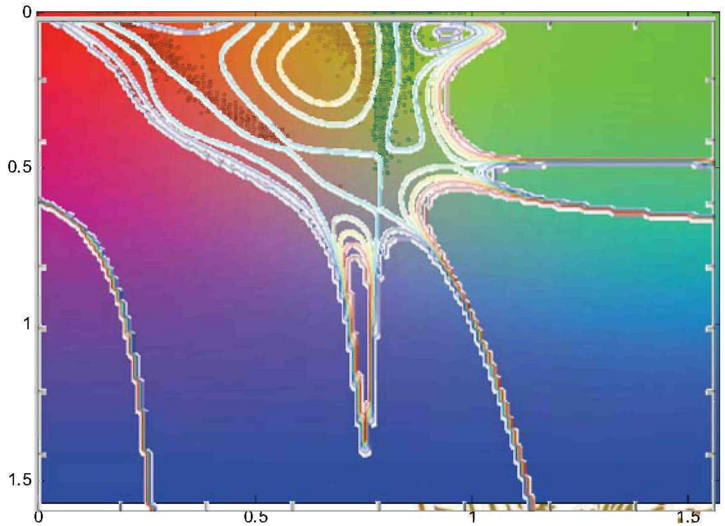
Red, green or red and green?



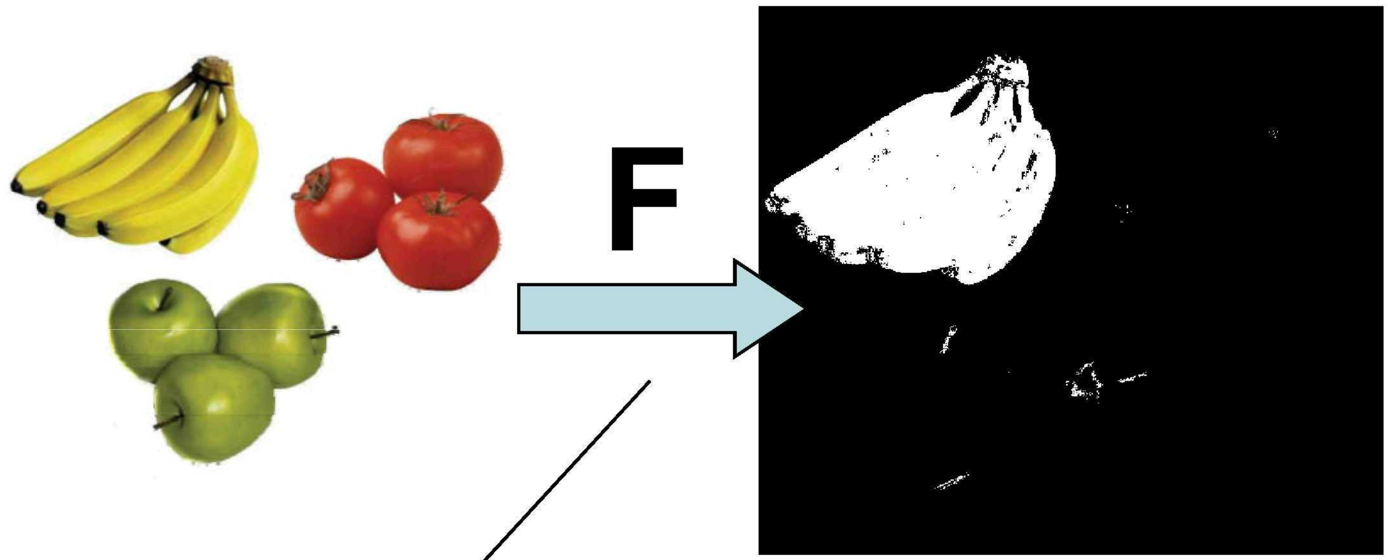


2D polynomial contrast function





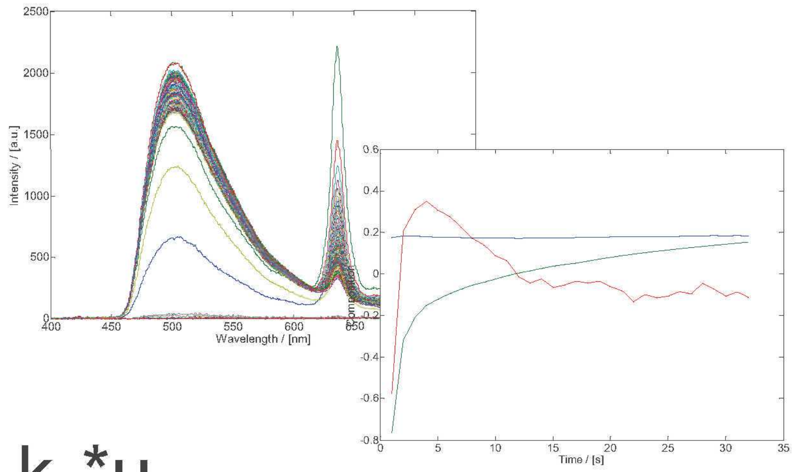
Evaluation



Nothing but a matrix multiplication!

Treshold -> True and false

Dynamical models



$$U_{n,1} = k_0 + k_1 * u_{n-1,1} + k_2 * u_{n-1,2} + k_2 * u_{n-1,3}$$

Time displaced

$$\begin{bmatrix} u_{1,1} & u_{1,2} & u_{1,3} \\ u_{2,1} & u_{2,2} & u_{2,3} \\ u_{N,1} & u_{0,1} & u_{0,1} \end{bmatrix} = \begin{bmatrix} 1 & u_{0,1} & u_{0,2} & u_{0,3} \\ 1 & u_{1,1} & u_{1,2} & u_{1,3} \\ 1 & u_{N-1,1} & u_{N-1,2} & u_{N-1,3} \end{bmatrix} * \begin{bmatrix} k_{1,1} & k_{1,2} & k_{1,3} & k_{1,4} \\ k_{2,2} & k_{2,2} & k_{1,3} & k_{2,4} \\ k_{3,2} & k_{3,2} & k_{3,3} & k_{3,4} \\ k_{4,2} & k_{4,2} & k_{4,3} & k_{4,4} \end{bmatrix}$$

First order system

Bleaching forecast based on initial values

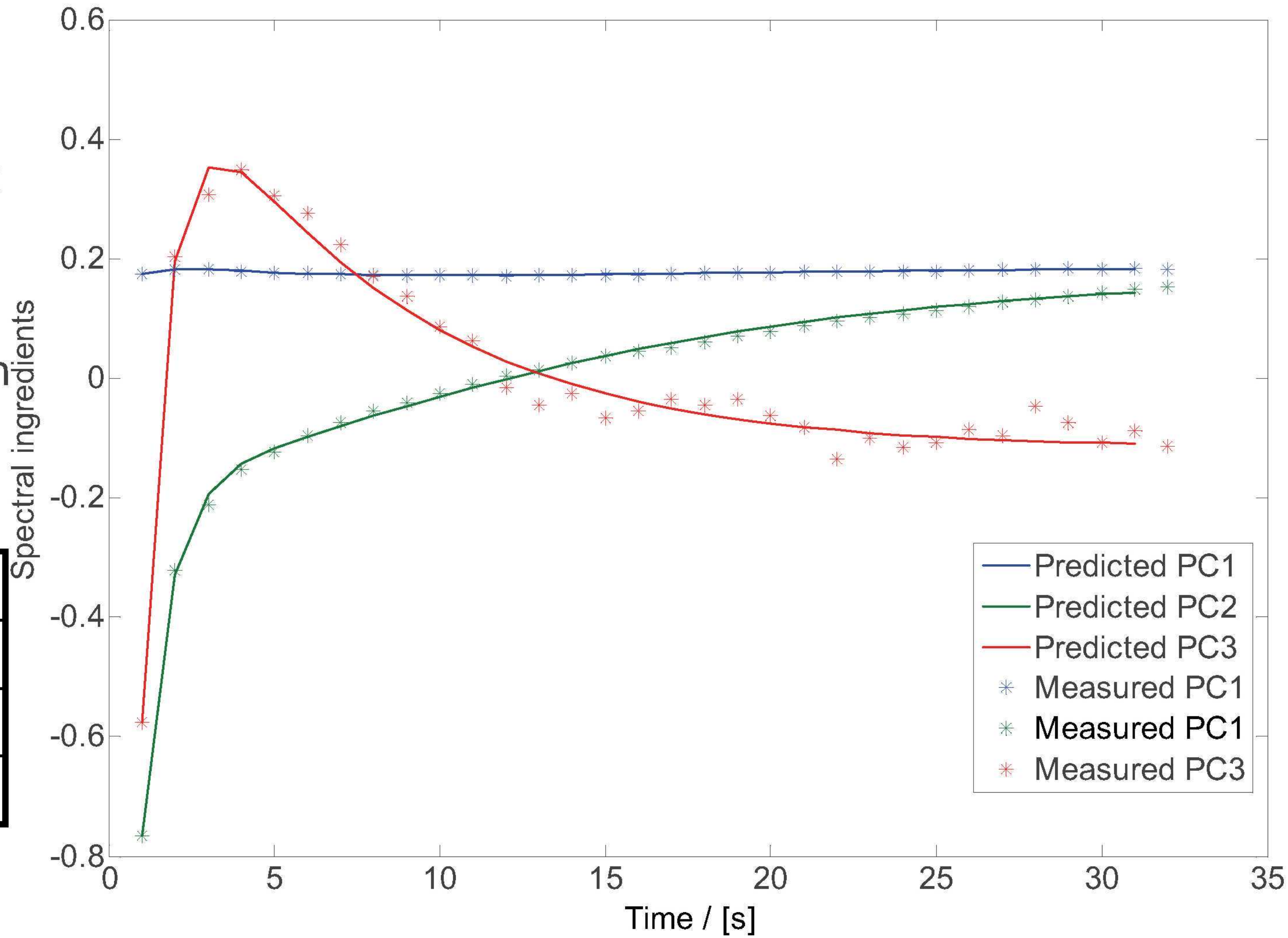
$M_{(t)}=U_{(t)}\ \Sigma\ V^*$

$U_{(t)}=[1\ U_0]^*K^t$

Complete description of dynamics

K=

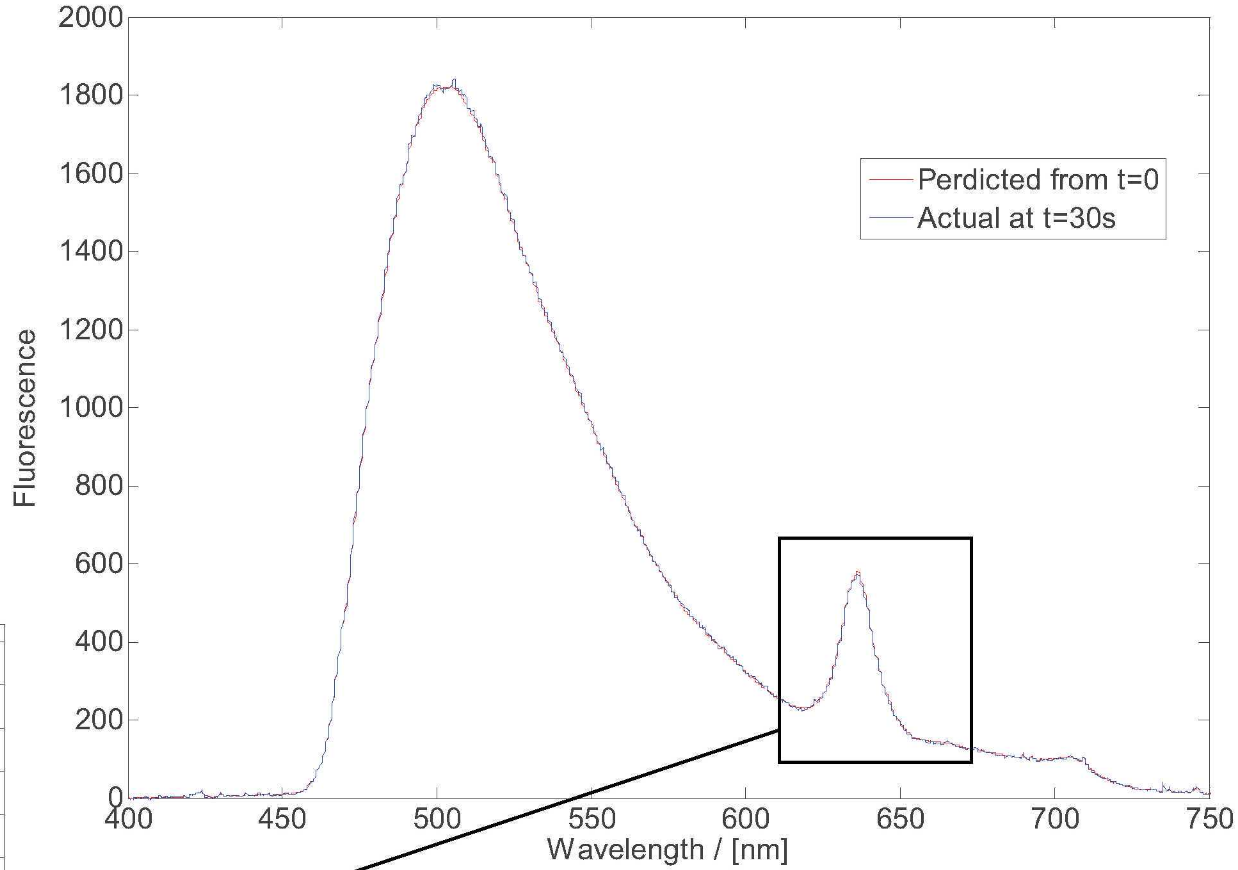
0.0125	-0.3214	-0.7518
0.9310	1.9789	4.3318
-0.0041	0.5985	-0.6455
-0.0095	-0.1882	0.5189

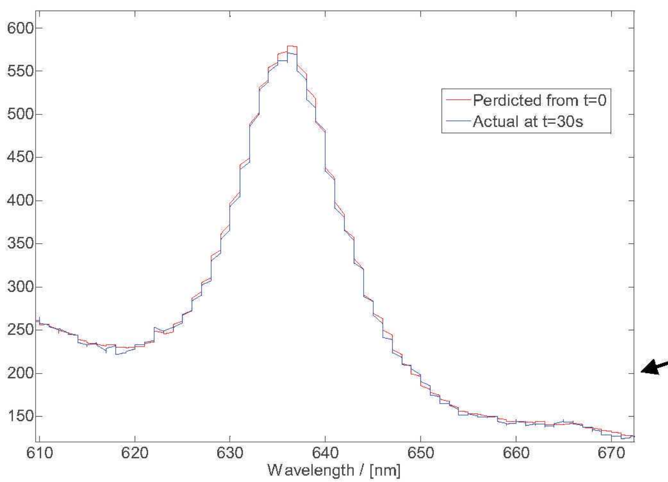


Comparison of actual and predicted spectrum

Time extrapolation

Reduced temporal representation





Have you been feeding

Applying Akaike's information criterion enables us to compare residuals with model parameters. Lowest AICc number provides the best choice, where **N** is the sample number and **m=1..N** the model order.

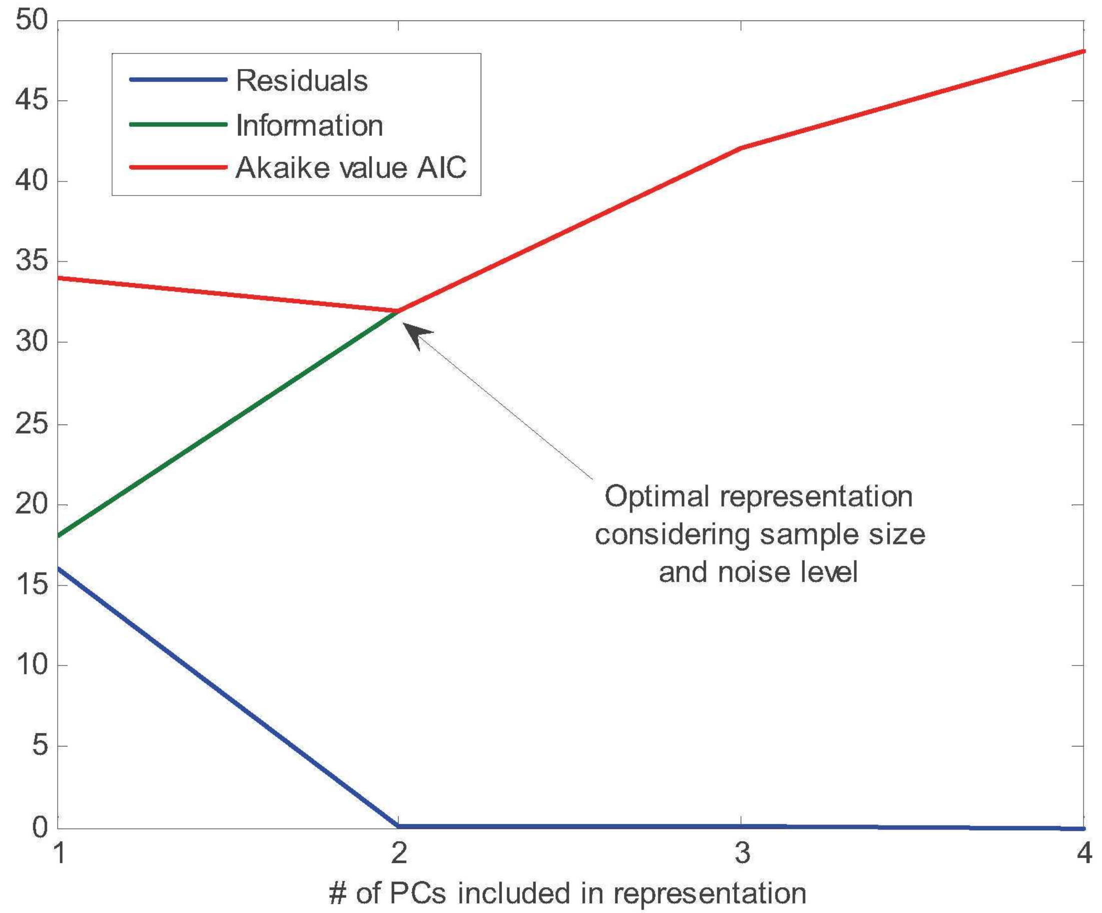
$$AIC(m) = -2N(N-m)Log\left(\frac{\prod_{i=m+1}^{N} \lambda_i^{\frac{1}{(N-k)}}}{\frac{1}{N-k}\sum_{i=m+1}^{N} \lambda_i}\right) + 2m(2N-m)$$

Penality for poor fitting

Penality for information added through parameters

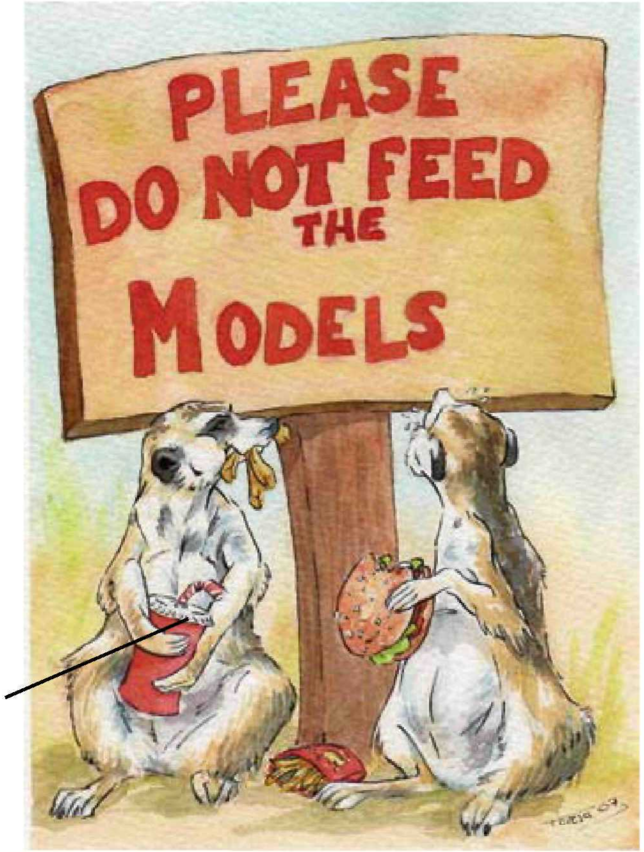
Akaike values:

By observing the residuals when representing spectra with different number of PC, we can estimate the number of PCs required to represent the data set by choosing the minimal AIC value.



Training and evaluation set

- Making sure not to feed the models with inside information



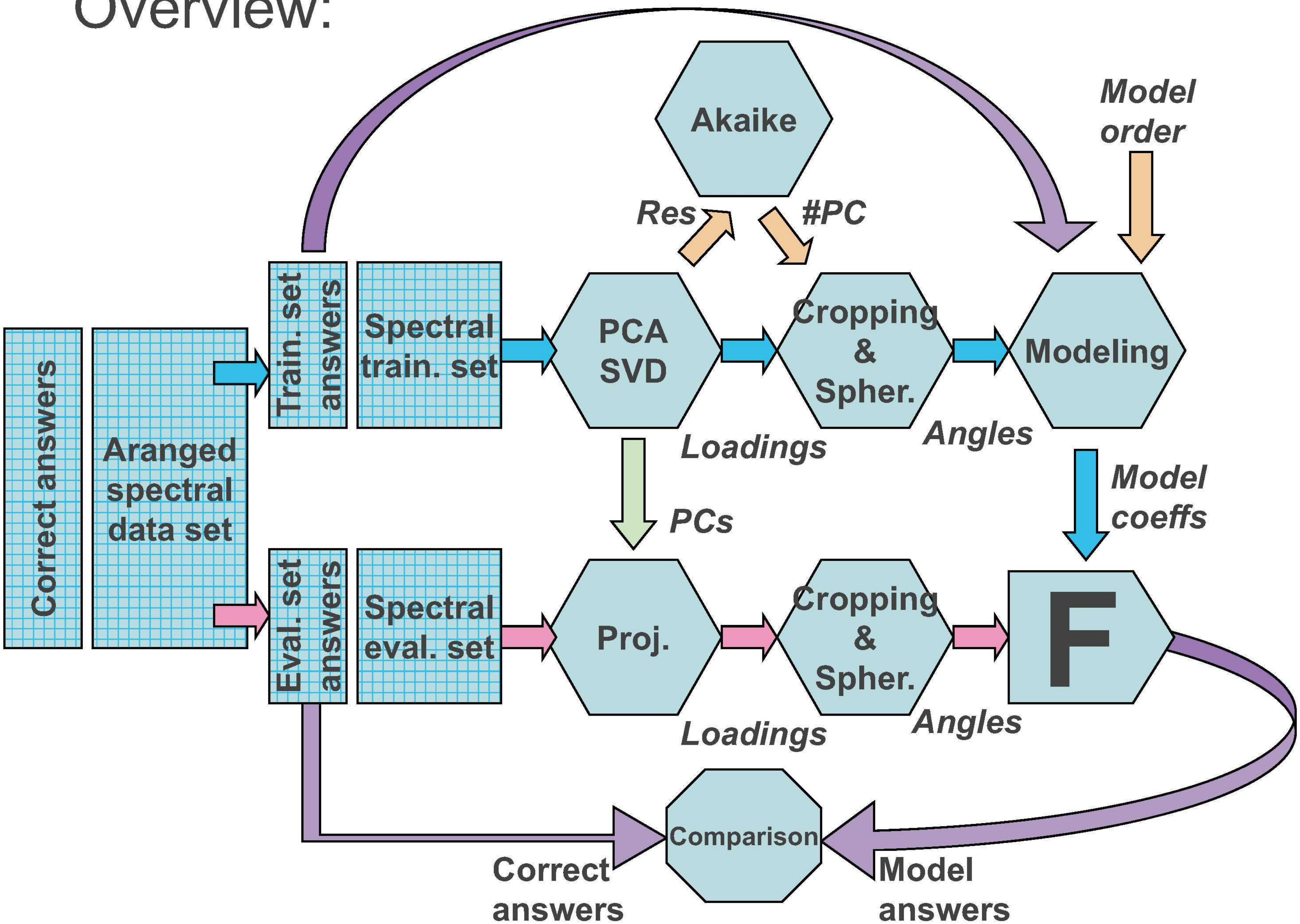
Model

$$Y = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \dots \\ 0 \end{bmatrix} \quad M = \begin{bmatrix} m_{11} & m_{12} & \dots & m_{1K} \\ m_{21} & m_{21} & \dots & m_{2K} \\ \dots & \dots & \dots & \dots \\ m_{N1} & m_{N2} & \dots & m_{NK} \end{bmatrix}$$

$$Y_{training} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ \dots \\ 1 \end{bmatrix} \quad M_{training} = \begin{bmatrix} m_{11} & m_{12} & \dots & m_{1K} \\ m_{31} & m_{31} & \dots & m_{3K} \\ \dots & \dots & \dots & \dots \\ m_{N1} & m_{N2} & \dots & m_{NK} \end{bmatrix}$$

$$Y_{evaluation} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ \dots \\ 0 \end{bmatrix} \quad M_{evaluation} = \begin{bmatrix} m_{21} & m_{22} & \dots & m_{2K} \\ m_{41} & m_{41} & \dots & m_{4K} \\ \dots & \dots & \dots & \dots \\ m_{N1} & m_{N2} & \dots & m_{NK} \end{bmatrix}$$

Overview:



Correct answers
Aranged spectral data set
Train. set answers
Spectral train. set
PCA SVD
Akaike
Res
#PC
Cropping & Spher.
Modeling
Model order
Loadings
Angles
PCs
Model coeffs
Eval. set answers
Spectral eval. set
Proj.
Cropping & Spher.
F
Loadings
Angles
Comparison
Correct answers
Model answers