STATE OF THE ART HYPERSPECTRAL DATA ANALYSIS AND PROSPECTS WITH RESPECT TO AI

ROSUS 2019, 21/03/19 Maribor Dr. Matthias Kerschhaggl

1. Presentation EVK
2. State of the art HSI data analysis
3. AI Prospects
4. Industrial apps using chemical imaging
5. Conclusions

EXPERT COMPANY

Industrial imaging solutions for sorting and inspection

CORE EXPERTISE

Hyperspectral, color and conductivity imaging technologies

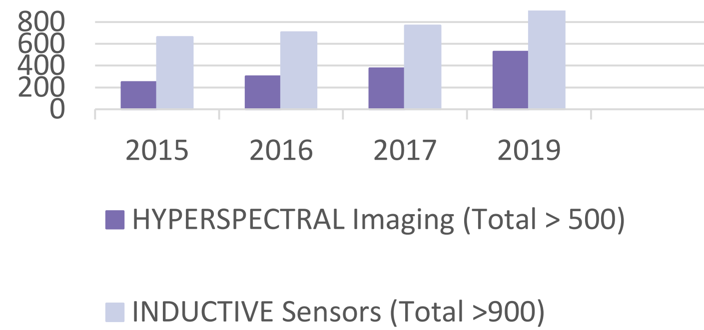
COMPLETE SOLUTIONS

From raw data acquisition to decision making

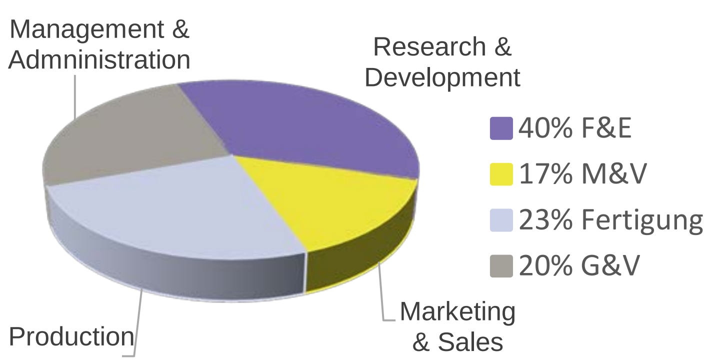
TARGET MEKETS

DEPLOYED PRODUCTS

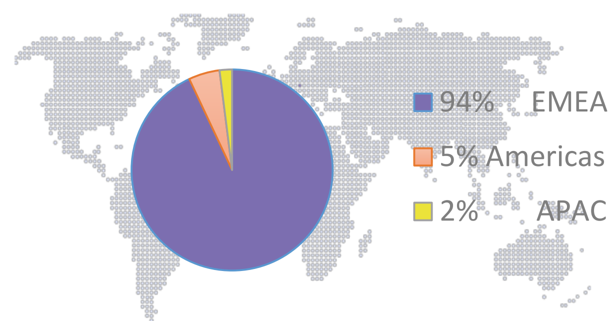
(kumuliert)



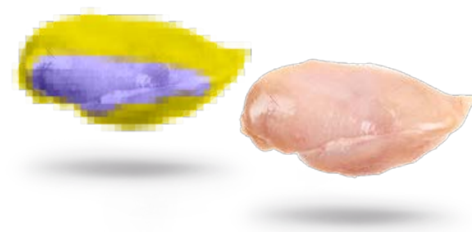
EVK TEAM



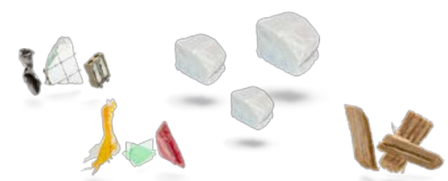
TURNOVER / REGION



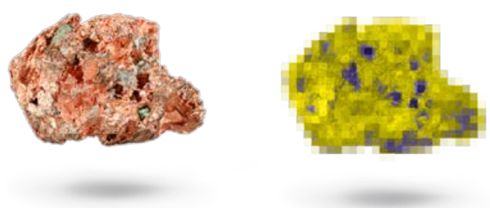
PROCESS ANALYTICS

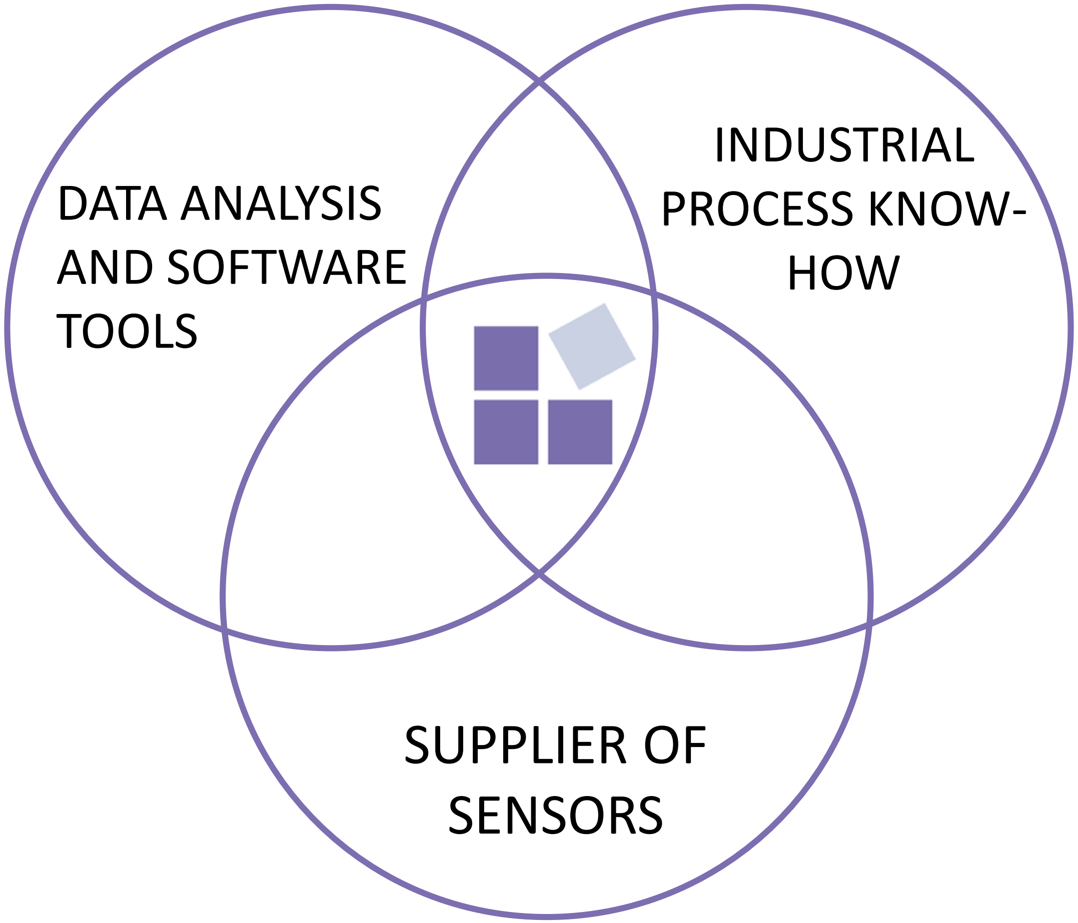



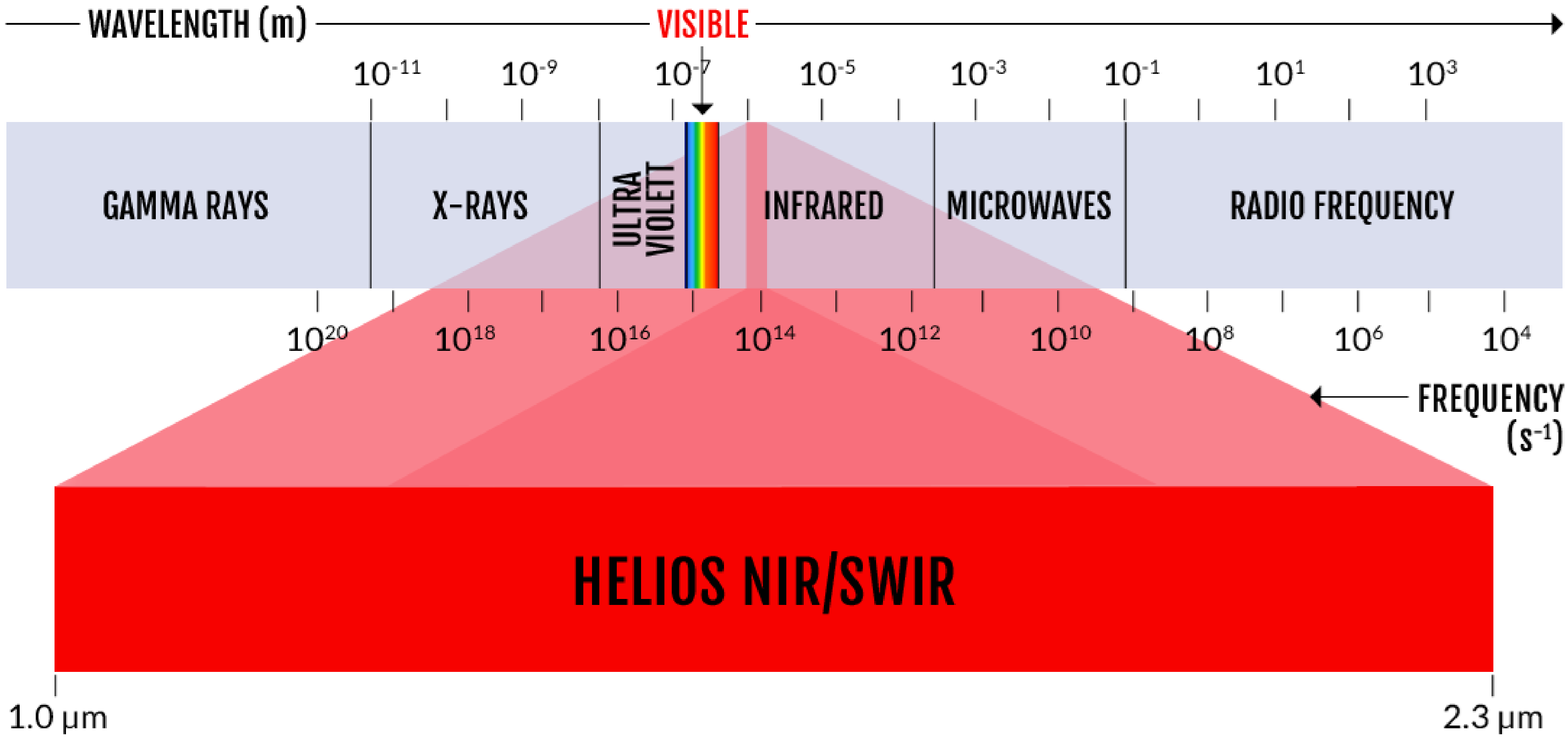
SMART SORTING

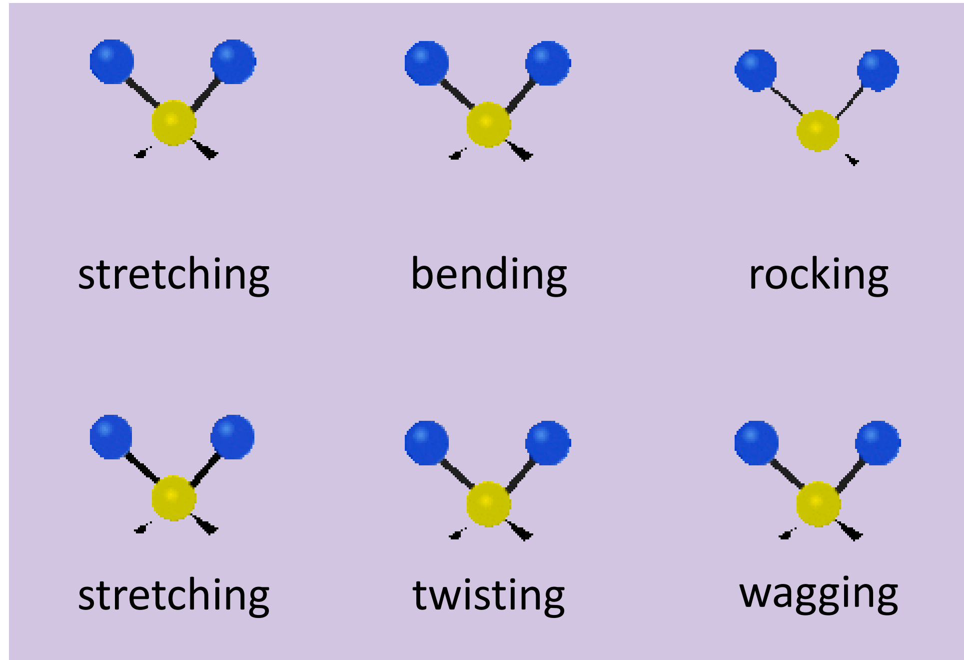



MONITORING AND INSPECTION





DATA ANALYSIS AND SOFTWARE TOOLS
INDUSTRIAL PROCESS KNOW-HOW
SUPPLIER OF SENSORS

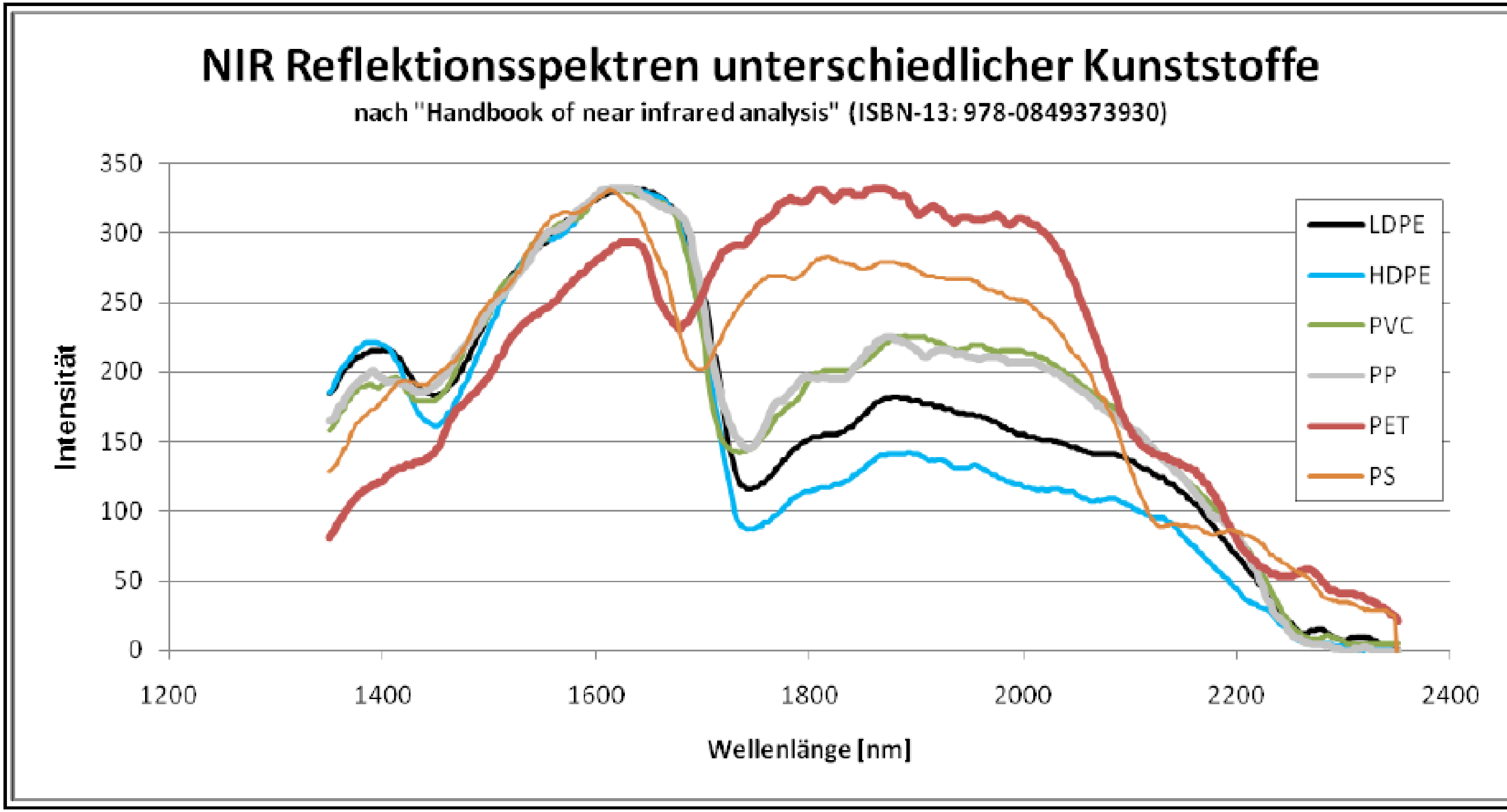

WAVELENGTH (m)
VISIBLE
10^{-11}
10^{-9}
10^{-7}
10^{-5}
10^{-3}
10^{-1}
10^{1}
10^{3}
GAMMA RAYS
X-RAYS
ULTRA VIOLETT
INFRARED
MICROWAVES
RADIO FREQUENCY
10^{20}
10^{18}
10^{16}
10^{14}
10^{12}
10^{10}
10^{8}
10^{6}
10^{4}
FREQUENCY (s^{-1})
HELIOS NIR/SWIR
1.0 µm
2.3 µm

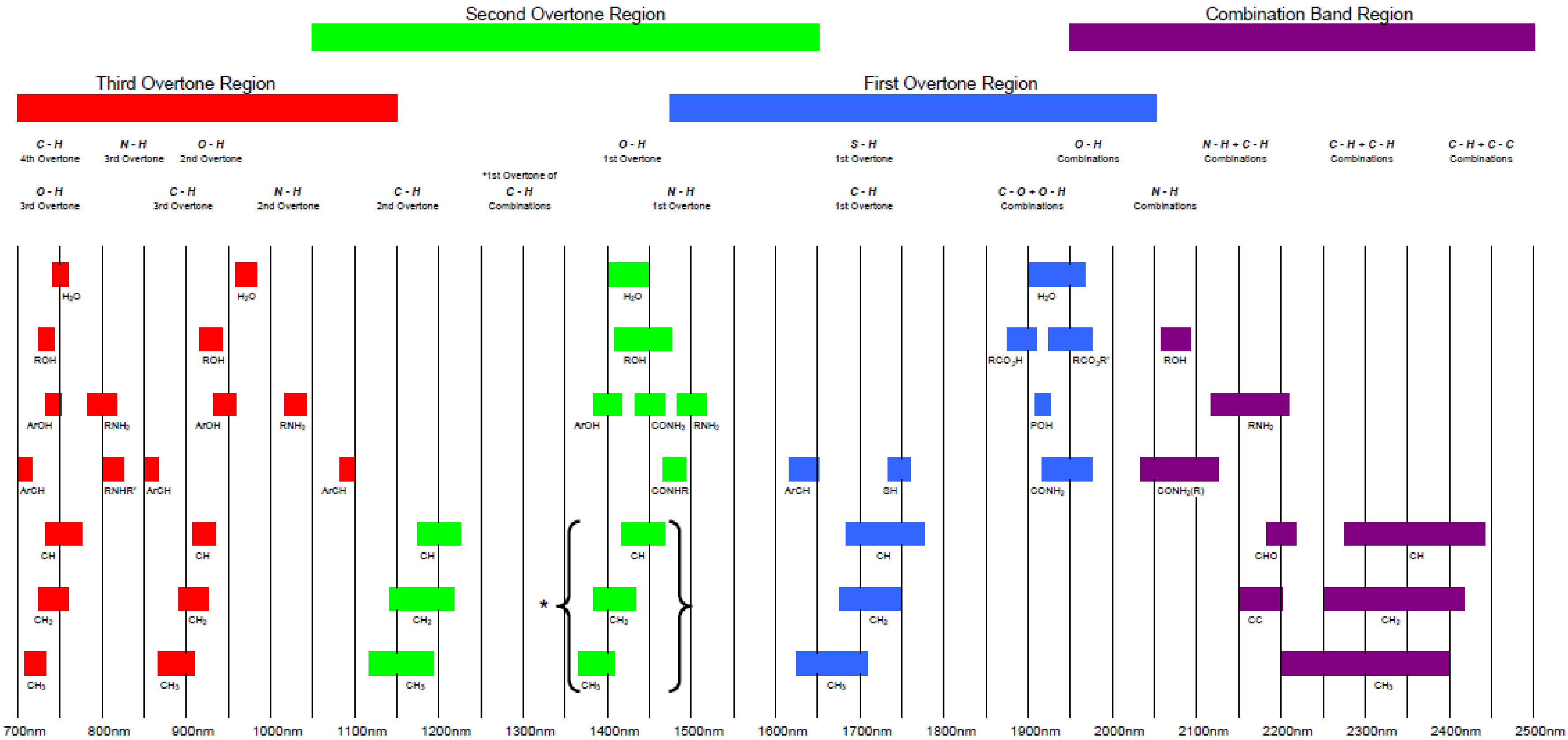


picture: Tiago Becerra Paolini

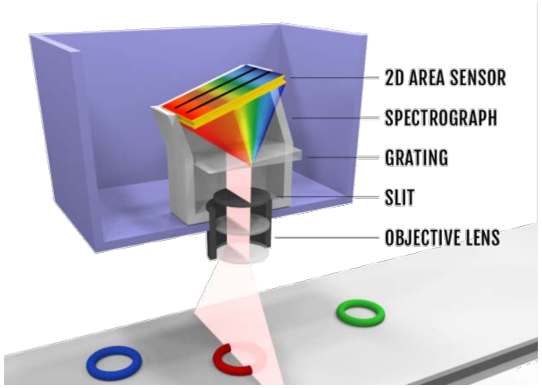
C-O-H-N molecular combinations

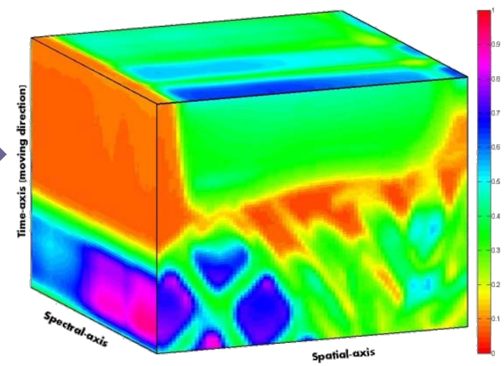
Absorption spectra according to the molecular structure

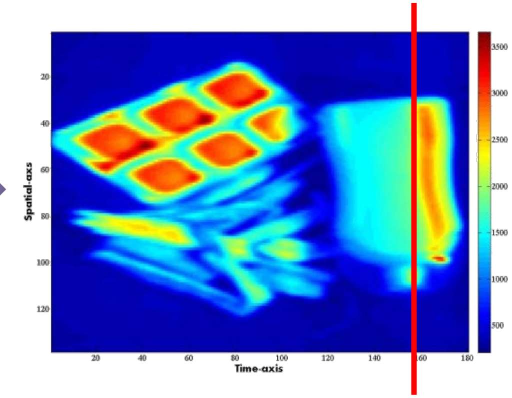

NIR Reflektionsspektren unterschiedlicher Kunststoffe
nach "Handbook of near infrared analysis" (ISBN-13: 978-0849373930)
Intensität
350
300
250
200
150
100
50
0
1200
1400
1600
1800
2000
2200
2400
Wellenlänge [nm]
LDPE
HDPE
PVC
PP
PET
PS



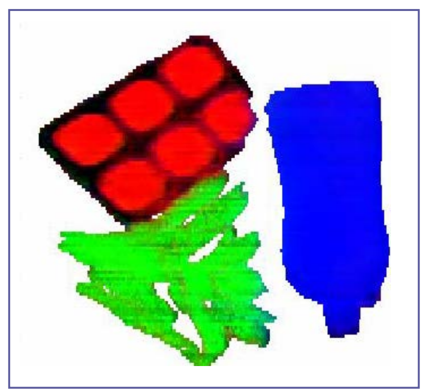
Seeing Spectroscopy Clearly ™

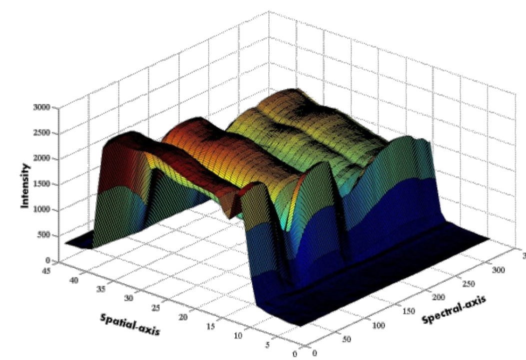






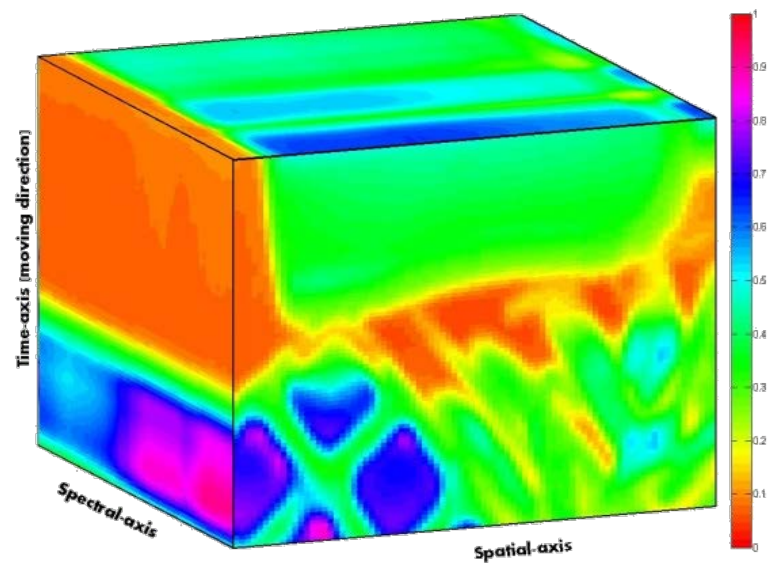
Camera internal classification and RGB visualization of chemical differences for real-time analysis





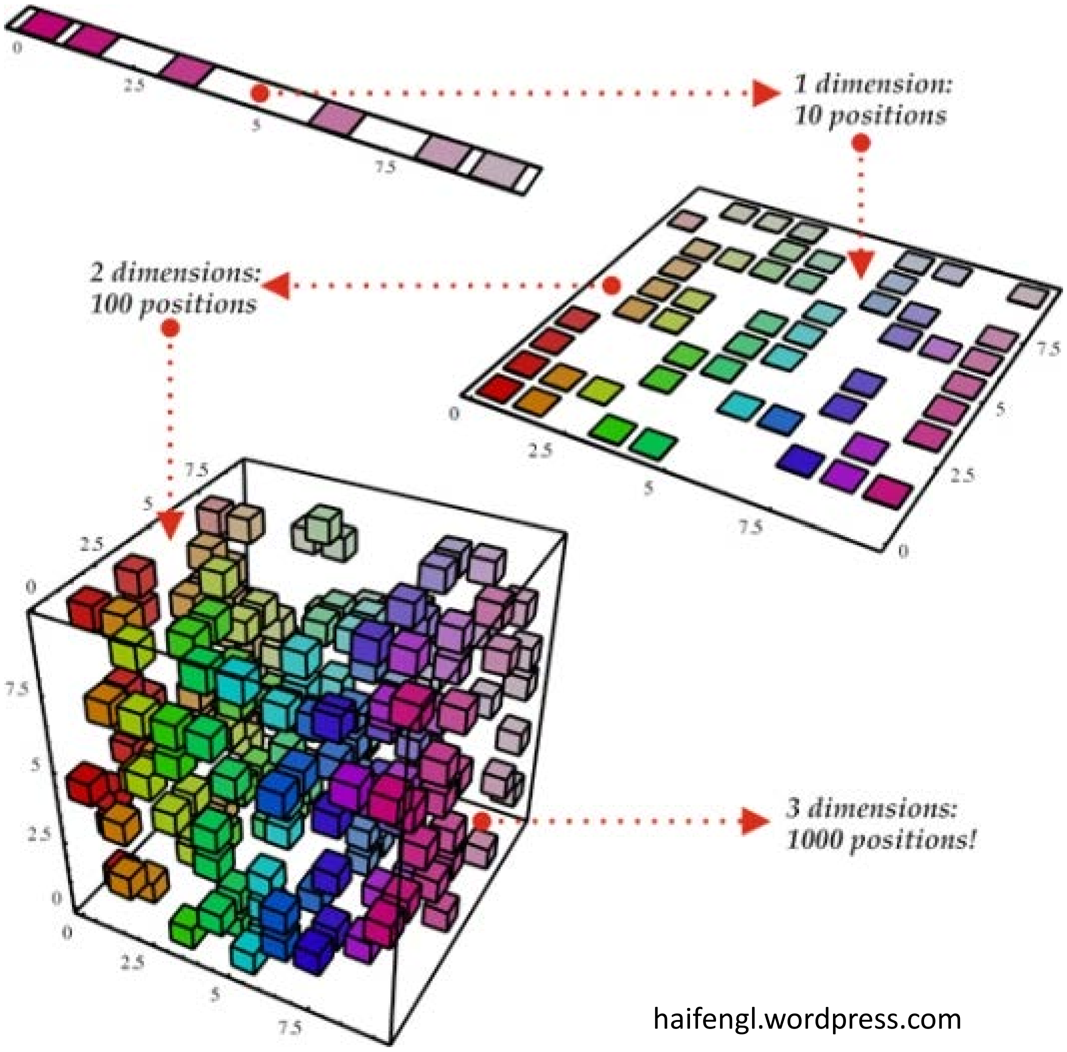
Hyperspectral data and its challenges in industrial applications

- High dimensionality (“curse of dimensionality”)
- Limited training samples
- Mixed spectral signatures
- Pre-Processing
- Real-time and inline suitable classification algorithms



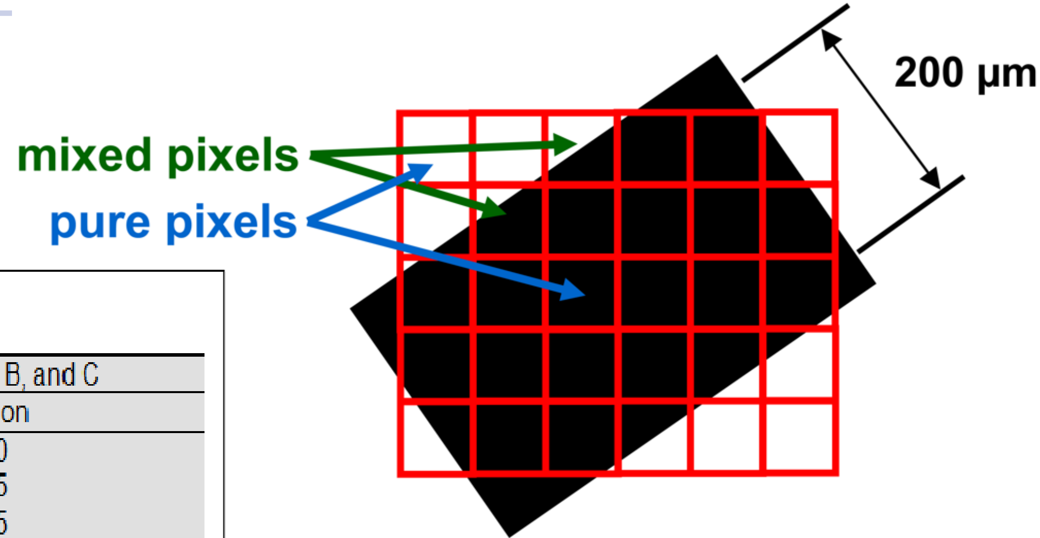
Curse of dimensionality

- Data coverage decreases with dimensionality (“empty space”)
- Data distributed mainly on the outskirts of data space
- Differences in min and max distances of data sets become marginal, i.e. distance measures lose meaning (e.g. knn search)
- Subspace methods (projections) can help!
- Manifold hypothesis

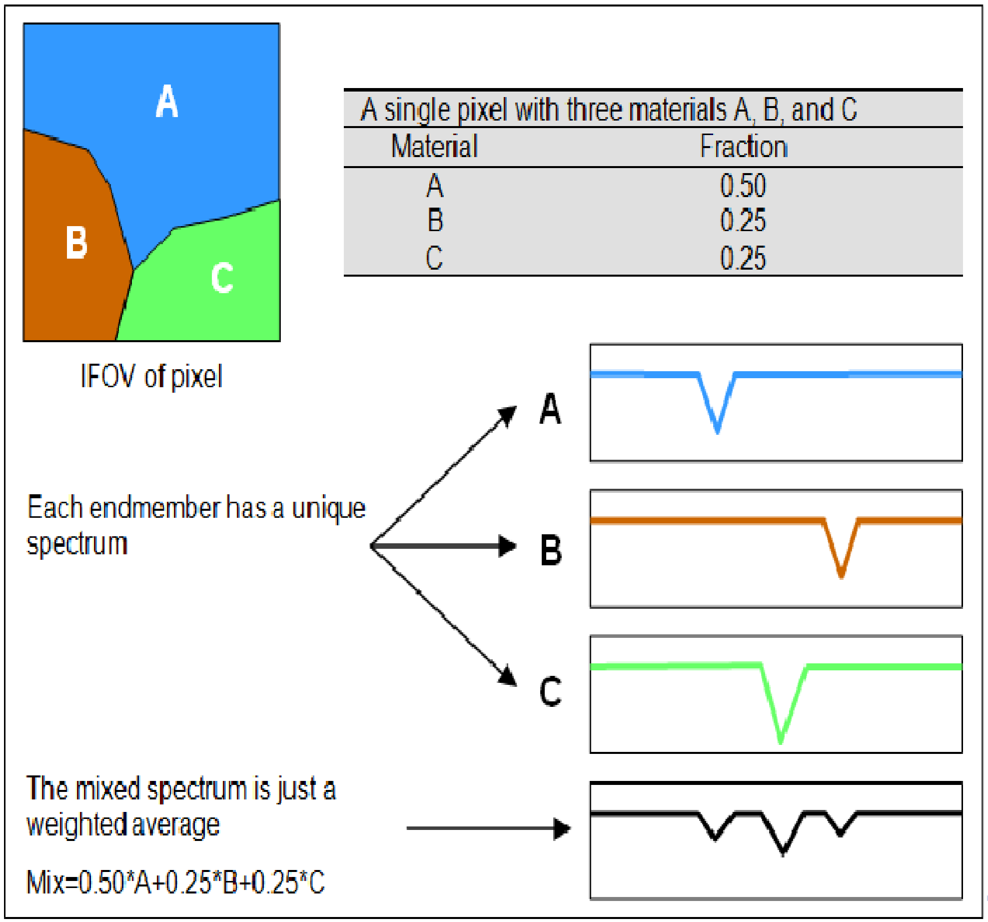


haifengl.wordpress.com

Mixed spectral signatures



M. E. Klein et al

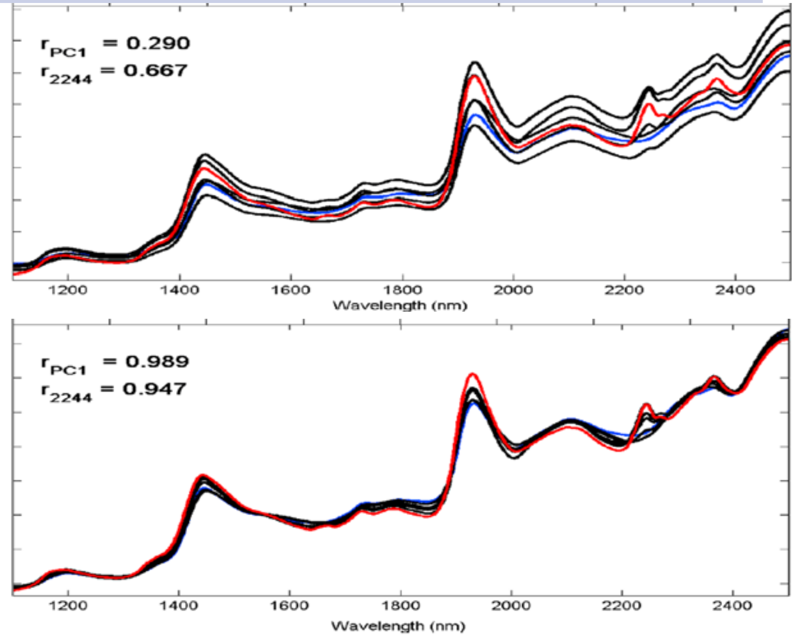


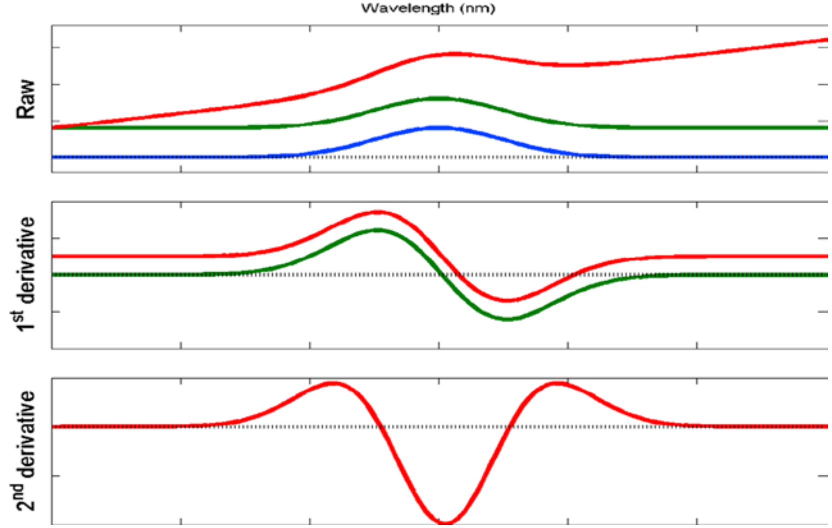
- Methods: MCR, LU, ICA

M. Kamal et al

Spectral Preprocessing

- Spectral Scatter:
  Norm, SNV, MSC
- High Frequency Noise:
  Savitzky-Golay Filtering
- Baseline correction
- Spectroscopic Feature Localisation:
  Derivative Spectroscopy





Å. Rinnan et al

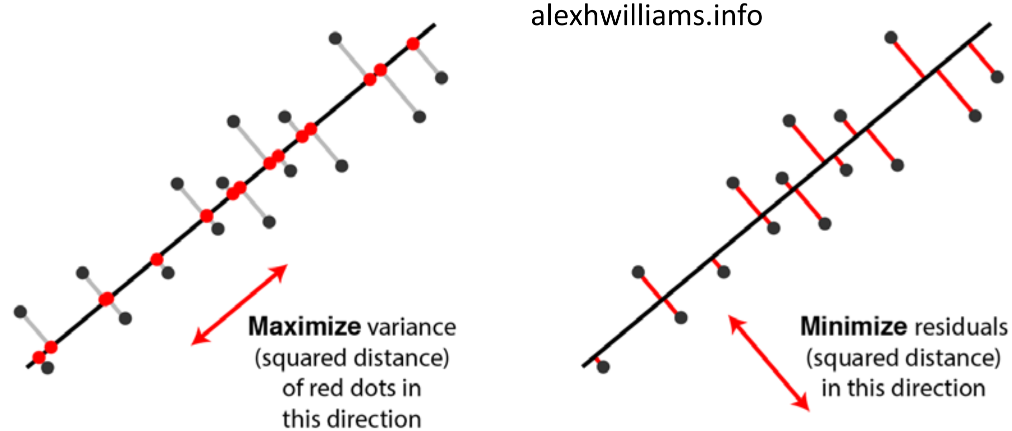
Real-time and inline suitable classification algorithms

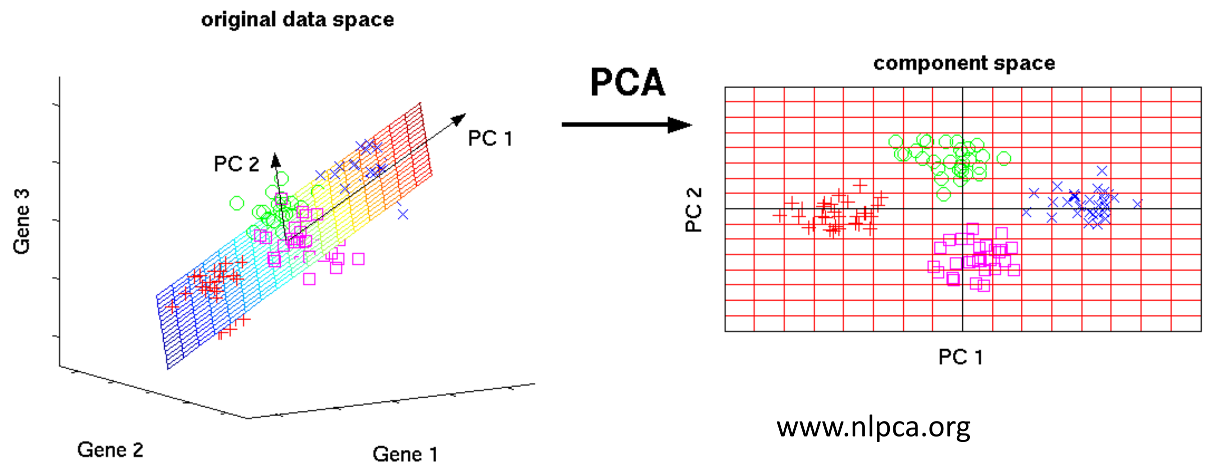
- Subspace Methods:
  PCA, PLS, ICA, LDA
- Cluster Methods:
  k-means, kNN, RDF
- And many more (SVM etc.)…

Real-time and inline suitable classification algorithms

- Subspace Methods:
  PCA, PLS, ICA, LDA
- Cluster Methods:
  k-means, kNN, RDF

alexhwilliams.info

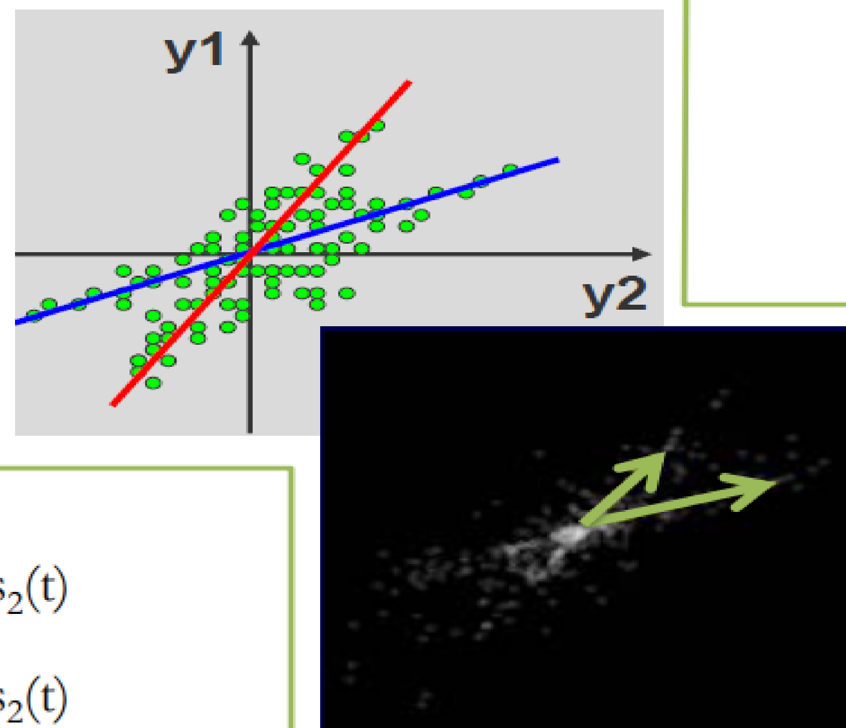


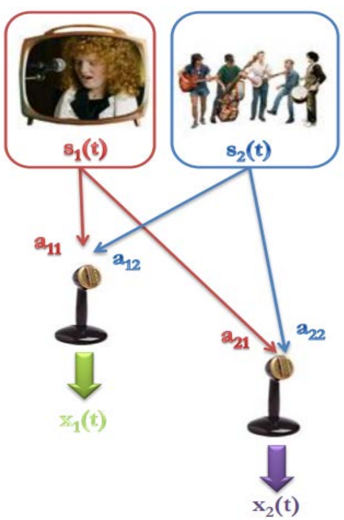


www.nlpca.org

Real-time and inline suitable classification algorithms

- Subspace Methods:
  PCA, PLS, **ICA**, LDA
- Cluster Methods:
  k-means, kNN, RDF





$$x_1(t) = a_{11}s_1(t) + a_{12}s_2(t)$$

$$x_2(t) = a_{21}s_1(t) + a_{22}s_2(t)$$

ICA finds directions which maximize independence (using higher order statistics)

www.sci.utah.edu

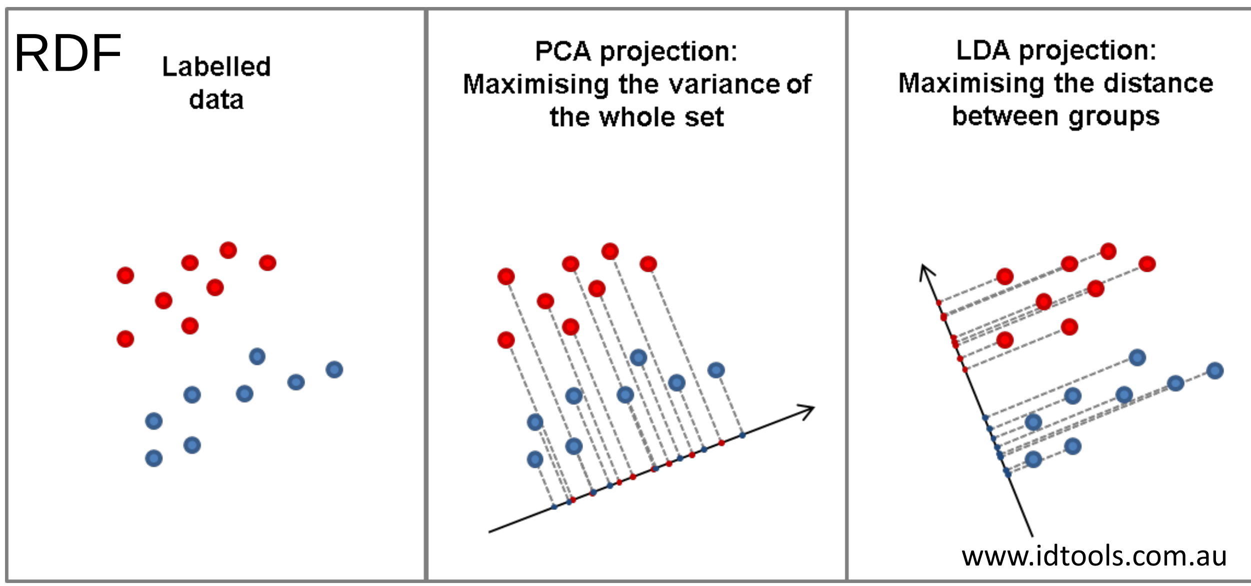
Real-time and inline suitable classification algorithms

$$J(w) = \frac{\left|\tilde{\mu}_1 - \tilde{\mu}_2\right|^2}{\tilde{s}_1^2 + \tilde{s}_2^2} = \frac{w^T S_B w}{w^T S_W w}$$

www.sci.utah.edu

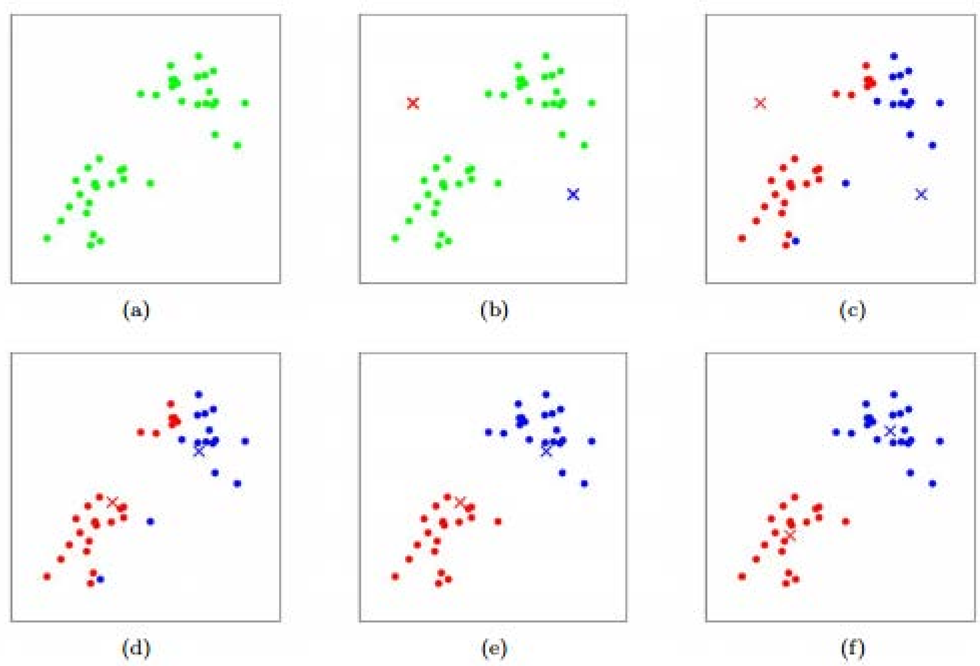
Hence ***J(w)*** **is a measure of the difference between class means (encoded in the between-class scatter matrix) normalized by a measure of the within-class scatter matrix.**

- Subspace Methods:
  PCA, PLS, ICA, **LDA**
- Cluster Methods:
  k-means, kNN, RDF



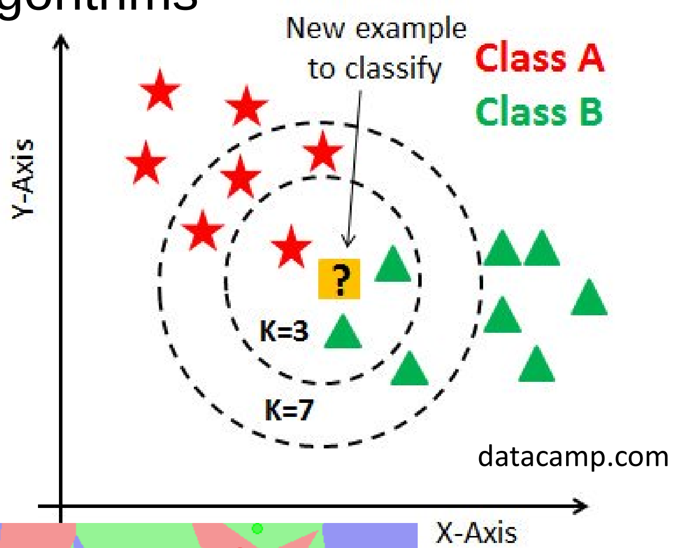
Real-time and inline suitable classification algorithms

- Subspace Methods:
  PCA, PLS, ICA, LDA
- Cluster Methods:
  k-means, kNN, RDF



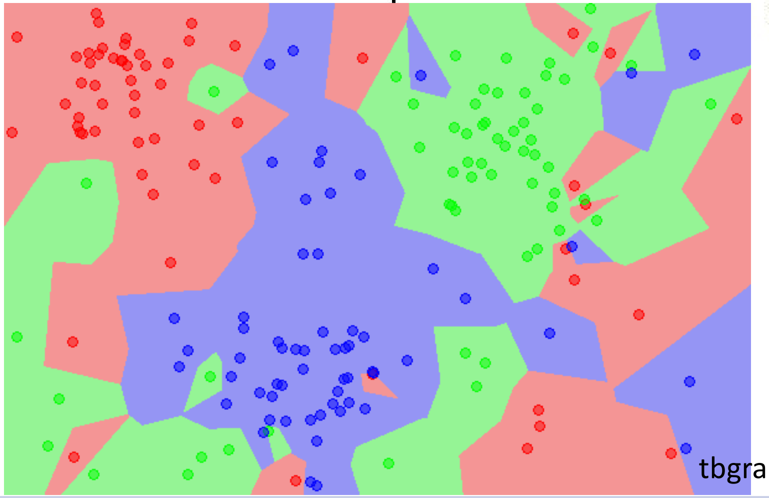
stanford.edu

Real-time and inline suitable classification algorithms



- Subspace Methods:
  PCA, PLS, ICA, LDA
- Cluster Methods:
  k-means, **kNN**, RDF

Problematic for high dimensional spaces!



tbgraph.wordpress.com

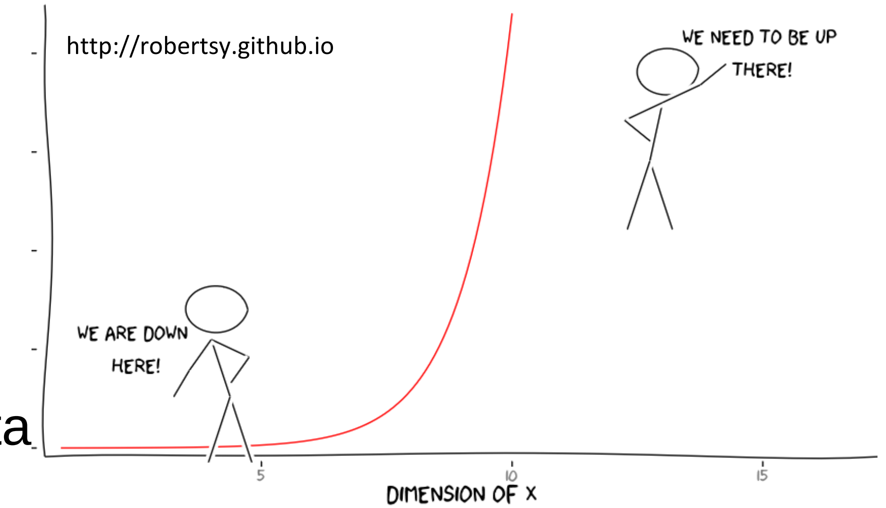
Real-time and inline suitable classification algorithms

- Subspace Methods:
  PCA, PLS, ICA, LDA
- Cluster Methods:
  k-means, kNN, **RDF**

www.analyticsvidhya.com

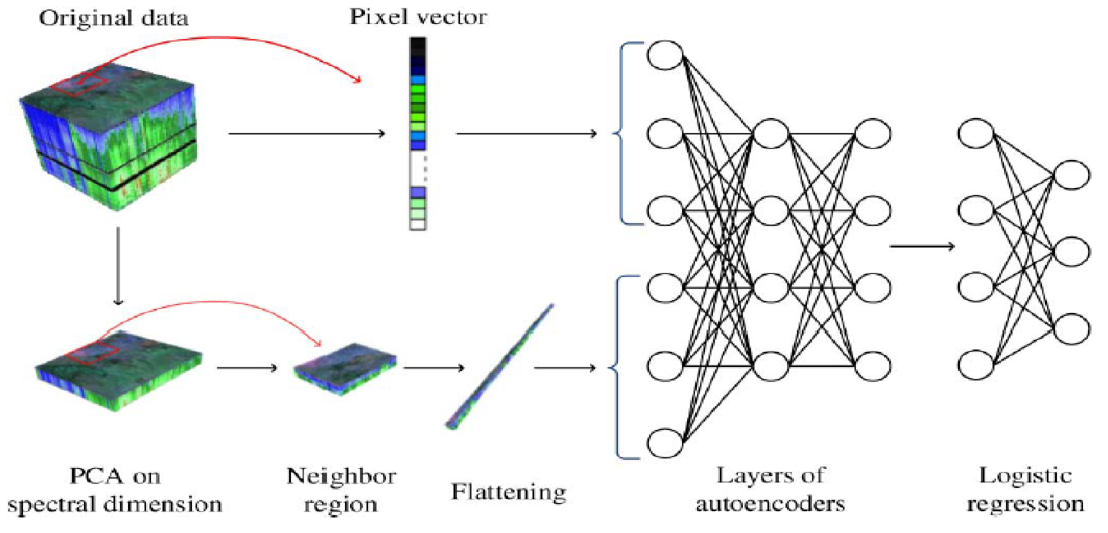
Problems to be solved:

- HSI-data are high dimensional
- Usually few labelled data
- Feature extraction is often difficult, tedious and operator dependent.
- Vital information is lost in current data reduction methods due to neglection of image features.

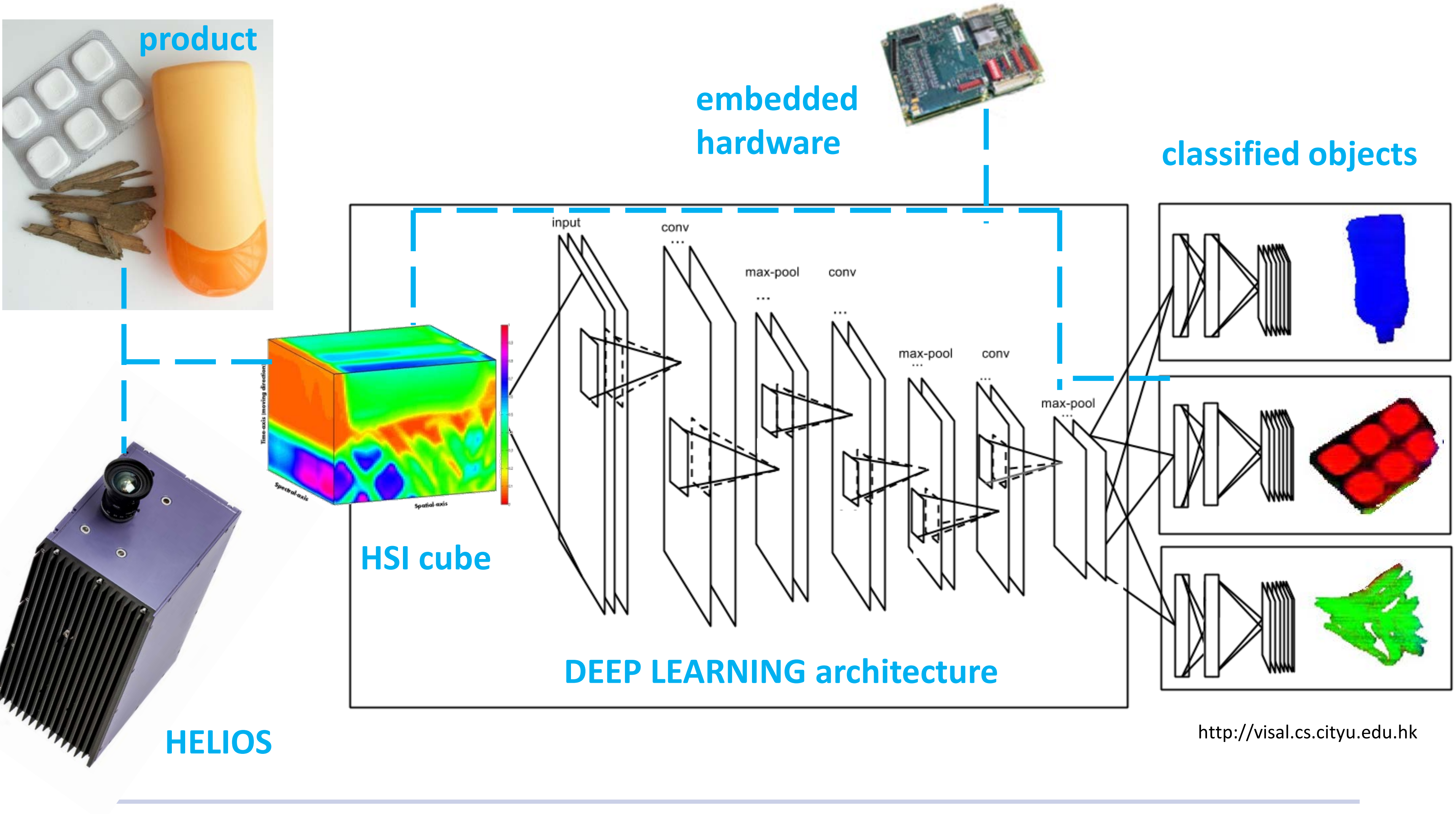


The promise of deep learning approach:

- Combination of spectral and spatial domains addressed in a single approach.
- Powerful and reliable computing hardware now available.
- Strong scientific evidence and technology penetration (e.g. CNN, SAE etc.) in leading high-tech companies.

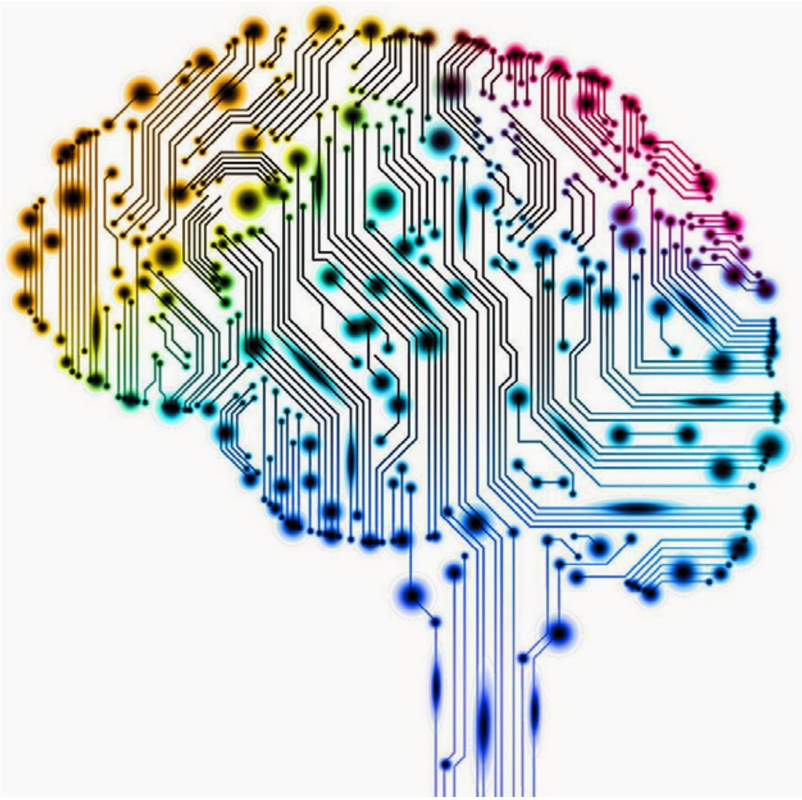


Chen et al.


product
embedded
hardware
classified objects
input
conv
max-pool
conv
max-pool
conv
max-pool
HSI cube
DEEP LEARNING architecture
HELIOS
http://visal.cs.cityu.edu.hk

Benefits:

- Best performing classification results in current bench mark computer vision problems -> application efficiency in food apps?
- Seamless classification model generation without operator variance -> cost reduction, process reliability
- New customer benefits via big data (mining) and predictive analysis/maintenance.



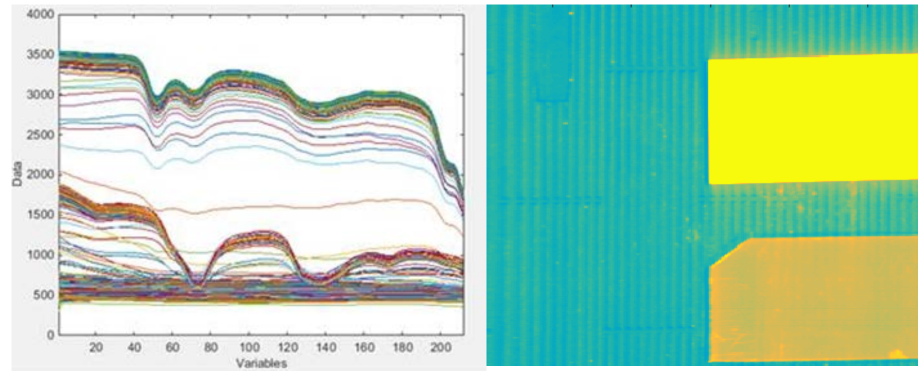
http://konwersatorium1-ms-pjwstk.blogspot.co.at

- Ongoing activities using deep neural network classification with HELIOS data

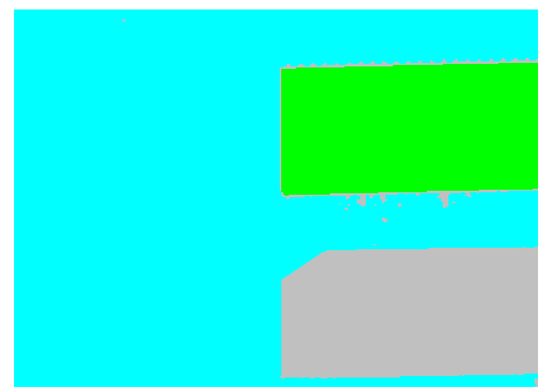
- Current results demonstrate a strong potential of this technology (reduced test times, superior classification power)

- Further R&D projects related to deep learning are planned.

raw HSI input cube of test samples on background



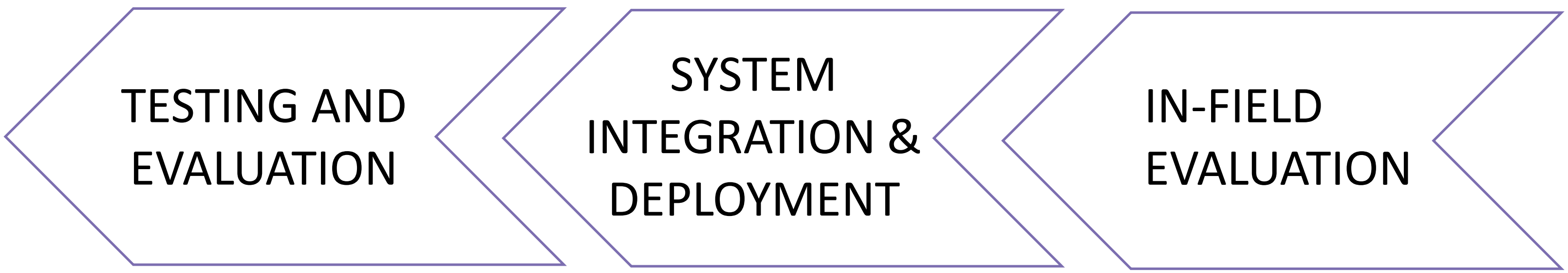
deep neural network classification result



WE BRING THE LAB INTO YOUR LINE

From at-line to in-line quality control

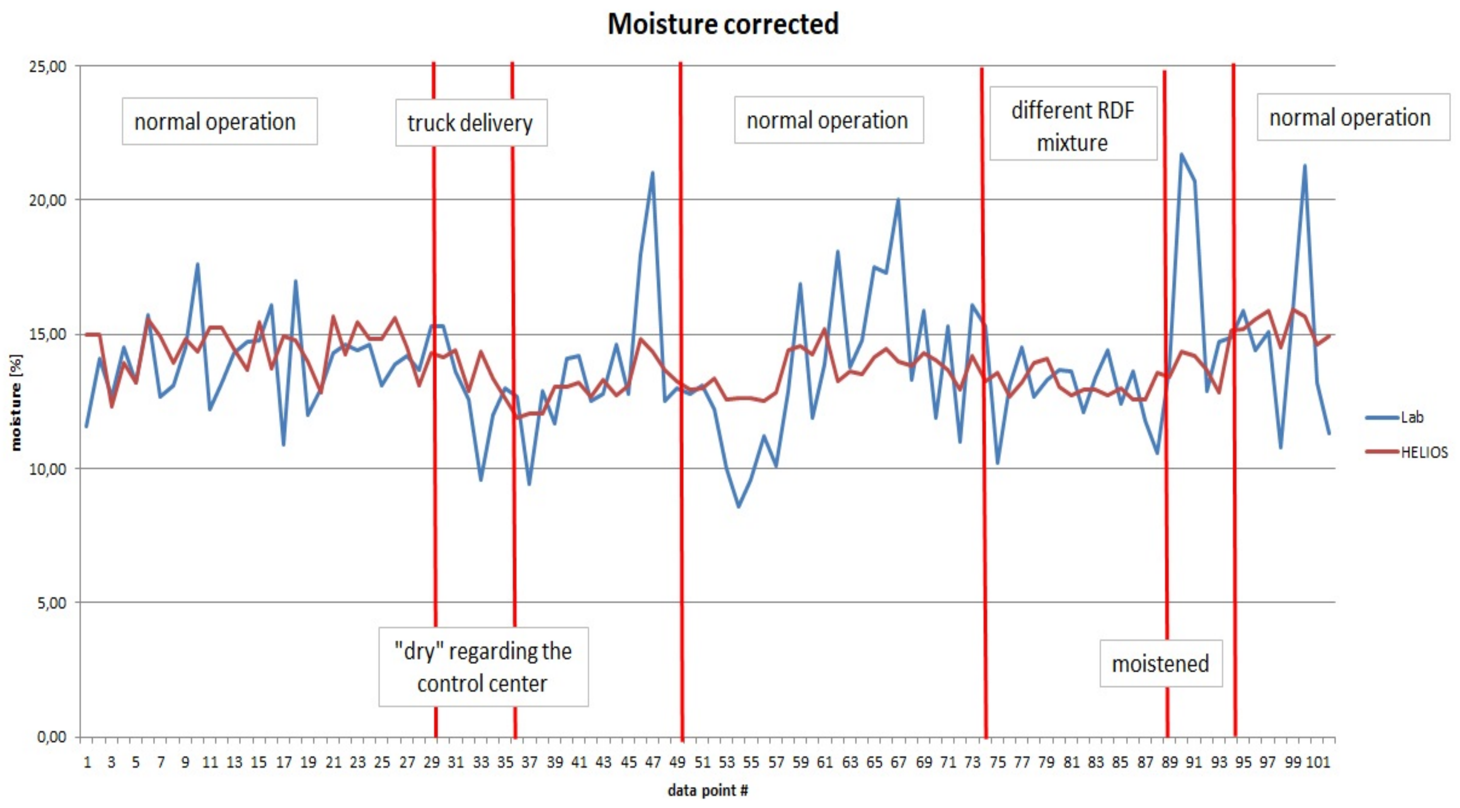
WHAT DOES IT TAKE



Refuse Derived Fuels

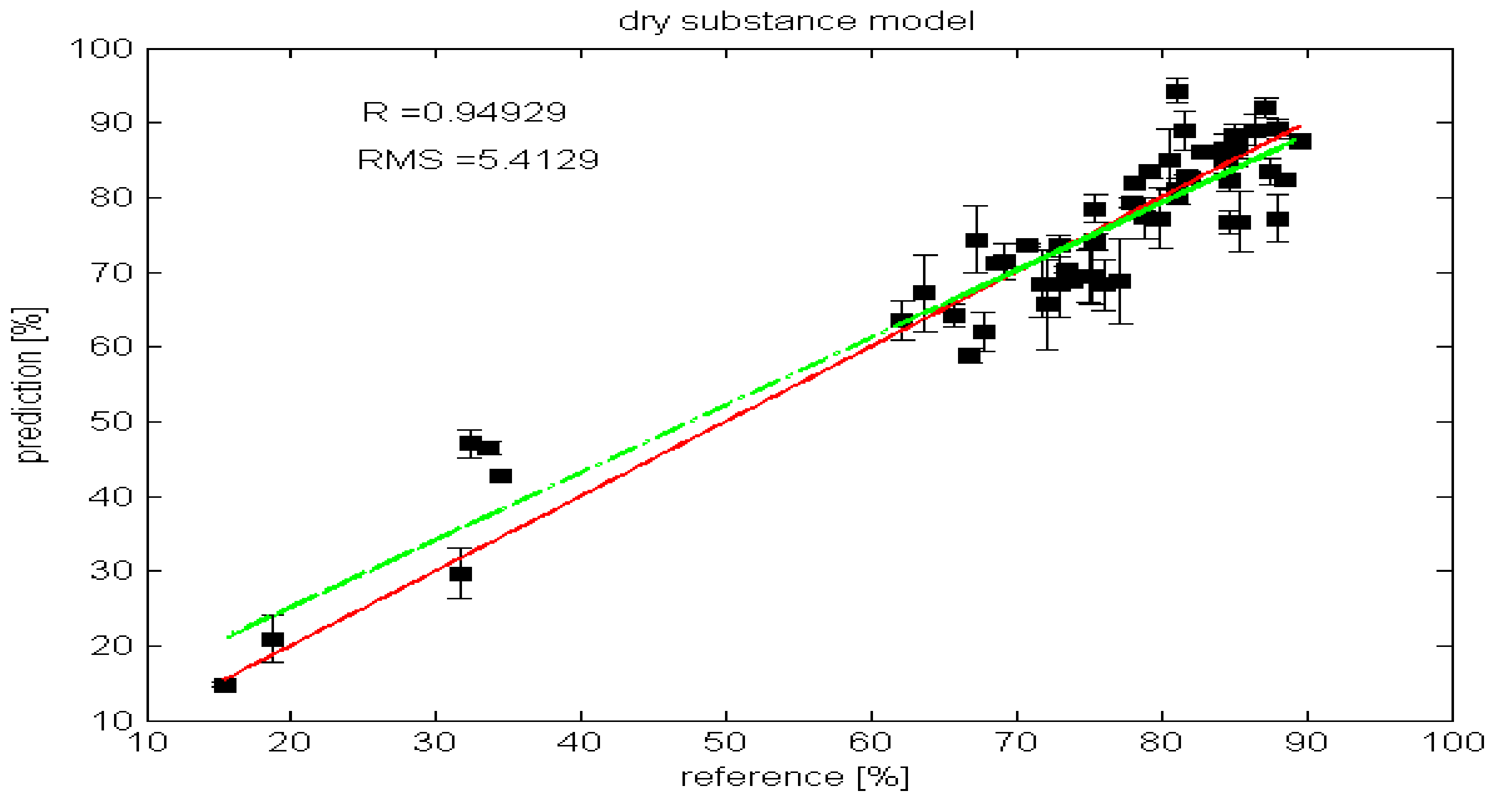
STATUS: **FIELD TESTS**

Humidity vs. Time in RDF stream: Lab vs. QCI HELIOS.



DRY MATTER DETERMINATION

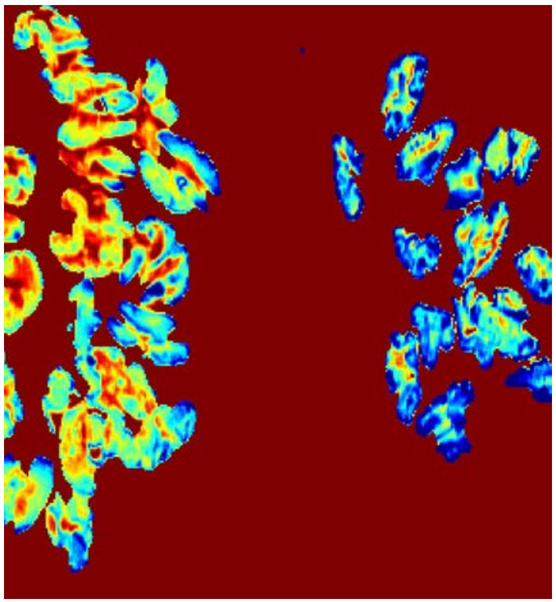
STATUS: **FULL PRODOCTION**


dry substance model
R =0.94929
RMS =5.4129
prediction [%]
reference [%]
100
90
80
70
60
50
40
30
20
10
10
20
30
40
50
60
70
80
90
100

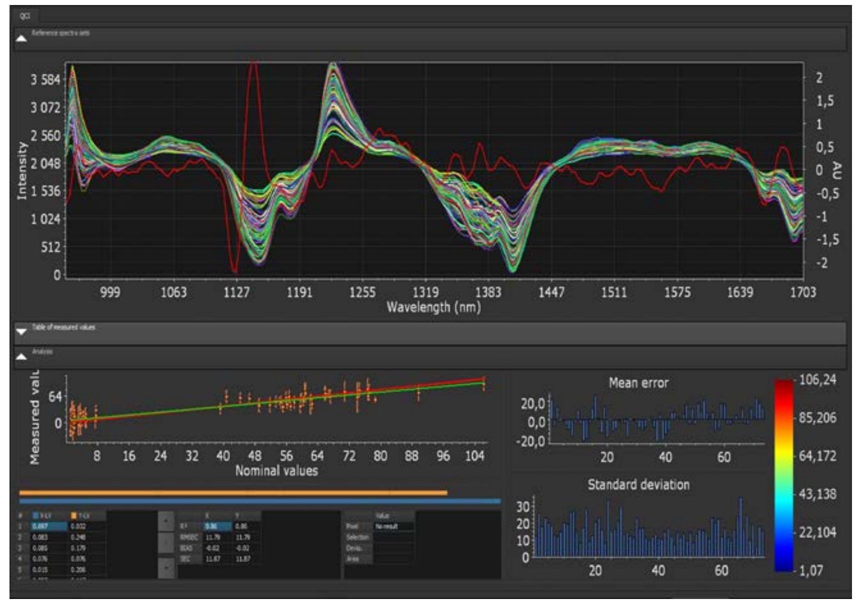
RANCID WALNUTS

STATUS: **MARKET INTRODUCTION**

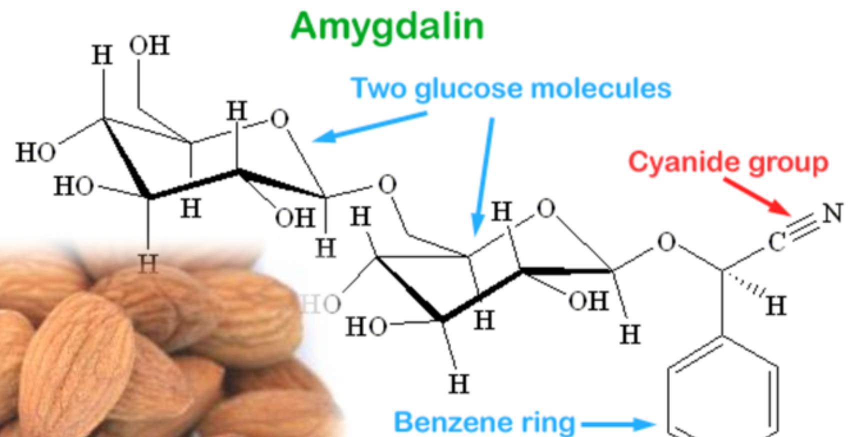
Heat map walnuts



Graphical user interface



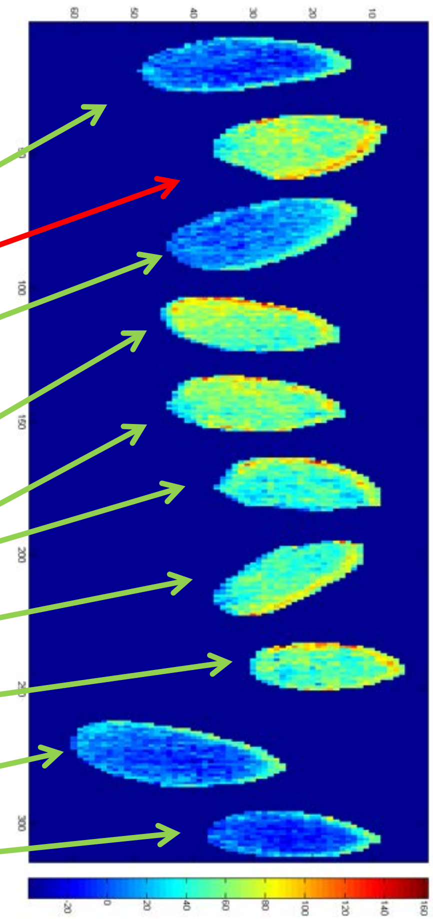
IN-LINE MEASUREMENT OF AMYGDALIN IN BITTER ALMONDS



Source: “The Chemistry of Almond Flavor”, Mitchell, University of California Davis The Almond Conference, 2015

STATUS: **MARKET INTRODUCTION**

Sample#	Amygdalin [mg/g]
SBA 1	4,6
SBA 2	3,7
SBA 3	3,9
SBA 4	63,3
SBA 5	58,6
SBA 6	57,2
SBA 7	66,3
SBA 8	55,5
SBA 9	4,8
SBA 10	4,1



Phenotypes in raw almonds:

__Non-bitter__

Sweet snacking almonds (nutty flavor)

__Bitter__

3-5% amygdalin content, cyanide aroma, hazardous to health

IDENTIFICATION OF MYCOTOXINS IN FIGS

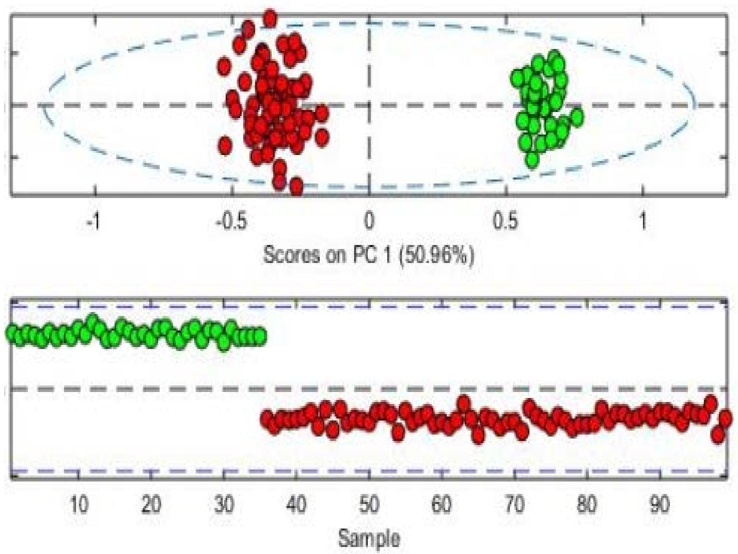
PRELIMINARY

PARTNER PROJECT

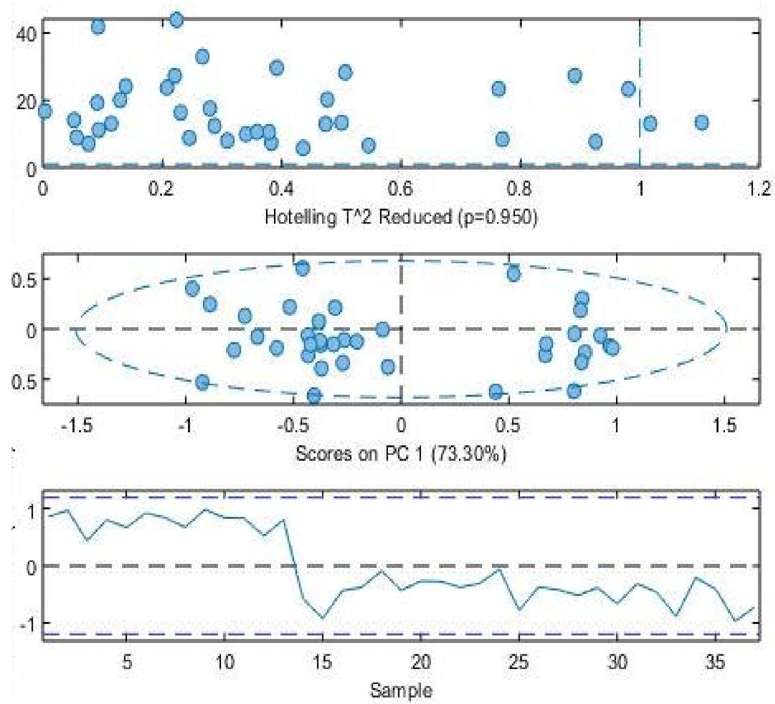
- Aflatoxins are toxic carcinogenic secondary metabolites produced by fungal species
- Great variety of crops is affected
- They are an increasing problem due to climate change
- This leads to huge economic losses

PRELIMINARY

Results



Tests



PRELIMINARY

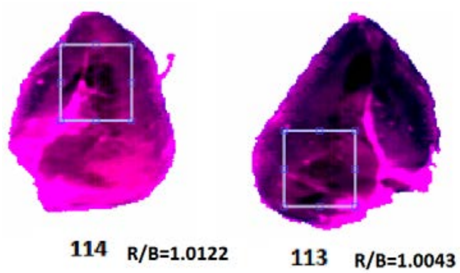
DETECTION OF WOODEN BREAST DEFECT

- Increasing prevalence in broiler production
- Considerable economic losses
- Currently manual inspection/hand sorting

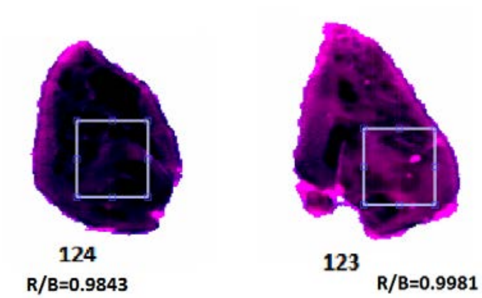
Ref: http://www.helsinki.fi/food-and-environment/research/groups/meat.html

STATUS: **FULL PRODUCTION**

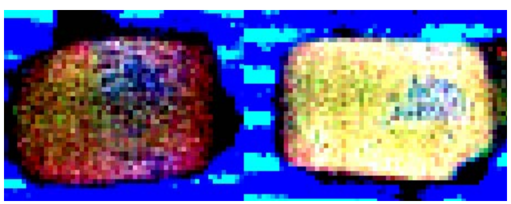
DETECTION OF WOODEN BREAST DEFECT



Bad product: Wooden breast filets



Good product: Ordinary chicken filets



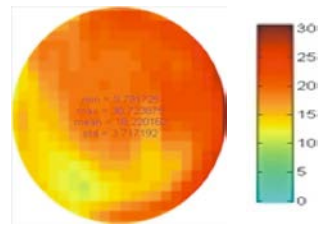
Chemical colour classified stream:

API CONCENTRATIONS

STATUS: **FULL PRODUCTION**

At-line quality control of pharmaceutical products

Chemical concentrations traced by spectral signatures, spatially resolved.



In-situ inference of quantitative values possible at bulk flow velocities of several m/s without destroying the product for every camera image pixel.

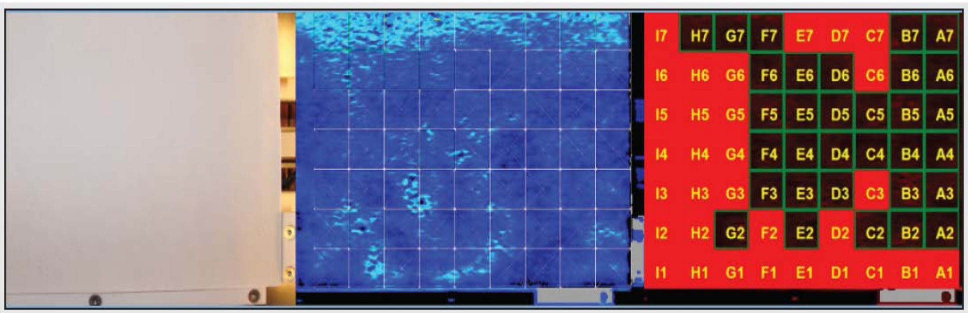
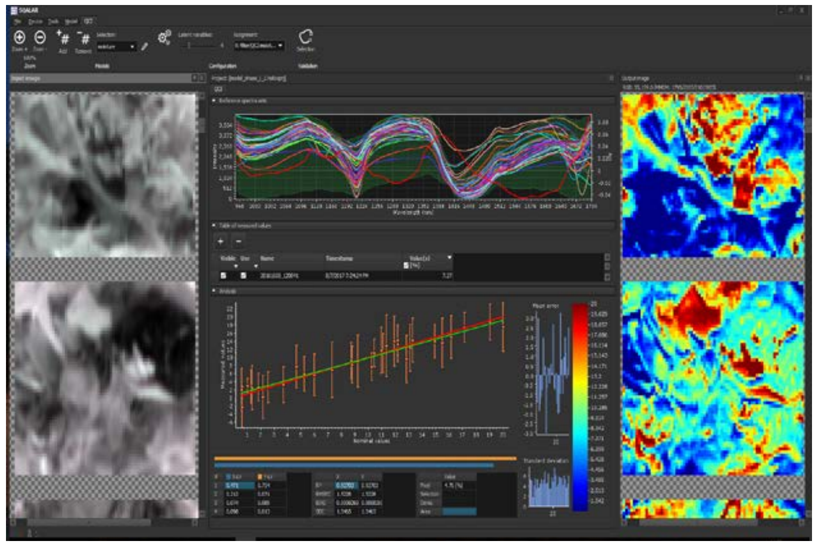


Figure 2: Left: pharmaceutical sponge with active substance on its surface. Centre: quantitative 'chemical colour' representation of concentration. Right: zones (squares) with sufficient and insufficient concentrations

EVK Software Tools for multivariate measurements
of **CHEMICAL COMPOSITIONS**

Brings the LAB INTO your LINE



inVISION **TOP** INNOVATION **2018**

For use IN-LINE in combination with
EVK Helios camera systems

- Currently used HSI data analysis methods are spectroscopy motivated and for the most part ignore imaging information.
- AI seems to be a good approach to overcome both the limitations of high dimensionality and the somewhat historical segregation of spectral and spatial information
- Quantitative Chemical Imaging (QCI) using modern algorithms opens a new window into the world of hyperspectral applications

- H. Bischof, TU Graz, “Machine Learning and Hyperspectral Imaging” (talk chii 2016)
- H. Lohninger, TU Wien, “Hyperspectral Imaging – Problems and Solutions from a Statistical Perspective” (talk chii 2017)
- Y. Chen et al., “Deep Learning-Based Classification of Hyperspectral Data”, IEEE Vol. 7, No. 6, June 2014

EVK