duced with a Trial Version of PDF Annotator - www.PDFAnnotator.com

→ The ultraviolet range :
↳ is not used because of the atmospheric absorption.

→ Visible range measurements :
↳ are usually made with passive systems.

→ The radiometric resolution :
↳ is expressed in equivalent ground meters.

→ A rectangular detector array :
↳ can be used to obtain multispectral or hyperspectral data.

→ A synthetic aperture radar (SAR) :
✓ ↳ allows an acceptable azimuth resolution at spacecraft altitude.

→ A radiometric error car result :
↳ from the effect of the atmosphere.

→ Because of sky irradiance :
↳ a particular pixel will be irradiated by energy scattered downwards from atmospheric constituents or from surrounding pixels.

→ Mie scattering :
↳ accounts for the white color of the clouds.

→ Haze removal :
↳ is a correction technique for radiometric distortions.

→ The matrix rotation $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}$
↳ represent the mathematical modeling for rotation correction.

→ Sequential Similarity Detection Algorithm (SSDA) :
↳ is an automatic method for control point localization which make use of correlation

→ the two images :

↳ PCA is more usefull for image (a).

a

b

→ the correlation matrix :

↳ shows that the real nbr of dimensions of data is <u>two</u>.

$$R = \begin{bmatrix} 1 & 0.377 & 1 \\ 0.377 & 1 & 0.377 \\ 1 & 0.377 & 1 \end{bmatrix}$$

→ Principal component decomposition :

↳ can be used for Bandwidth compression.

→ The human photointerpretation of an image :

↳ allows only a limited multispectral analysis.

→ In the statistical supervised classification of the pixel in an image :

↳ each spectral class is modeled with a Gaussian distribution

→ Let the spectral classes for an image be represented by w_i, $i = 1, \dots, M$. M is the total nbr of classes, and take a column of brightness values for one pixel x. In the Bayes classification process, the <u>posterior</u> probabilities are :

↳ ~~$p(w_i/x)$~~ ↳ ~~$p(x/w_i)$~~ ↳ $\underline{p(w_i/x)}$.

→ To avoid poor classification, max. likelihood classification can make use of thresholds that :

↳ define regions where the discriminant functions for adjoining spectral classes are equal.

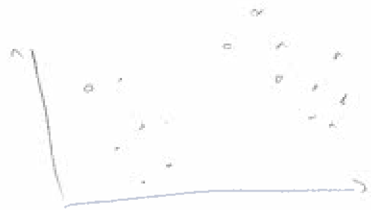
-> The KNN classifier:

↳ takes into account the spatial context of the pixels.

-> The effect of spatial contexts can be incorporated into a classificator using:

↳ KNN ???

-> The N-dimensional data in fig.:



↳ $w^T x + w_{N+1} = 0$.

-> The training operations during the linear discrimination of data:

↳ non-parametric iterative procedures.

-> A support Vector classifier:

↳ provides a training approach that depends only on those pixels in the vicinity of the separating hyperplane.

-> "Slack variables" in SVM:

↳ training set ~~available~~ reliable but the data is affected by noise.

-> The clustering algorithm referred to Kmeans:

↳ nbr of clusters must be selected beforehand by the ~~d~~ user.

-> A single pass clustering technique:

↳ use the 1st row of samples to start a center of cluster.

-> Clustering by histogram peak selection:

↳ usefull with low dimentionality of the data.

→ ~~The correlation matrix between the bands in HSI data.~~

→ SIFT detector is based on a pyramid scale space of

↳ Difference of Gaussian.

→ High variance:

↳ overfitting.

→ In CNNs, the convolutional layers reaches Shift invariance by:

↳ shared weights.

→ Max pooling layer in DNN:

↳ reduce activation by taking the max value with a certain stride.

→ → A side looking RADAR:

↳ is used for microwave wavebands.

→ Rayleigh scattering:

↳ strongly λ dependent.

→ Panoramic distortion:

↳ depends on the Earth curvature.

→ $$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1.411 \end{bmatrix}$$

↳ aspect ratio correction.

→ Ground Control Points (GCP)

↳ relationship between r.s. image and map.

→ Taylor method for contrast enhancement

↳ relies on PC decomposition.

→ Band ratios :
↳ can be used to find vegetated regions .

→ unsupervised classification in an image :
↳ clustering methods.

→ classes w_i $(i = 1, \dots, M)$, x : brightness . In Baye's classification process, the a priori or prior probabilities are :
↳ $p(w_i)$

27] X → In the Bayes classification process, $p(w_i)$:
↳ can be estimated by the analysit . / ↳ ~~are always unknown~~

→ In the Bayes classification, the Decision surfaces :
↳ are used to compute the class membership of a pixel .

→ The minimum distance classification :
↳ is faster than the maximum likelihood classification .

→ The KNN classifier :
↳ is not well suited for HSI data .

→ The concept of spatial context :
↳ are usefull to remove individual pixel labelling errors, that might result in noisy data .

→ The ISODATA algorithm :
↳ allows for different numbers of clusters .

→ Classification time :
↳ increases quadratically for Max. likelihood decision and parallelpiped classification . linearly for min. distance and

① Source & char. of remote sensed image data

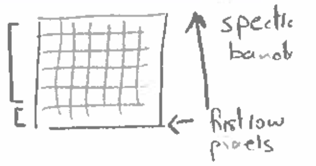
→ UV measurement are not made because of atmospheric absorption.

→ ~~"optical" remote sensing → visible to I.Red~~

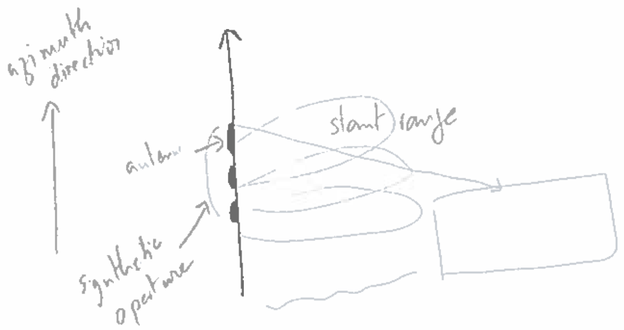
→ Visible - Infrared - microwave ranges → remote sensing.

→ Passive sensors : visible - Infrared. 0.4 — 12 um.
Microwave 30 — 300 mm.

→ rectangular detectors placed on a satellite → used to get HSI data, since it records the brightness of every pixel in multiple spectral bands.



→ Side looking Radar → in microwave region.



→ SAR : Synthetic Aperture Radar
↳ uses a synthetic aperture for the antenna
↳ better resolution.

② Error correction and registration of image data

→ Scattering (effect by the atmosphere)

① Rayleigh scattering → wavelength dependent [blueness of the sky

② Aerosol or Mie scattering → smoke / haze / fumes / clouds / dust

→ Correction of atmospheric effect on the radiometric errors :
↳ "Haze" removal : shifts back the histogram of the blue color band.

(3) Multispectral transformations of image data

PCA ——

Taylor method of contrast enhancement
(decorrelation stretch)

→ Principle Component decomposition.

↳ used for Bandwidth compression.

↳ used for HSI.

(4) The interpretation of digital image data

Applications of ML:

- Classification
- Detection
- Segmentation

→ Non-parametric methods: KNN - SVM

$$\text{Precision} = \frac{TP}{TP + FP}$$

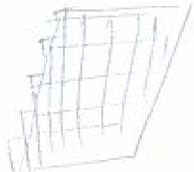
$$\text{Recall} = \frac{TP}{TP + FN}$$

$$\text{Accuracy} = \frac{TP + TN}{\#\text{ of samples}}$$

→ Cross-validation: used to find a good balance between bias and variance.

→ During radiometric correction :
↳ correction of instrumentation error → contrast enhancement
[destriping]
↳ using mean and standard deviation.

→ Rotation of the earth :



↳ the image acquired is shifted at the bottom
↳ for correction : ~~we correc~~ offset the bottom of the image by amount of movement of the ground during acquisition.

→ Space crafts in low orbit, are not affected by the Earth curvature.

→ Aspect Ratio distortion : distortion of the size of the image
↳ because of underscanning or overscanning.
↳ because scan are taken too quickly compared to IFOV.

→ oscillating mirrors → non-linearities.

→ Ground control points (GCPs) : set of features on the map that can also be identified on the image.

→ Interpolation
① Nearest neighbor resampling :
② Bilinear interpolation :
③ Cubic convolution interpolation

→ Aspect Ratio correction :

$$\underset{\text{map}}{\begin{bmatrix} x \\ y \end{bmatrix}} = \begin{bmatrix} 1 & 0 \\ 0 & 1.411 \end{bmatrix} \underset{\text{image}}{\begin{bmatrix} u \\ v \end{bmatrix}}$$

→ SIFT algorithm : used to localize and match interest points in both images.

→ RANSAC : to filter out bad matches which affect the result of registration.

(5) Supervised (statistical) classification techniques

Bayes classification

→ we need to find $p(w_i/x)$. $i = 1, \dots, M$

→ classification performed as:

$x \in w_i$ if $p(w_i/x) > p(w_j/x)$ at all $i \neq j$

→ $p(x/w_i)$ is available, estimated from training data.

$$p(w_i/x) = \frac{p(x/w_i) \cdot P(w_i)}{P(x)}$$

$p(w_i)$ → a priori or prior probabilities.
$p(w_i/x)$ → posterior probabilities.

Maximum likelihood:

→ Decision surface: surface that separates classes

→ Threshold: is used with poor classification.
↳ pixels with prob. below the threshold, are not classified.
↳ thresholds are applied to the discriminant functions, not to the probabilities.

→ N dimentional multispectral space, covariance matrix $= N \times N$
↳ it has $\frac{1}{2} N (N+1)$ distinct element.

→ Max. likelihood classification depends on "mean" and "covariance matrix" Σ

Minimum distance classifier:

→ determines the "means" only.
→ faster than max. likelihood.
→ not as flexible.

→ Classification time for max. likelihood increases " quadrically ".

→ " Spatial context " : taking into account adjacent pixels when performing classification.
↳ solution to " acquisition noise ".

→ "Probalistic label relaxation" : allow the spatial properties to be part of the classification.

→ Markov Random fields : by considering the full image.
↳ maximizes the global posterior prob. $p(\Omega / X)$.

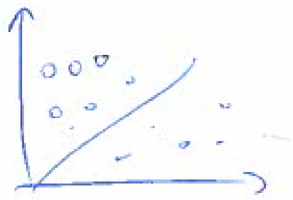
6 Supervised non-parametric classification

→ KNN : K - Nearest Neighbor classifier

→ time - consuming.
→ assumes that pixels that are close to each others, belong to the same class.
→ A pixel is classified by checking to which groups of pixels (that are already trained) is it the closest.
→ computes the distance from the " unknown " pixel to all the classified ones and compares to take the shortest distance.
→ " Euclidian distance ".
→ discriminant function $g_i(x) = k_i$
→ decision rule : $x \in w_i$ if $g_i(x) > g_j(x)$ $j \neq$
→ NOT suited for HSI.

→ The decision rule :

$x \in$ class if $w^t x + w_{N+1} > 0$



→ **SVM** : Support Vector Machine

→ SVM provides a training approach that depends only on the pixels that are on the vicinity of the hyperplane.
[called the support vectors].

↳ It provides an OPTIMAL hyperplane that helps in the training patterns.

↳ The optimal orientation of the hyperplane is when there is a max. separation between the patterns in the two classes.

↳ we can draw two further hyperplanes parallel to the optimal one, bordering the nearest training pixels from the two classes.

→ **Slack** variables :

→ when we have "class overlap", we introduce a slack variable to relax the requirement for a hyperplane and thus, find a max. marging solution.

→ **Kernel** trick :

→ use of feature space transform ϕ

→ map data in a new space, higher dimension and then classify it with a linear separation.

7 Clustering and unsupervised classification

→ Clustering : used for data mining

→ an image is segmented into unknown classes. It is the task of the user to label it afterwords.

→ Clustering implies a grouping of pixels in a mutidimentional space

→ we calculate the Euclidean distance between two pixels to check their similarity.

→ once a candidate clustering has been found, it is desirable to have a means by which the "quality" of clustering can be measured.

→ K- means : Iterative optimization algorithm

(1) . . . $\hat{m}_i$

(2) . . .

(3) The new set of means that resulted from step (2) are computed. m_i , $i = 1, \ldots, c$.

(4) if $\hat{m}_i = m_i$ for all i, the procedure is terminated, otherwise $\hat{m}_i$ becomes as (m_i) and the procedure returns to step (2).