Analyzing images

Anders Kaestner :: Laboratory for Neutron Scattering and Imaging

Outline – Analysis

Introduction

Introduction

The previous lectures we looked into methods to prepare images.

Filtering

- Intensity adjustments
- Denoising
- Edge detection

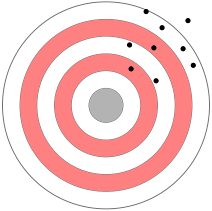
Segmentation

- Classification
- Labeling
- Post segmentation cleaning

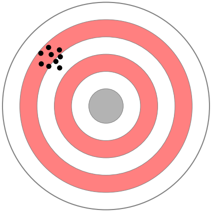
These steps are needed to provide information for the analysis.

Precision and Accuracy

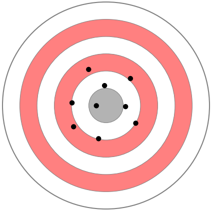
Low Precision and Accuracy



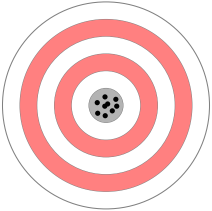
High Precision and Low Accuracy



Low Precision, High Accuracy



High precision and accuracy



Basic image metrics

Counting pixels

Counting pixels is the most simple measurement on an image:

Area and volume

2D The number of pixels corresponds to the area.
Each pixel has the unit area (pixel size)^2

$$Area_i = {L_{pixel}}^2 \cdot \#(\text{pixels} \in \text{label}_i)$$

3D The number of voxels corresponds to the volume.
Each pixel has the unit volume (voxel size)^3

$$Volume_i = {L_{voxel}}^2 \cdot \#(\text{voxels} \in \text{label}_i)$$

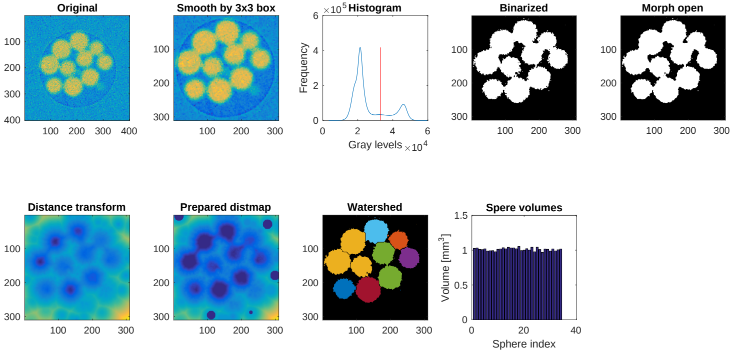
Efficient area measurements in labeled images

Area/volume per item in a labeled image: Using a histogram with N_{Labels} bins.

Measure volume per item

Example

We want to measure the volume of items in a packing of spheres

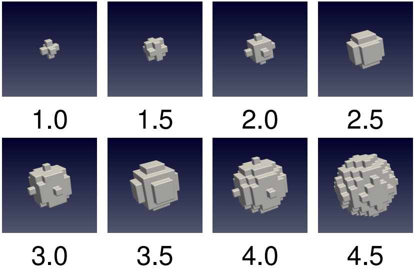


How exact can the volume be measured?

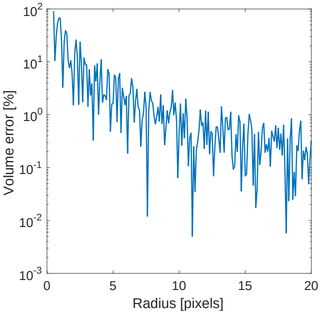
Problem with a discrete grid

- The accuracy decrease when the item size approach the voxel size.
- It is difficult reproduce fine details.

Discrete spheres

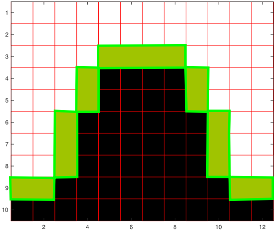


Volume error



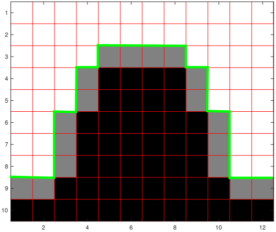
Boundary length

Count pixels



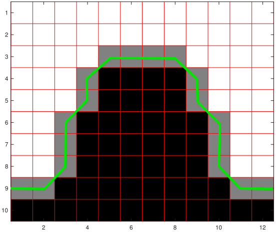
Simple processing.

Pixel outline



Requires neighborhood.

Pixel midline



Requires neighborhood.

Some words about accuracy and precision

- Boundary position depends on threshold.
- The precision depends on noise and metric.

How to measure boundary length

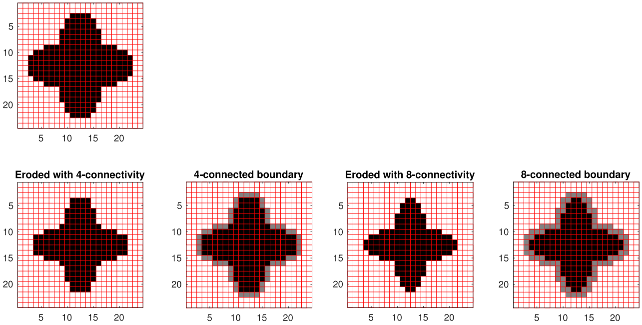
Identify boundary pixels

The boundary pixels of a bilevel image are identified by

$$b(f) = f - \varepsilon_{SE}(f)$$

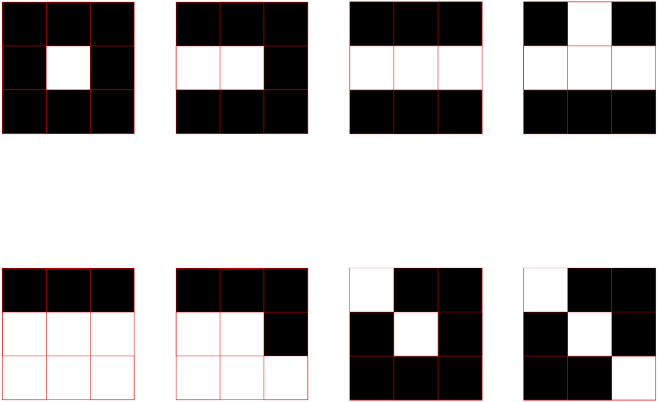
with structure element *SE* representing the direct neighbours.

Example



Pixel patterns on the edge

Each edge pixel type must be identified as a rotation of



In 3D more patterns are added.

Increase edge accuracy

Interpolate edge shape

Sub-pixel accuracy can be achieved with edge interpolation.

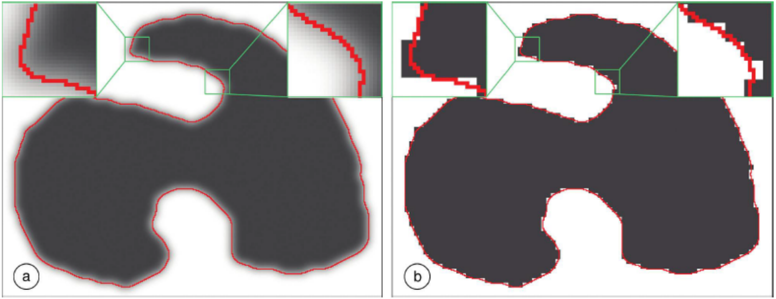


Figure 10.72 Fitting a smooth boundary for perimeter measurement: (a) contour line on a smoothed (antialiased) feature; (b) the contour line superimposed on the original pixels.

Russ (2016)

Metrics using area and volume

The surface area and volume can be used to characterize the particle shape using a sphere as reference shape:

Sphericity

Measures how round a particle p is using the ratio between surface area and volume

$$\Phi = \frac{\pi^{1/3}(6\,V_p)^{2/3}}{A_p}$$

The sphericity is

- One for a sphere
- Other shapes have $\Phi<1$.

Equivalent radius

Assuming the items to be approximate spheres.

$$R = \left(V\frac{3}{4\,\pi}\right)^{1/3}$$

Measures the radius of the sphere with the same volume.

Distance guided measurements

Motivation

Some applications require the knowledge how the intensity depends on the distance.

- Physical quantities like density or water content.
- Edge sharpness

Segmentation

Identify features.

$$f \rightarrow b$$

Distance map

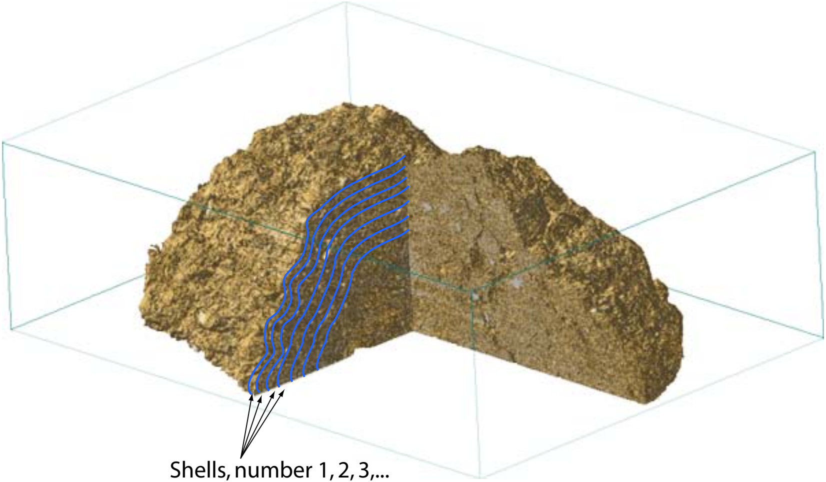
Apply distance transform.

$$D = \mathcal{D}(b)$$

Collect profile

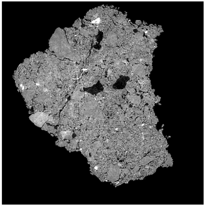
- $s(d) = \{p | D(p) = d\}$.
- Compute average intensity $E[s(d)], \forall d \in D$.

Shell peeling

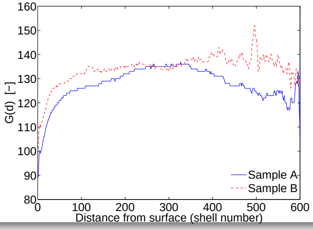


Distance averages

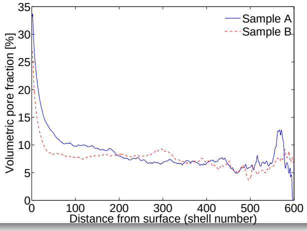
CT slice



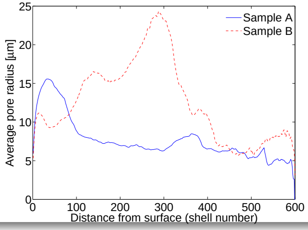
Density



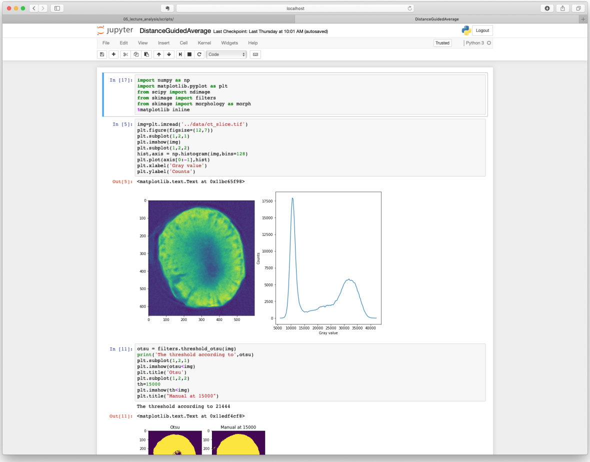
Porosity



Pore radius



Demo notebook



Characterizing shapes

Position

Center of box

- Identify bounding box (smallest box including entire item)
- Using box coordinate to find center point

Simple but may be misleading.

Center of gravity

The mid point of item f is

$$CoG_x = \frac{\sum_{i \in f} x_i}{\text{Area(f)}}$$
$$CoG_y = \frac{\sum_{i \in f} y_i}{\text{Area(f)}}$$

For 3D a third equation is added, normalization by volume.

van Assen et al. (2002)

Variations of the center of gravity

Density weighted

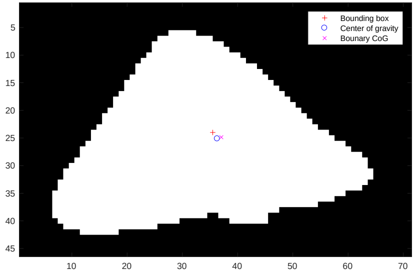
$$CoG_x = \frac{\sum_{i \in f} f(i)\, x_i}{\sum_{i \in f} f(i)}$$

Boundary only

$$b(f) = f - \varepsilon_{SE}(f)$$

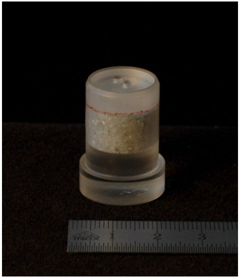
$$CoG_x = \frac{\sum_{i \in b(f)} x_i}{\sum b(f)}$$

Example

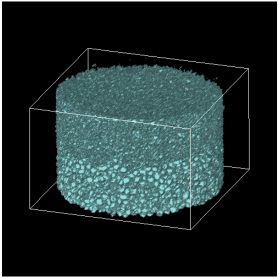


Segmenting two grain size fractions

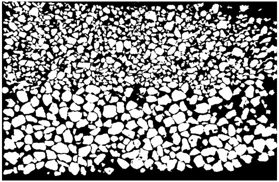
Sample



Volume



Vertical slice



Kaestner et al. (2005)

Separation of sand grains

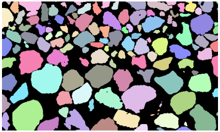
Part I

Segmenting sand grains

1. Create elevation map

$$s_d = d_{\mathcal{E}}(s) - d_{\mathcal{E}}(s^c)$$

2. Use Watershed segmentation



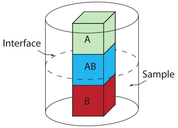
3. Compute
   - CoG
   - Equivalent grain radius

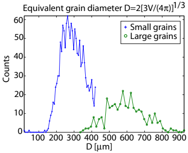
   for each item.

Part II

Classifying sand grains

1. Collect class statistics



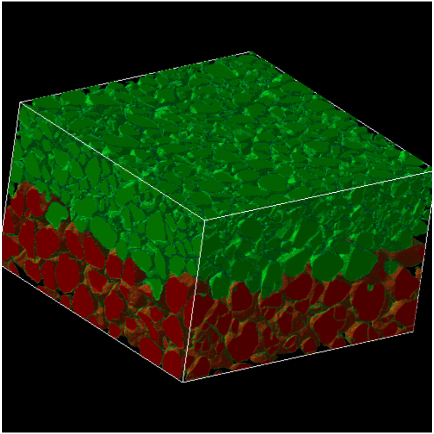


2. Use a spatial Bayesian threshold

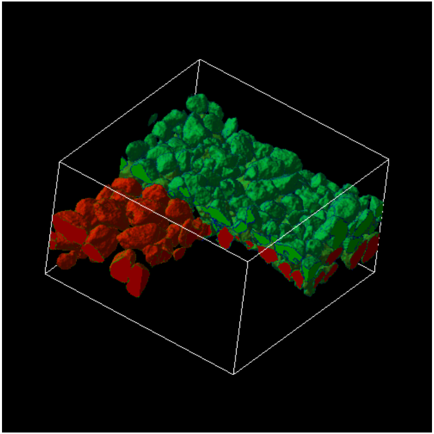
$$\mathcal{R}(d) = \frac{1}{1+e^{-d/\sigma}} \frac{P_B}{P_A}$$

Segmented sands

Segmented sand fractions



Sand grains in the interface



Item separation

Centroid distances

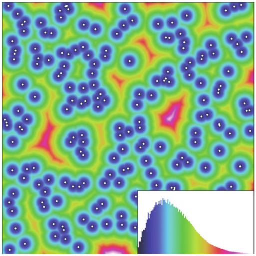


Figure 10.20 Random distribution of points (white) and the Poisson distribution of distances from

Different particle distributions

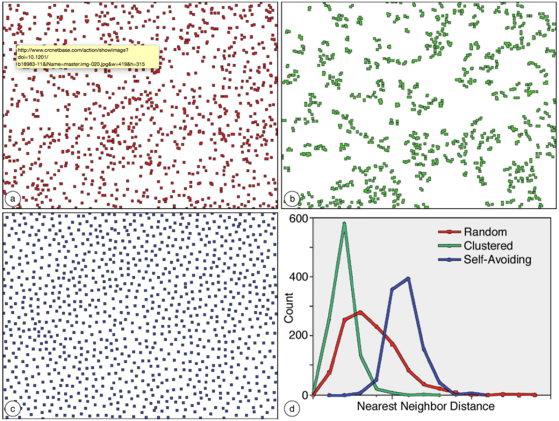


Figure 10.21 Feature distributions illustrating (a) random, (b) clustered, and (c) spaced or self-avoiding arrangements; and (d) histograms of nearest-neighbor distances for each of the point distributions. The mean value of the clustered distribution is less than, and that for the self-avoiding distribution is greater than, the mean for the random one.

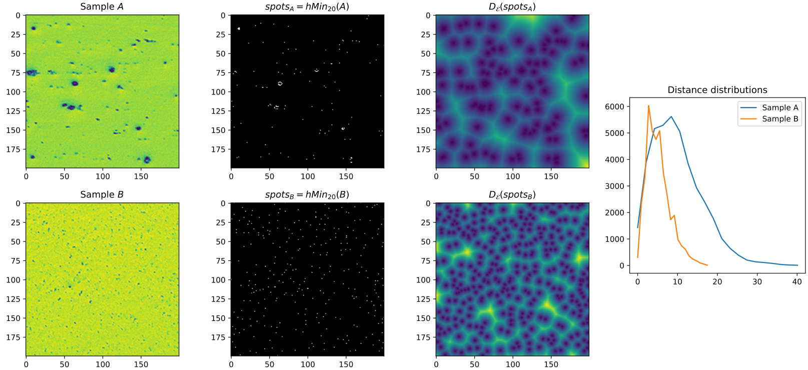
Russ (2016), p. 576

Application of centroid distribution

Procedure

1. Microscope images of samples with defects
2. Find spots using hMin operation
3. Measure the distance between the defects using an Euclidean distance transform

Results



Orientation

Moments

$$\mu_{m,n} = \sum_{i,j} x_i^m y_j^n \cdot v(i,j)$$

$\mu_{0,0}$ Area

$\mu_{1,0}, \mu_{0,1}$ Corresponds to centroid

x, y Positions in item

v Optional weight, e.g. the density.

Normalized moment

$$\eta_{n,m} = \frac{\mu_{m,n}}{\mu_{0,0}{}^{(m+n+2)/2}}$$

Used for size independent shape descriptors.

Orientation

$$\theta = \frac{1}{2}\tan^{-1}\left(\frac{2\mu_{1,1}}{\mu_{2,0} - \mu_{0,2}}\right)$$

Grain orientation

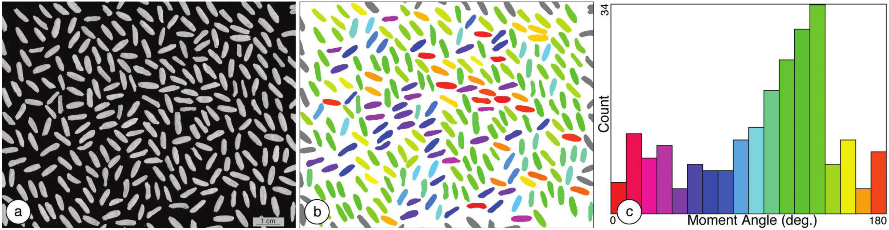


Figure 10.17 Measurement of feature orientation: (a) rice grains; (b) individually color-coded according to moment angle; (c) distribution of grains according to angle.

1. Segment – Otsu will do just fine here
2. Label image
3. Compute moments $\mu_{1,1}$, $\mu_{2,0}$, $\mu_{0,2}$ for each item
4. Compute the orientation item using the moments
5. Plot the angle histogram of the items.

Russ (2016), p. 574

Invariant moments

The following normalized moments are invariant to

- Translation
- Orientation
- Size

$$
\begin{aligned}
m_1 &= \eta_{20} + \eta_{02} \\
m_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11}{}^2 \\
m_3 &= (\eta_{30} - 3\,\eta_{12})^2 + (3\,\eta_{21} - \eta_{03})^2 \\
m_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \\
m_5 &= (\eta_{30} - 3\,\eta_{12})^2(\eta_{30} + \eta_{12})\left[(\eta_{30} + \eta_{12})^2 + 3\,(\eta_{21} + \eta_{03})^2\right] \\
&\quad +(3\,\eta_{21} - \eta_{03})^2(\eta_{21} + \eta_{03})\left[(\eta_{30} + \eta_{12})^2 + 3\,(\eta_{21} + \eta_{03})^2\right] \\
m_6 &= (\eta_{20} - \eta_{02})\left[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2\right] \\
&\quad +4\,\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \\
m_7 &= (3\,\eta_{21} - \eta_{03})^2(\eta_{30} + \eta_{12})\left[(\eta_{30} + \eta_{12})^2 + 3\,(\eta_{21} + \eta_{03})^2\right] \\
&\quad +(3\,\eta_{12} - \eta_{30})^2(\eta_{21} + \eta_{03})\left[(\eta_{30} + \eta_{12})^2 + 3\,(\eta_{21} + \eta_{03})^2\right]
\end{aligned}
$$

Euler’s Sunday walk in Königsberg

Königsberg and graph theory

Question: Is it possible to walk each bridges only once?

Wikipedia

Connectivity analyis

Definition

The Euler-Poincare characteristic $\chi = V - E + F$ combines the number of

- V vertices/nodes
- E edges/connections
- F faces/loops

to quantify the topology and connectivty of a shape.

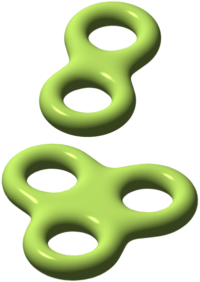
$$\begin{aligned}
\chi_4 &= \mathcal{N}\left(\begin{matrix}\blacksquare & \square \\ \square & \square\end{matrix}\right) + \mathcal{N}\left(\begin{matrix}\blacksquare & \square \\ \square & \blacksquare\end{matrix}\right) - \mathcal{N}\left(\begin{matrix}\blacksquare & \blacksquare \\ \blacksquare & \square\end{matrix}\right) \\
\chi_8 &= \mathcal{N}\left(\begin{matrix}\blacksquare & \square \\ \square & \square\end{matrix}\right) - \mathcal{N}\left(\begin{matrix}\square & \blacksquare \\ \blacksquare & \square\end{matrix}\right)
\end{aligned}$$

The configurations are identified using a hit and miss transform (a morphological operation).
It tells us about the complexity of porous structure.

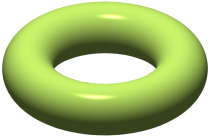
Soille (2002), Vogel (2002)

Some shapes and their EP numbers

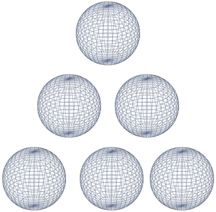
Negative



Zero



Positive



Displacement analysis

Digital image correlation

- Digital image correlation is a method to measure local image displacements.
  - Absolute displacement
  - Direction
- Requires two images for the measurement
- It uses correlation for the analysis.
- Can also be applied to volume data.
- Does not need segmentation to work.

Measure displacement by correlation

Definition

Correlation measures how similar two signals X and Y are, using

$$c(k) = \frac{\mathrm{cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{\sum_{i=1}^{N}(x_{k-i}-\overline{x})(y_i-\overline{y})}{(N-1)\,s_x\,s_y}$$

This process can be implemented more efficiently using the FFT

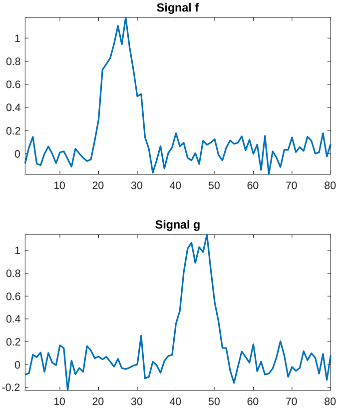
$$c(k) = \mathcal{F}^{-1}\{\mathcal{F}\{x\}\cdot\mathcal{F}\{y\}^*\}$$

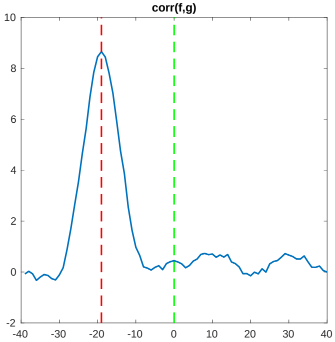
The relative displacement distance is given as the k where $c(k)$ is at maximum.

1D Correlation

Example

Two phase shifted signals *f* and *g*

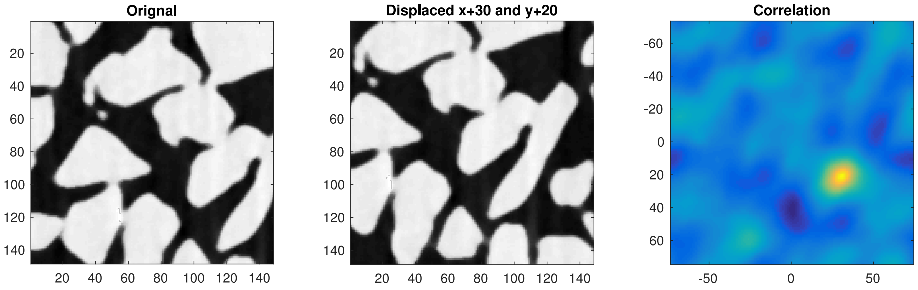




The correlation function computed with FFT gives the displacement distance.

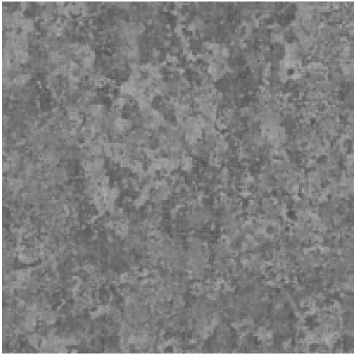
2D Correlation

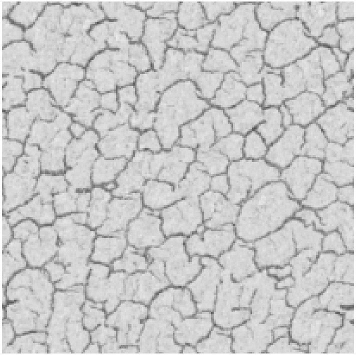
Global correlation



Textures and stress

- A texture is an image pattern often with a noise like appearance.
- It can represent barely resolved sample features.



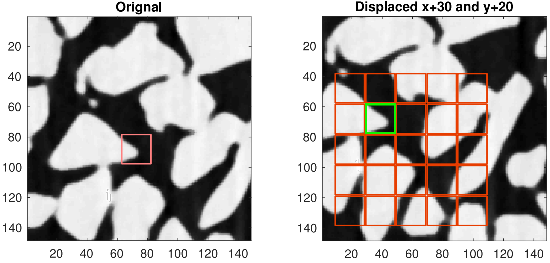


Samples under stress have local texture displacements.

Displacement field

Principle

- The global correlation can only detect if the sample was displaced.
- We want to measure local displacements/changes relative to a reference.
- A displacement field can be computed as correlation with gauge regions.



- The gauge region with best correlation is chosen as local displacement
- A vector field can be generated this way.
- Amplitudes or directions

Displacement field

Example: Water flow in shear band

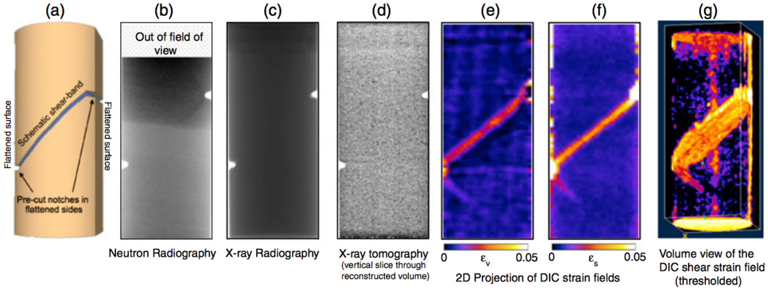


Figure 1. (a) Schematic of the sample geometry (the distance between the flattened faces and the width of the following images are 35 mm). (b) Neutron radiograph of the sample prior to water imbibition; note dark region (higher attenuation) to the top where oil had invaded the sample. (c) X-ray radiography of the sample before water imbibition (note that the slightly darker area at the top is the additional attenuation due to the holder for the sample and not the effect of the oil). (d) Vertical slice though the middle of the X-ray tomography volume (before water imbibition). (e, f) Volumetric and shear strain fields from the volumetric DIC presented as a 2-D median image across the sample parallel to the flattened surfaces. Note that positive volumetric strain indicates compaction and also that some artifacts of the DIC procedure can be seen in the volumetric strain field in the form of subvertical stripes and also subhorizontal stripes emanating from the two notches. (g) Volume rendering of the shear-strain field after thresholding to remove the lower strain values.

Hall (2013)

Classification

Classification – second round

- Last week we used classification methods on gray levels to identify regions.
- Classification methods can also be used on more abstract information:
  - Size
  - Shape
  - Position

  to learn more about the sample.

Identifying relevant metrics

Our first task is to find metrics carrying most information.

Tools in this quest are:

- Single and multivariate histograms
- Statistical methods
  - Correlation $\mathrm{corr}(X, Y)$
  - Mutual information

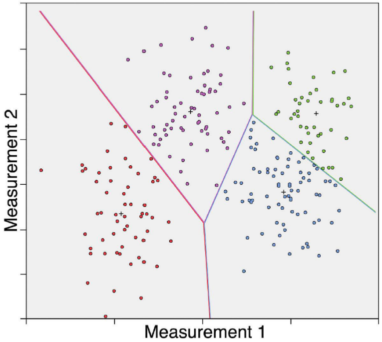
$$I(X;Y) = \sum_{y \in Y} \sum_{x \in X} p(x,y) \log\left(\frac{p(x,y)}{p(x)\,p(y)}\right)$$

$I(X;Y) = 0$ if X and Y are independent. Wikipedia

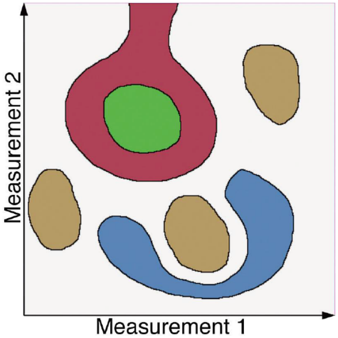
  - Principal compnent analysis.

Types of classification method

Linear separation



Non-linear separation



k-Nearest Neighbors

Principle

Given a training set of observations, select the class for a new observation as the one that have the local majority. The value *k* is the number of voters.

Example

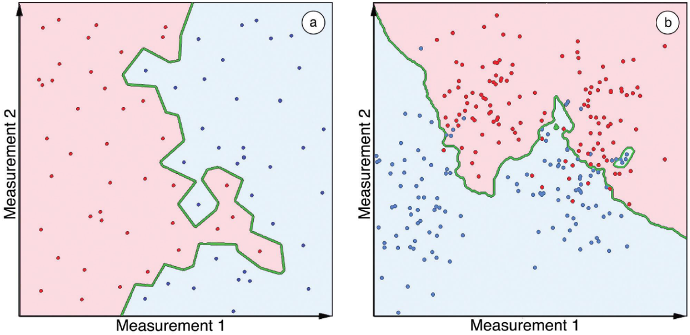


Figure 12.58 *The irregular boundary between two classes using k-nearest-neighbor classification: (a) k = 1; (b) k = 9.*

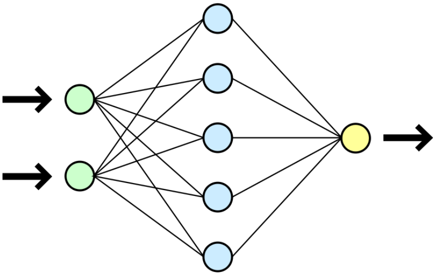
Other non-linear methods

Methods

- Artificial neural networks (ANN)
- Support vector machines
- Self-organizing maps

Require a learning phase to set the weights.

ANN



Haykin (1994), Duda et al. (2001)

Working with machine learning

All machine learning techniques must be trained. . . this requires data.

Training data

All machine learning techniques must be trained to make a prediction. Must be labeled!

Validation data

After training, the trained method must be validated to know the performance. Must be labeled, not same as training data.

Working data

Let it lose on the real problem.

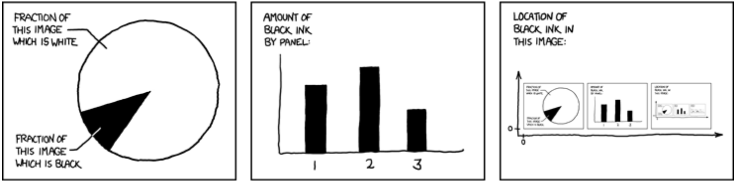
Summary

Summary

We have looked at different ways to quantify information from images

- Size, volume, surface area (Ch 10).
- Shape features (Ch 11).
- Shape guided measurements
  - Distance guided
  - Digital image correlation
- It was not possible to cover all metrics... chapter 10-11 provides more.
- Classification of features (Ch 12).

In the end you want to make statistics



xkcd.com

References I

Duda, R., Hart, P., and Stork, D. (2001). *Pattern classification*. John Wiley & Sons, New York, 2 edition.

Hall, S. (2013). Characterization of fluid flow in a shear band in porous rock using neutron radiography. *Geophysical Research Letters*, 40:2613–2618.

Haykin, S. (1994). *Neural Networks - A comprehensive foundation*. Prentice Hall.

Kaestner, A., Lehmann, P., and Fluehler, H. (2005). Identifying the interface between two sand materials. In *Proc. 5th Int. Conf on 3-D Digital Imaging and Modelling*, pages 410–415. IEEE Computer Science Press.

Russ, J. (2016). *Image processing handbook*. CRC Press.

Soille, P. (2002). *Morphological image analysis*. Springer Verlag, 2nd edition.

van Assen, H., Egmont-Petersen, M., and Reiber, J. (2002). Accurate object localization in gray level images using the center of gravity measure: Accuracy versus precision. *IEEE Transactions on Image Processing*, 11(12):1379–1384.

Vogel, H.-J. (2002). *Topological Characterization of Porous Media*, pages 75–92. Springer Berlin Heidelberg, Berlin, Heidelberg.