

STATISTICS OF IMAGING

Copyright Mats Rudemo, Department of Mathematical Statistics, Chalmers University of Technology, 412 96 Gothenburg, Sweden

E-mail: rudemo@chalmers.se

March 15, 2017

Introduction

The object of these notes is to provide an introduction to several subjects connected with statistical inference from images and spatial data. Image analysis and spatial statistics are sister fields, which both are extensive and grow at considerable speed. Thus only some selected parts can be covered here and the choice of subjects is, of course, heavily influenced by my experience and interests.

The first part “Images” includes a very brief introduction to basic digital image processing, including image acquisition, image filtering and object feature measurements. Pattern recognition, for instance based on features obtained from objects identified in images, is treated at some length. Both the case with known classes, called *discrimination* or *supervised learning* and the case with unknown classes, called *clustering* or *unsupervised learning* are covered. The first part is concluded by a chapter on statistical models for images. One class of models discussed consists of Bayesian models with a Markov random field prior and with observation noise that is pixel-wise independent and identically distributed.

The second part “Spatial Statistics” starts with some basic properties of spatial random processes: covariance properties and prediction (kriging). Spatial point processes are treated in some detail including image models constructed from point processes. The second part is concluded by a brief introduction to shape analysis and the related problems of image warping and image matching.

The third part “Applications” contains examples of image analysis applied to problems in biology, bioinformatics and remote sensing. The examples cover analysis of data from microarray (DNA chip) images, two-dimensional electrophoresis and aerial photographs of forests.

PART 1. IMAGES

1 Digital images

A digital image may be regarded as a matrix of pixels (picture elements), $f = (f_{ij}) = (f_{ij}, i = 1, \dots, m, j = 1, \dots, n)$. Here $f_{ij} \in V$ where V is the set of possible pixel values, e.g. $V = \{0, 1\}$ for a binary image, $V = \{0, \dots, 255\}$ for a grey level image with 256 grey levels, conveniently coded as bytes, and $V = \{0, \dots, 255\}^3$ for a colour image with 256 levels in each of the three colours Red, Green and Blue. Thus each pixel is specified both by a location (i, j) and a pixel value f_{ij} . The first location index i specifies the row and the second index j the column. Rows are counted either from above (most common in the image processing literature) or from below, while columns are counted from the left.

1.1 Examples of images

Example 1.1. *Aerial photographs of a thinning experiment.*

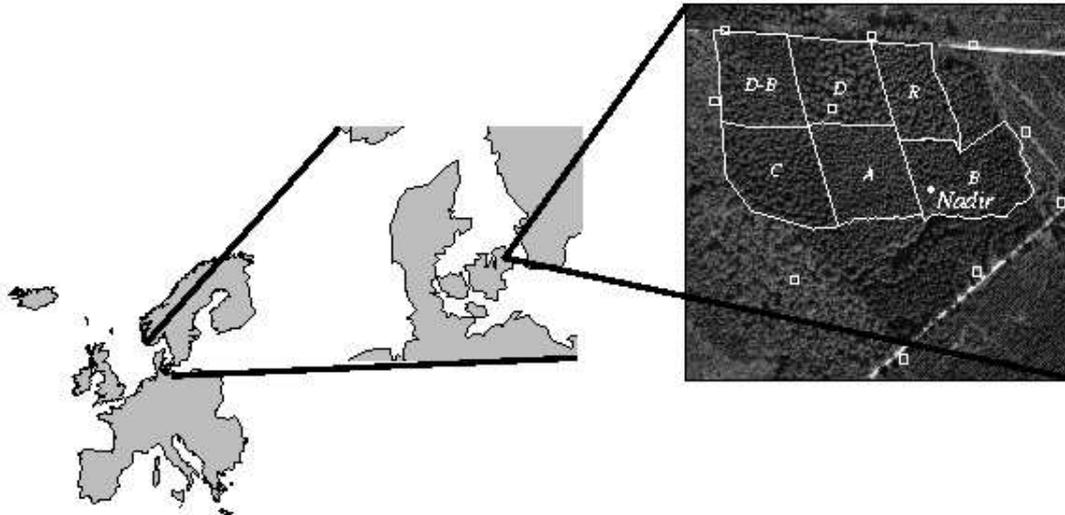


Figure 1: Aerial photograph of the thinning experiment KU in northern Sealand with Norway spruce trees. The position of the airplane at image acquisition was 560 m above “Nadir”.

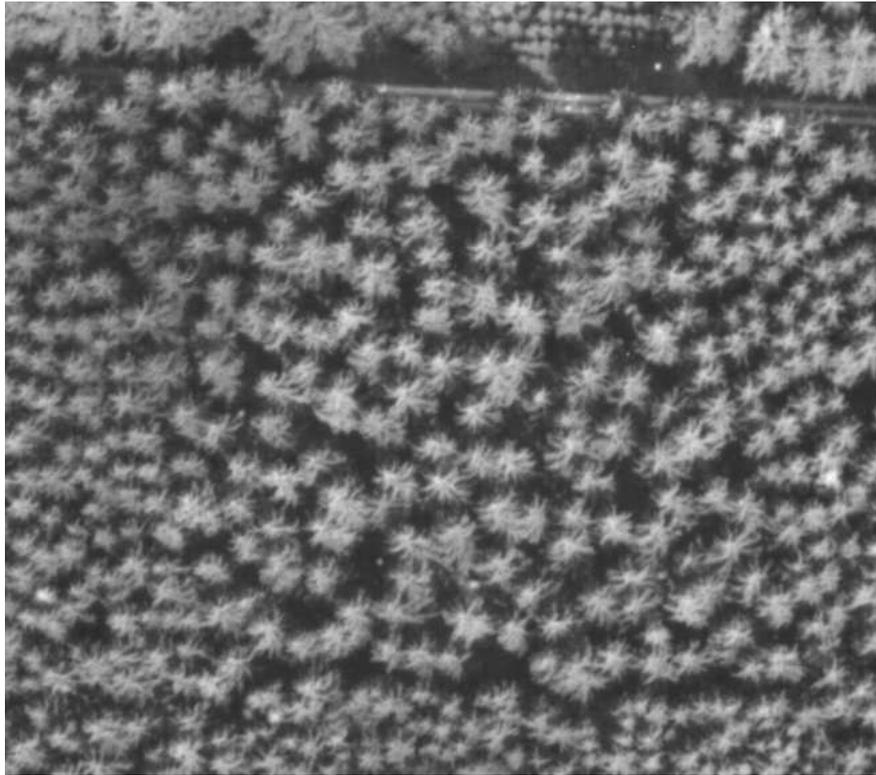


Figure 2: Detail of the aerial photograph in Figure 1 covering the subplot D with very heavy thinning.

Figure 1 shows an aerial photograph of the thinning experiment KU, in northern Sealand, with six subplots which were subject to different thinning treatments (Dralle & Rudemo, 1996). The six treatments were

- A No thinning
- B Light thinning
- C Medium-heavy thinning
- D Very heavy thinning
- D-B In the youth very heavy thinning, later light thinning
- R Heavy row thinning

The photograph was acquired from an airplane at the altitude 560 m above the point "Nadir" in Figure 1. An enlargement of the subplot D is shown Figure 2.

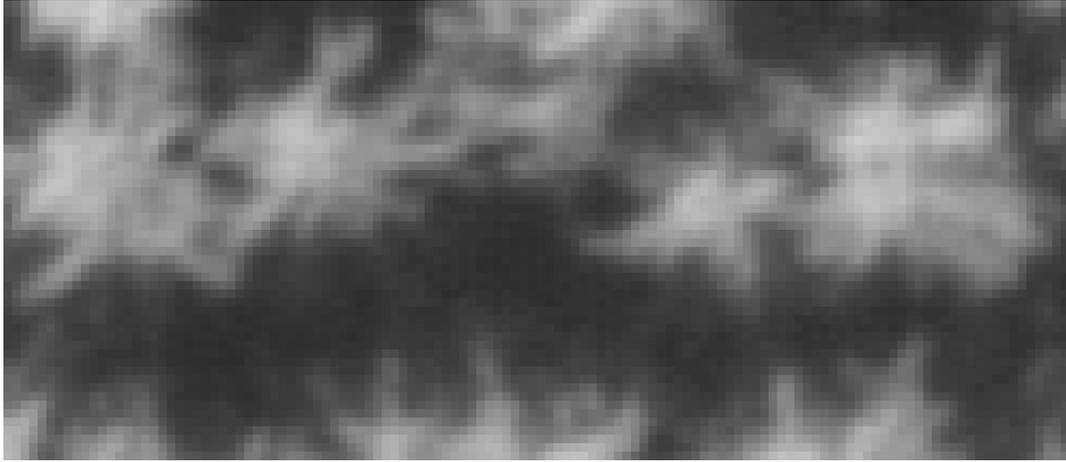


Figure 3: Detail of the aerial photograph in Figure 2 showing part of the southeastern corner of subplot D.

A further enlargement of the southeast corner of subplot D is shown in Figure 3. Here the individual pixels, each corresponding to a square of about $15\text{ cm} \times 15\text{ cm}$ at ground level, are visible.

In Figure 4 we see subplot D from a photograph acquired with the airplane in a position to the northwest of the experimental area. The time of acquisition was August 4 at 10:08 AM, which implies that the sun was in the direction southeast, and the trees were thus backlit in Figure 4.

One object of the image analysis of the photographs obtained in this experiment was to estimate the number of trees in the different subplots and to estimate the positions of the tree tops. This application is further discussed in Chapter 10 in Part 3.

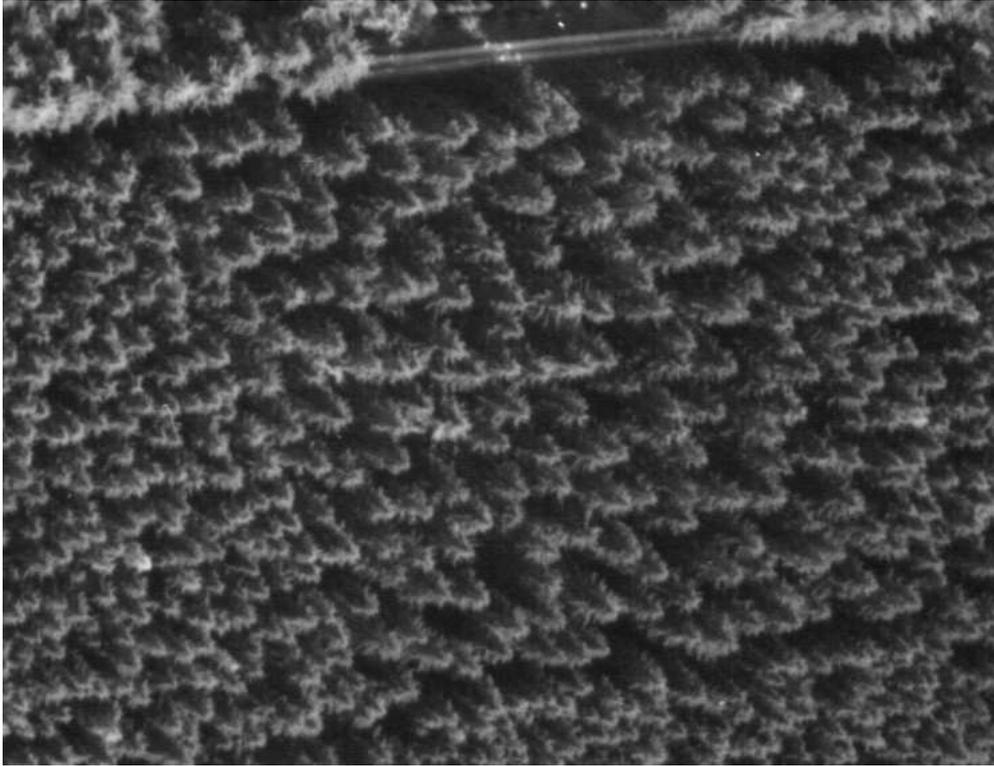


Figure 4: Detail of aerial photograph of subplot D of backlit Norway spruce trees acquired from an oblique angle with the airplane located to the northwest of the experimental area shown in Figure 1.

□

Example 1.2. *Weed seeds.*

Figures 5 and 6 show images of 25 seeds of each of two weed species: curly dock, *Rumex crispus*, and thyrses sorrel, *Rumex thyrsiflorus*. The images were obtained in the study Petersen (1992), where seeds from 40 weed species were studied. The object was to find features from images of the weed seeds which enable recognition of the individual species. Problems of this type will be discussed in Chapter 2 on pattern recognition.

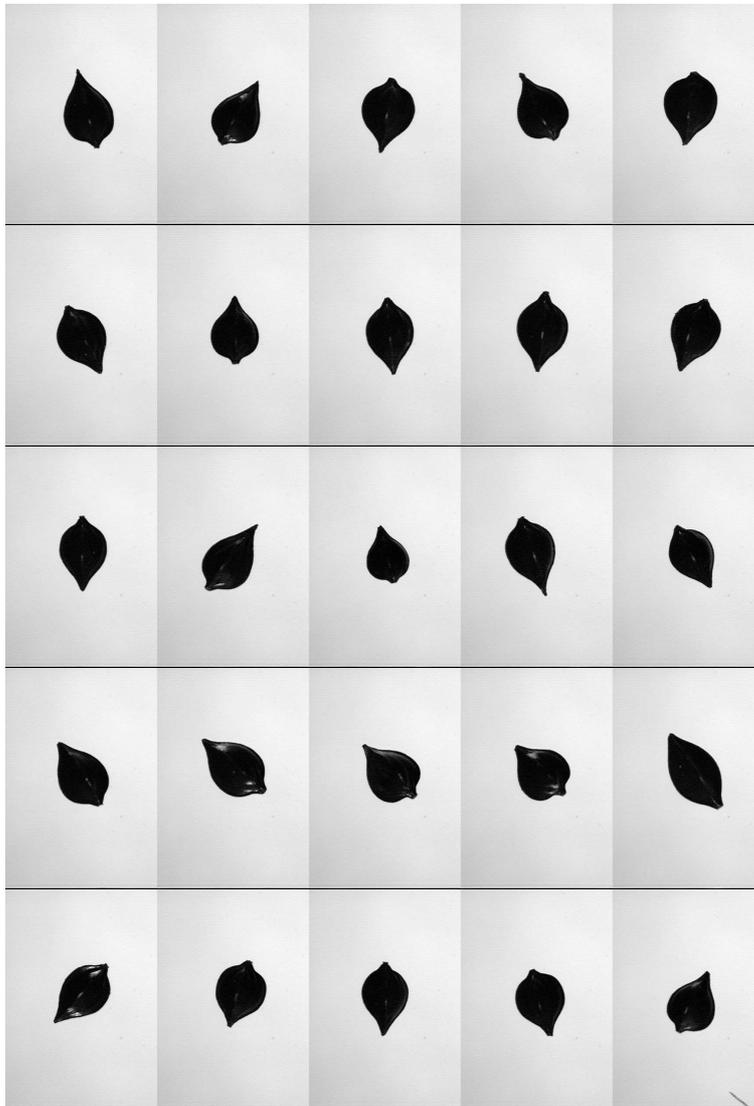


Figure 5: Images of seeds of *Rumex crispus*.

In Figures 5 and 6 we see varying orientations and sizes of the seeds but also some additional variation in the form of the contours. An important problem for series of images of this type, in addition to the previously mentioned pattern recognition, is to estimate some kind of average shape of a seed from a given species, and also to quantify in terms of statistical distributions the probable deviations from this average shape. In Chapter 7 on image warping and image matching such problems will be treated.

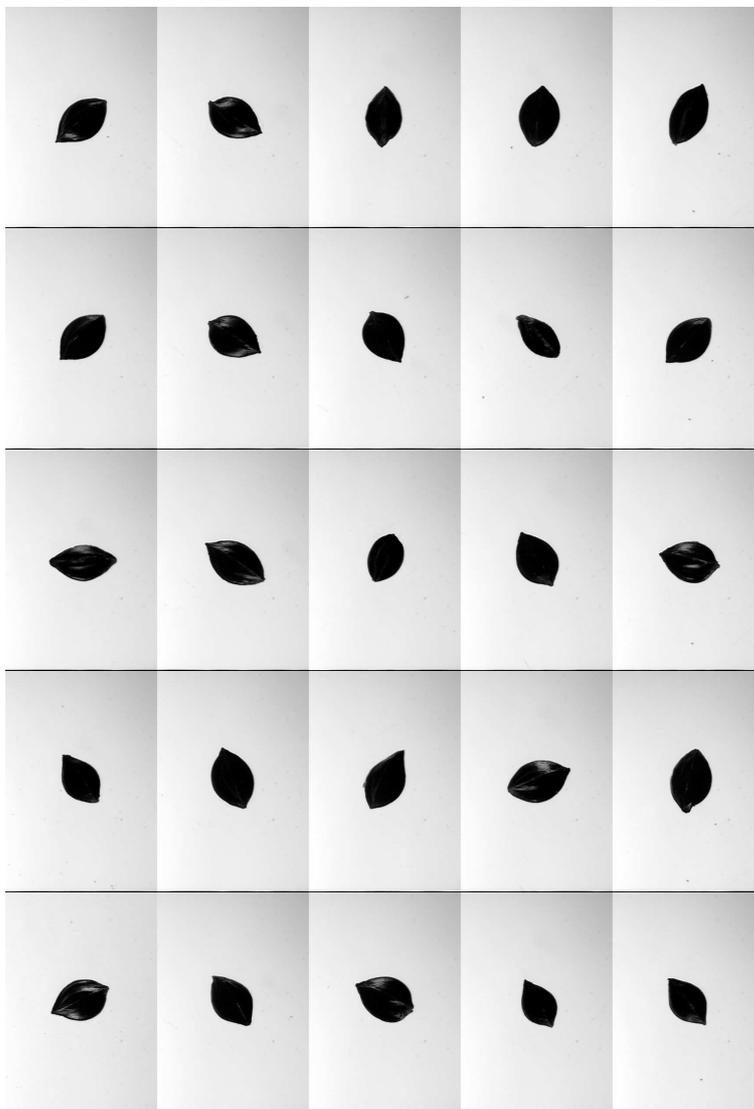


Figure 6: Images of seeds of *Rumex thyrsoiflorus*.

□

Example 1.3. *Weed plants at an early stage.*

Mattias Andersson (1998) studied weed and crop classification using a dataset with 27 images from each of 8 species: carrot, *Daucus carota*, which was the crop, and 7 weed species. Figure 7 shows photographs of two carrot plants and two ladythumb smartweed plants. Similarly, Figure 8 shows photographs of two fumitory plants and two corn spurry plants.

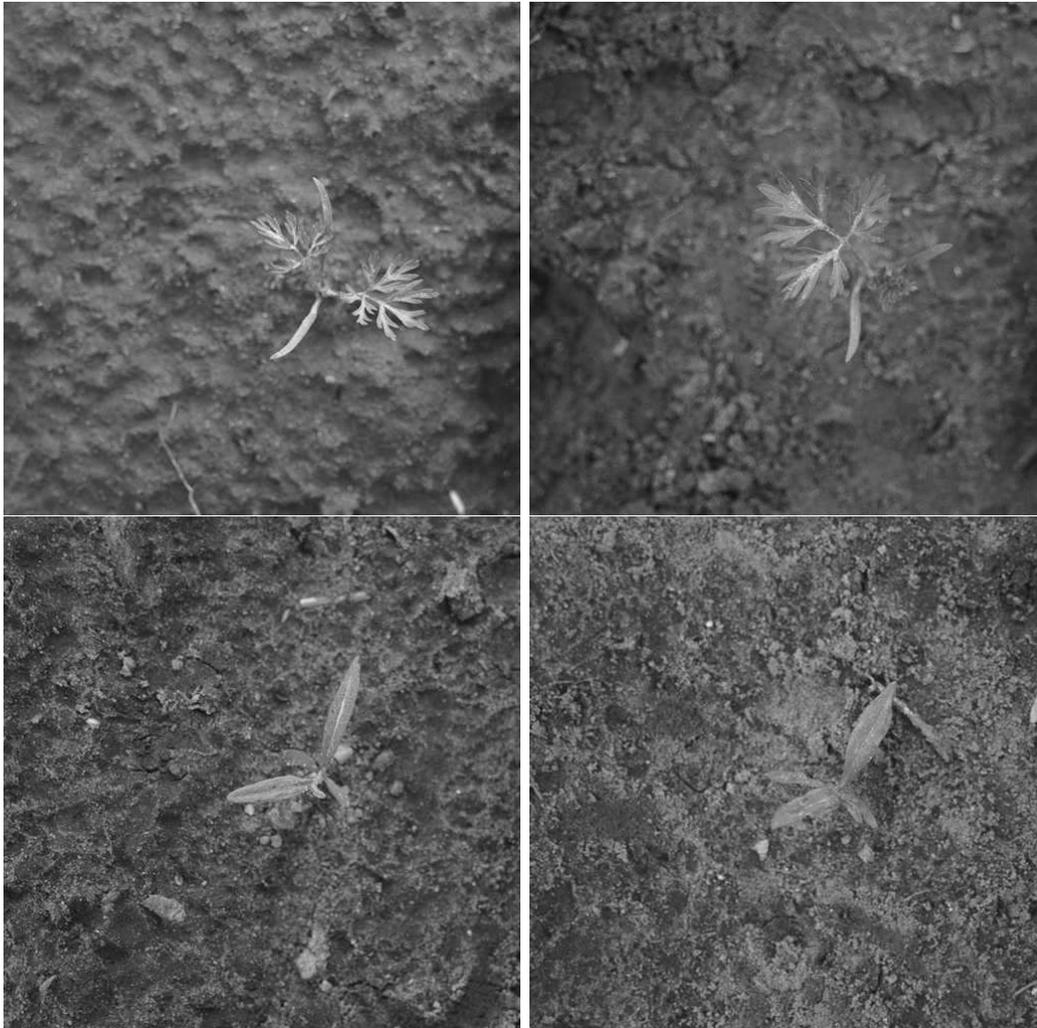


Figure 7: Above two images of plants of carrot, *Daucus carota*, L., and below two images of plants of ladythumb smartweed *Polygonum persicaria*, L.

The images were obtained with a Canon EOS500N still camera with a 80 mm zoom lens and mounted on a tri-pod pointing directly towards ground. The images obtained were in colour, although they are shown as grey-level images in Figures 7 and 8. The corresponding colour images may be obtained from

<http://www.math.chalmers.se/~rudemo/Images/WeedPlants/WeedPlants.html>

The number of pixels of the images was originally 512×768 but was reduced to 512×512 by cutting. The pixel width corresponds to 0.195 mm at ground level.

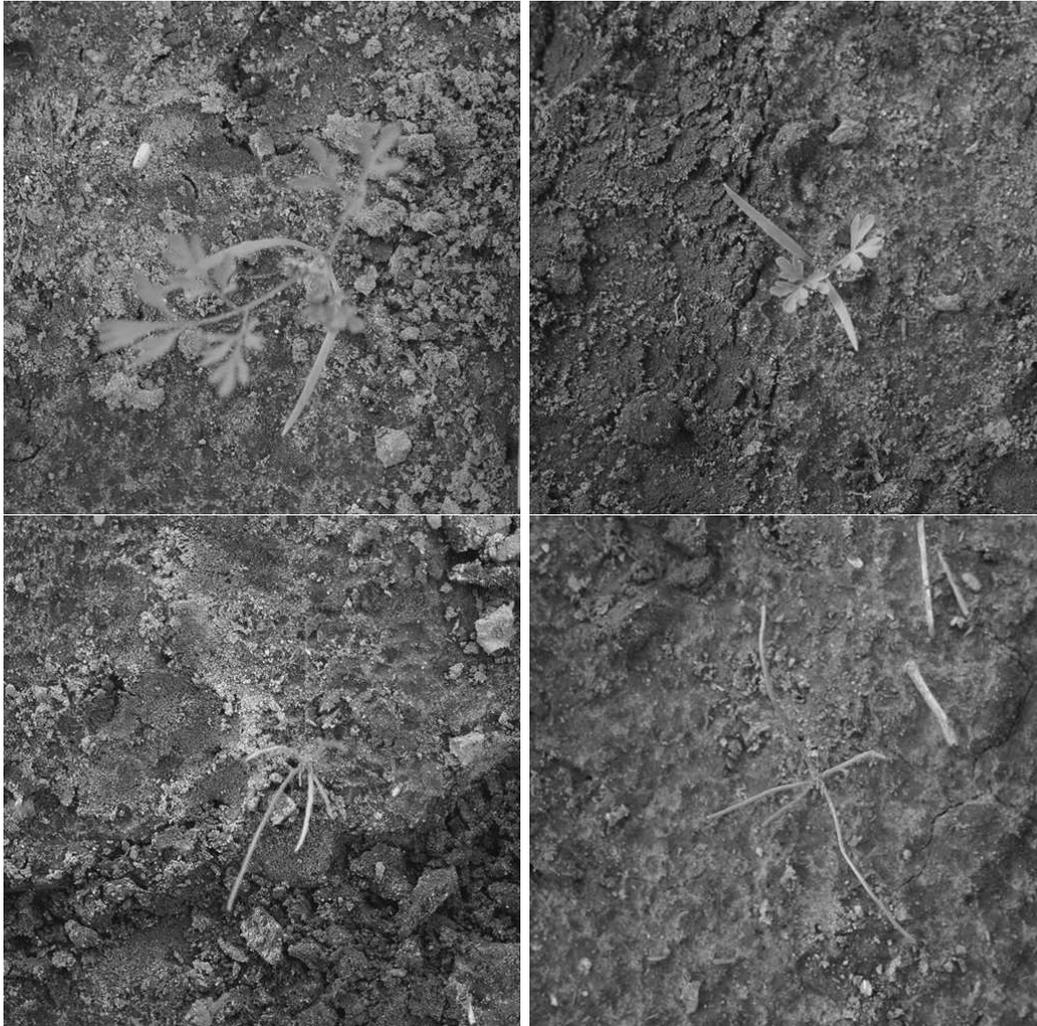


Figure 8: Above two images of plants of fumitory, *Fumaria officinalis*, L., and below two images of plants of corn spurry, *Spergula arvensis*, L.

□

Example 1.4. *Two-dimensional electrophoresis images.*

Yeasts are uni-cellular fungi which reproduce rapidly and thus are highly suitable as model systems for more complicated eucaryotic species such as mammals. In particular, the genome of baker's yeast, *Saccharomyces cerevisiae*, was fully sequenced by Goffeau et al. (1996).

Figures 9 and 10 show four images from an experiment with baker's yeast and two treatments corresponding to growth under normal conditions and growth under stress with salt added to the nutrition solution, see Alipour et al. (2000) and Gustafsson et al. (2002). In the experiment there were five repetitions both for the standard treatment, corresponding to growth in a standard solution, and the treatment with growth under salt stress, which in this experiment corresponds to growth in a 1 M sodium chloride solution. Figure 9 shows the images obtained from two repetitions with the standard treatment and Figure 10 shows images from two repetitions with salt added.

Each spot in a 2D electrophoresis image such as in Figures 9 and 10 corresponds to a protein with a specific isoelectric point (pI) determined by isoelectric focusing in the horizontal direction as a first step and a specific molecular weight determined by vertical separation in a second step. For instance, under ideal conditions the protein molecules perform in the second

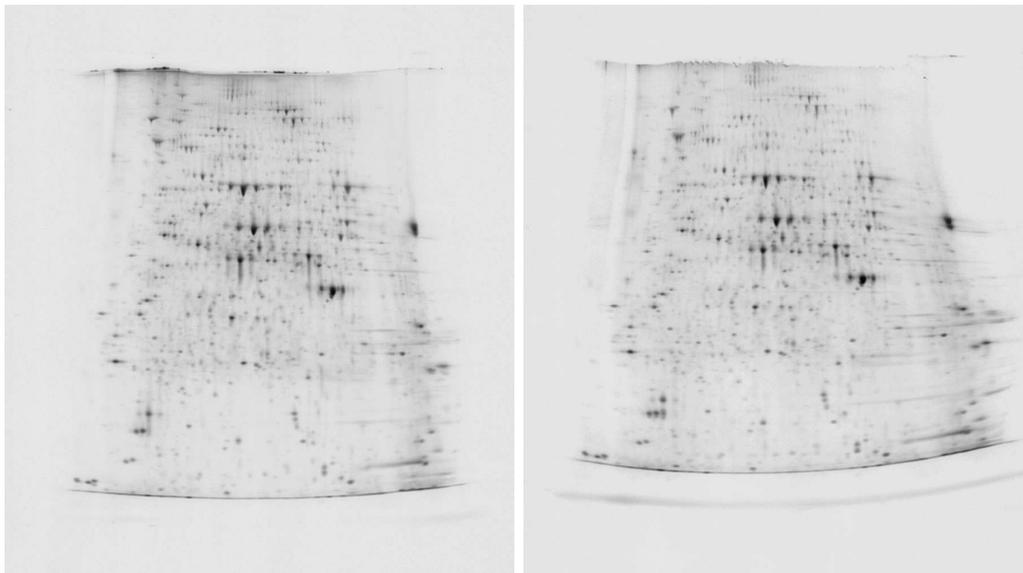


Figure 9: Images from 2D gel electrophoresis of baker's yeast grown in a standard solution.

step a vertical Brownian motion with drift from a starting position at the top such that small molecules travel a longer way than large molecules. Typically one can separate proteins in the pH range, or more precisely the pI range, 4–7 and with molecular weights in the range 5–200 kDa. Under favourable conditions thousands of proteins may thus be resolved, and the size of a spot in the electrophoresis image is a measure of the level of the corresponding protein.

The basic problem in an experiment such as the one described with yeast grown under standard conditions and under salt stress is to find those proteins that are upregulated and those that are downregulated under stress. As a first step we need to find those spots in the four images in Figures 9 and 10 that correspond to each other, that is, which measure the same protein. This is called matching of the images and may be performed by a warping of images onto each other. It is clear from an inspection of the two images in Figure 9, and similarly the two images in Figure 10, that also for experimental units that have received the same treatment the locations of spots corresponding to one protein can vary considerably due to random variation. And this random variation seems to be more complicated than the variation corresponding to a Brownian motion as referred to above.

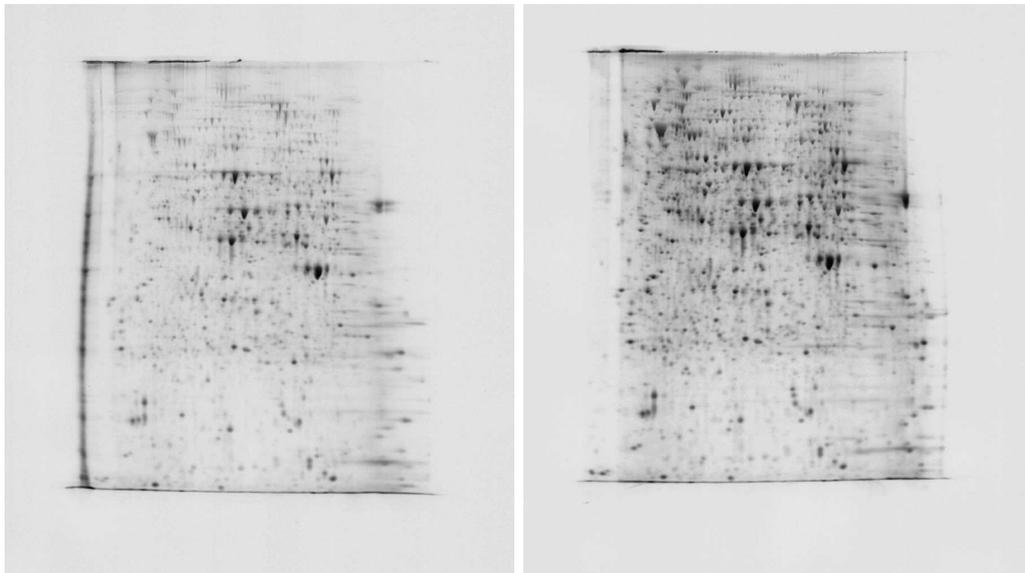


Figure 10: Images from 2D gel electrophoresis of baker's yeast grown under stress in a solution with salt added.

□

Example 1.5. *Two-colour spotted microarrays.*

In microarray analyses the expression level of thousands of genes can be estimated simultaneously. In two-colour spotted microarray analysis DNA fragments corresponding to different genes are typically arrayed on glass slides in spots with a diameter of the order $100\ \mu\text{m}$.

Gray scale image, 020725cy53x8l30g40avg4.tif, log-transformed Gray scale image, 020725cy3wtl30g40avg4.tif, log-transformed

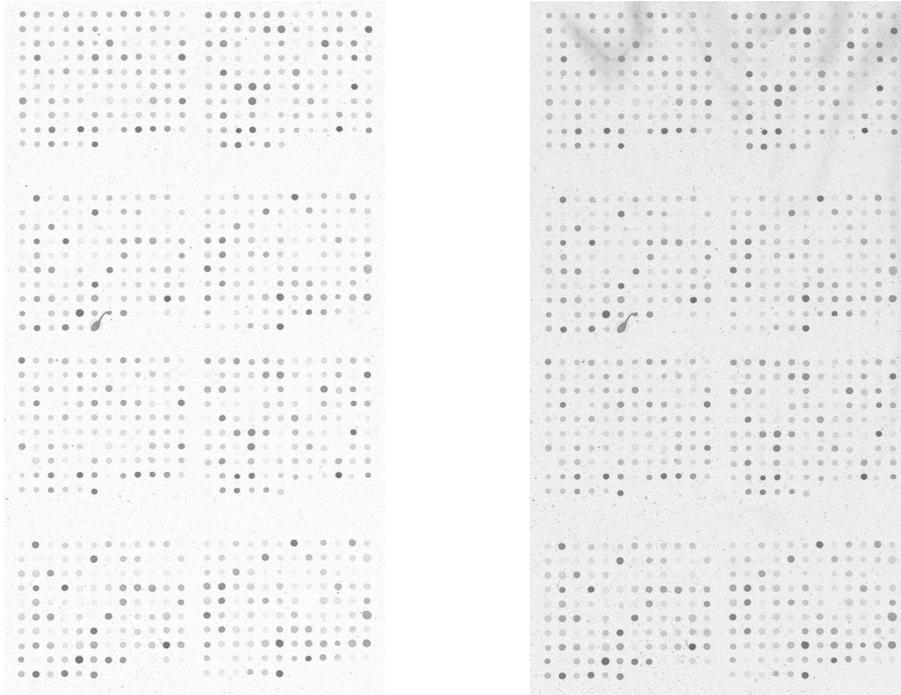
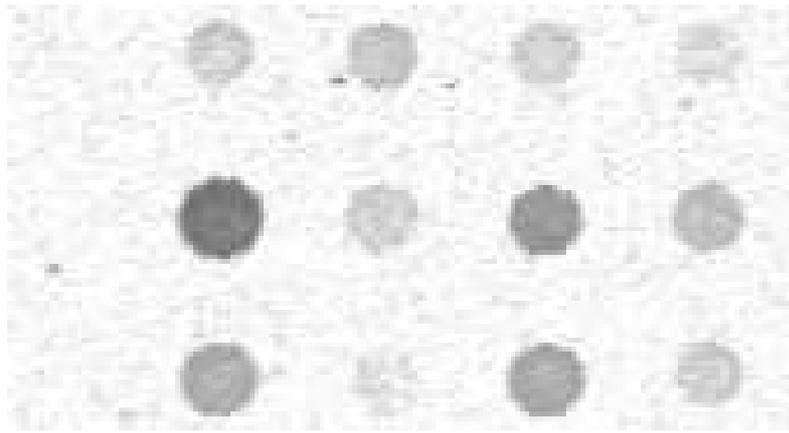


Figure 11: Images from an experiment with two varieties of *Arabidopsis*, Cy5 channel (left) for a transgenic line and Cy3 channel (right) for the wild-type in a two-colour spotted microarray experiment with 452 genes. The upper half with 20 rows contains all the 452 genes and the lower half is a repetition of the upper half. The images are shown inverted (high intensity shown as black) and a logarithmic scale transformation of intensities is also used.

Complementary DNA (cDNA) is synthesized from two sources of RNA of different origins and labeled with different fluorescent dyes, for instance, one with the green dye Cy3 and the other with the red dye Cy5. The pools of labeled cDNA are mixed together and allowed to hybridize with the DNA fragments in the different spots on the glass slide. The slide is illuminated with two laser light sources exciting the two fluorescent dyes and the intensity of emitted fluorescent light is measured at two suitably chosen wavelengths.

Figure 11 shows grey-level images for the two channels of one array in an experiment comparing RNA from two varieties of *Arabidopsis* plants, transgenic line 3x8 and wild-type wt (Kristensen et al. 2005). For clarity of display the images are shown inverted, that is black corresponds to high intensity levels and before inversion a logarithmic transformation is also used. Data transformations and spot shape models for spotted microarrays are discussed in (Ekstrøm et al. 2004) and applied to data from this experiment.

Gray scale image, 020725cy53x8l30g40avg4.tif



Gray scale image, 020725cy3wtl30g40avg4.tif

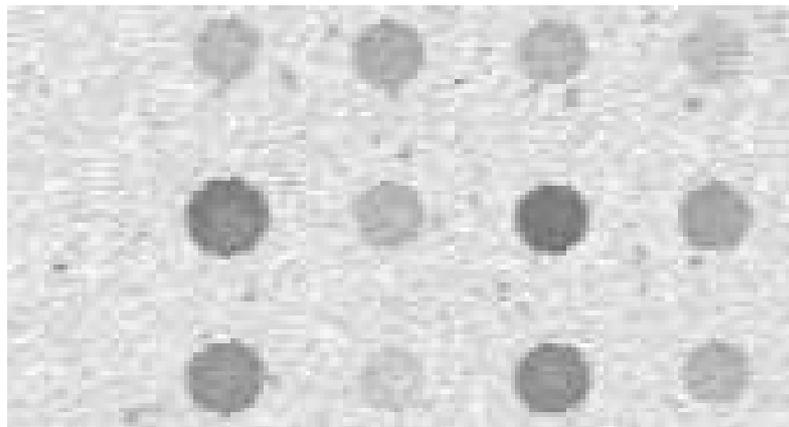


Figure 12: Blow-up of rows 6–8 and columns 1–4 in Figure 11 with the Cy5 channel for the transgenic line above and the Cy3 channel for the wild-type below.

Figure 12 shows a blow-up with 3 rows and 4 columns for both channels. One crucial question analysed in experiments of this type is to find out which genes that are differentially expressed, that is show significantly higher or lower intensities. In this experiment it turned out that remarkably few genes in the transgenic line were affected in the comparison with the wild-type. One of the few genes affected was the gene that corresponds to the first spot in the middle row in Figure 12. As indicated in the figure it was upregulated in the transgenic line. However, random errors are large in this type of experiments and typically one needs to repeat the experiment for several slides and make a subsequent statistical analysis of the results, cf. Chapter 8. □

Example 1.6. *Diffusing particles.*

Colloidal particles in a suspension perform random motion essentially as a three-dimensional Brownian motion with the diffusion coefficient as a crucial parameter. However, as the particles come close they interact and this interaction may be described by an interaction potential.

A series of images were obtained by video microscopy, see Kvarnström (2000a,2002b), in a joint project with Lennart Lindfors, AstraZeneca, Mölndal. The object in this project was to estimate the diffusion coefficient and, if possible, also the particle interaction potential.

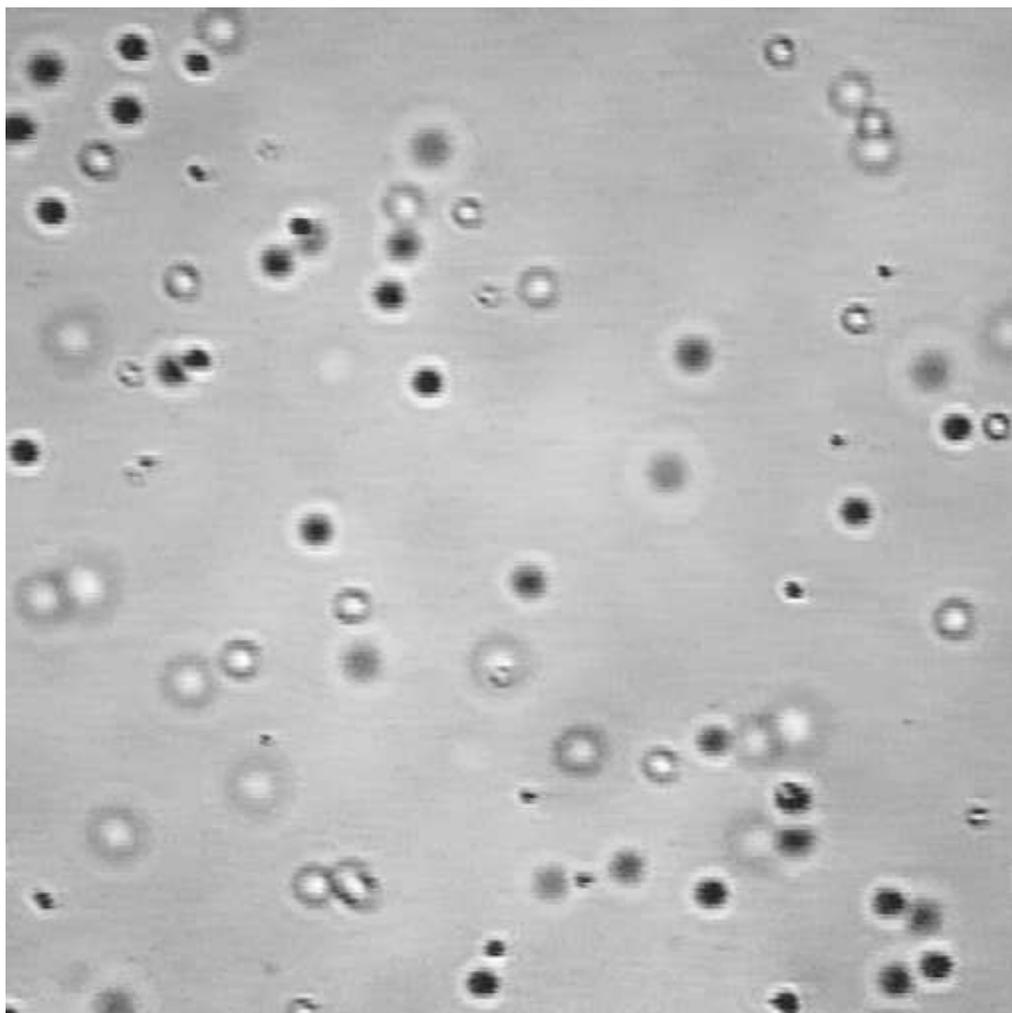


Figure 13: Image obtained by video microscopy showing diffusing particles. Particles in phocus are shown as small distinct black objects.

Images of the diffusing particles were obtained with a time interval of 0.02 seconds between images, and two consecutive images are shown

shown as small distinct black objects, while particles out of phocus are extended, the degree of extension depending on the distance to the phocal plane. An object corre-

sponding to a particle out of phocus is further either white or black in its central part corresponding to the particle being above or below phocus, respectively.

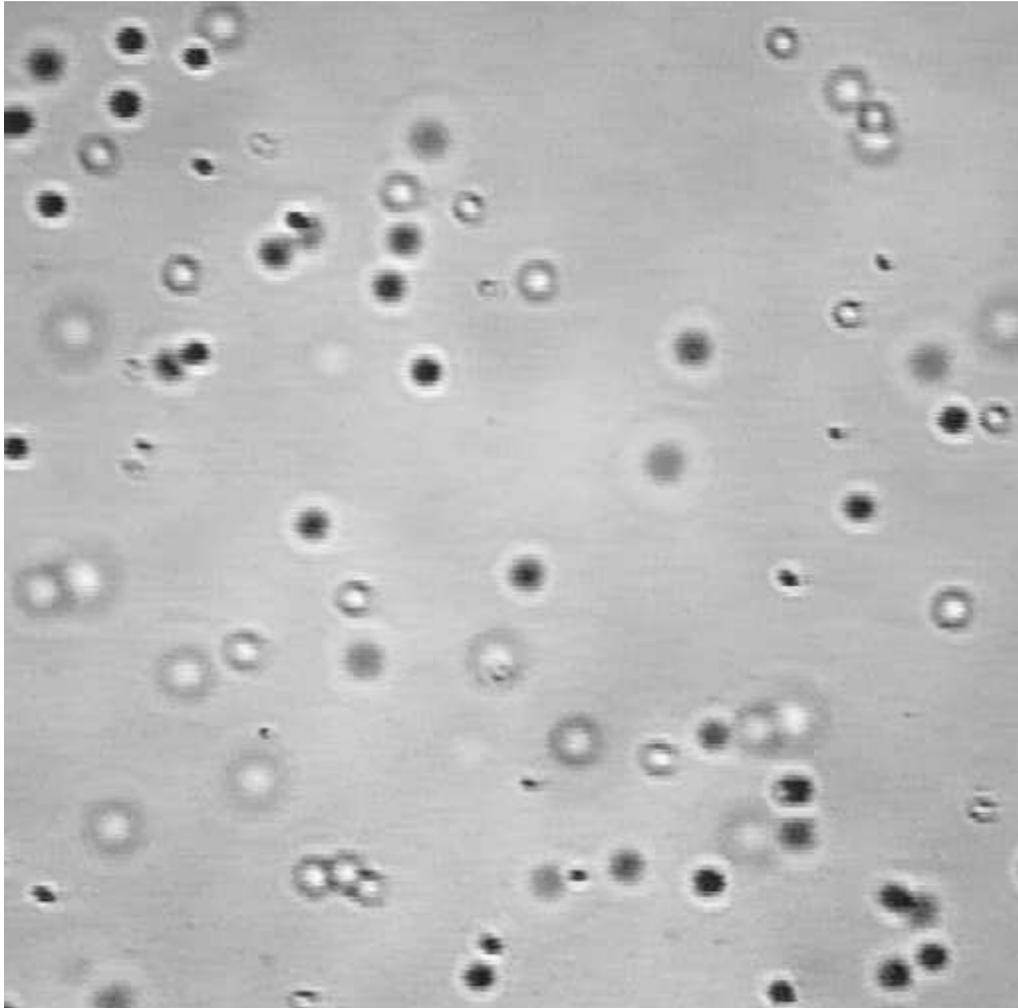


Figure 14: Image obtained by video microscopy showing diffusing particles. This image was obtained 0.2 seconds after the image in Figure 13.

□

Example 1.7. *Handwritten digits.*

The MNIST database of handwritten images consists of a training set with 60000 digits and an evaluation set of 10000 digits, see

<http://yann.lecun.com/exdb/mnist/>

Examples of images from this set is given in Figure 15, actually the first 100 digits from the training set. The digit images are 28x28 pixel grey level images obtained from 20x20 pixel binary black and white images. The MNIST dataset has been used extensively as a proving ground for pattern recognition methods and it will also be used in these notes in Chapter 2.

5	0	4	1	9	2	1	3	1	4
3	5	3	6	1	7	2	8	6	9
4	0	9	1	1	2	4	3	2	7
3	8	6	9	0	5	6	0	7	6
1	8	7	9	3	9	8	5	9	3
3	0	7	4	9	8	0	9	4	1
4	4	6	0	4	5	6	1	0	0
1	7	1	6	3	0	2	1	1	7
8	0	2	6	7	8	3	9	0	4
6	7	4	6	8	0	7	8	3	1

Figure 15: Examples of 100 handwritten digits from the MNIST database.

□

1.2 Image filtering

Let $w = (w_{k,l}) = (w_{k,l}, k = -p, -p+1, \dots, p, l = -p, -p+1, \dots, p)$ be a matrix of real numbers. A new image g may be constructed from a given image f by linear filtering,

$$g_{ij} = \sum_{k=-p}^p \sum_{l=-p}^p w_{k,l} f_{i+k, j+l}. \quad (1)$$

A simple filter example is a 3×3 averaging filter

$$w = \begin{bmatrix} w_{-1,-1} & w_{-1,0} & w_{-1,1} \\ w_{0,-1} & w_{0,0} & w_{0,1} \\ w_{1,-1} & w_{1,0} & w_{1,1} \end{bmatrix} = \frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}. \quad (2)$$

A more smooth averaging filter is obtained by use of circular 2D Gaussian filter with a variance parameter σ^2 ,

$$w_{k,l} = c \exp\left(-\frac{1}{2\sigma^2}(k^2 + l^2)\right), \quad (3)$$

where c is chosen such that

$$\sum_{k=-p}^p \sum_{l=-p}^p w_{k,l} = 1, \quad (4)$$

and p is chosen so that $w_{k,l}$ is small outside the region determined by $|k| \leq p$ and $|l| \leq p$. Chose, for instance, p to be the smallest integer which is at least as large as 3σ .

Care has to be taken in (1) when the indices in the summation fall outside the original image. One possibility is to restrict the filtering to those pairs (i, j) for which all indices $i+k$ and $j+l$ in (1) fall inside the image f , another possibility is to extend the original image in a suitable way, and a third possibility is to modify the filter close to the image edges.

The averaging filter (2) is relatively vulnerable to large errors in individual pixels. A more robust filter is the nonlinear *median* filter which for 3×3 neighbourhood is given by

$$g_{ij} = \text{median}\{f_{i+k, j+l} : |k| \leq p, |l| \leq p\} \quad (5)$$

with $p = 1$. Here $\text{median}(A)$ denotes the median for a finite set A of real numbers.

Image filtering can also be used to emphasize edges. Thus a linear filter with

$$w = \begin{bmatrix} w_{-1,-1} & w_{-1,0} & w_{-1,1} \\ w_{0,-1} & w_{0,0} & w_{0,1} \\ w_{1,-1} & w_{1,0} & w_{1,1} \end{bmatrix} = \frac{1}{6} \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}. \quad (6)$$

will tend to emphasize vertical edges, and similarly the filter

$$w = \begin{bmatrix} w_{-1,-1} & w_{-1,0} & w_{-1,1} \\ w_{0,-1} & w_{0,0} & w_{0,1} \\ w_{1,-1} & w_{1,0} & w_{1,1} \end{bmatrix} = \frac{1}{6} \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}. \quad (7)$$

will tend to emphasize horizontal edges.

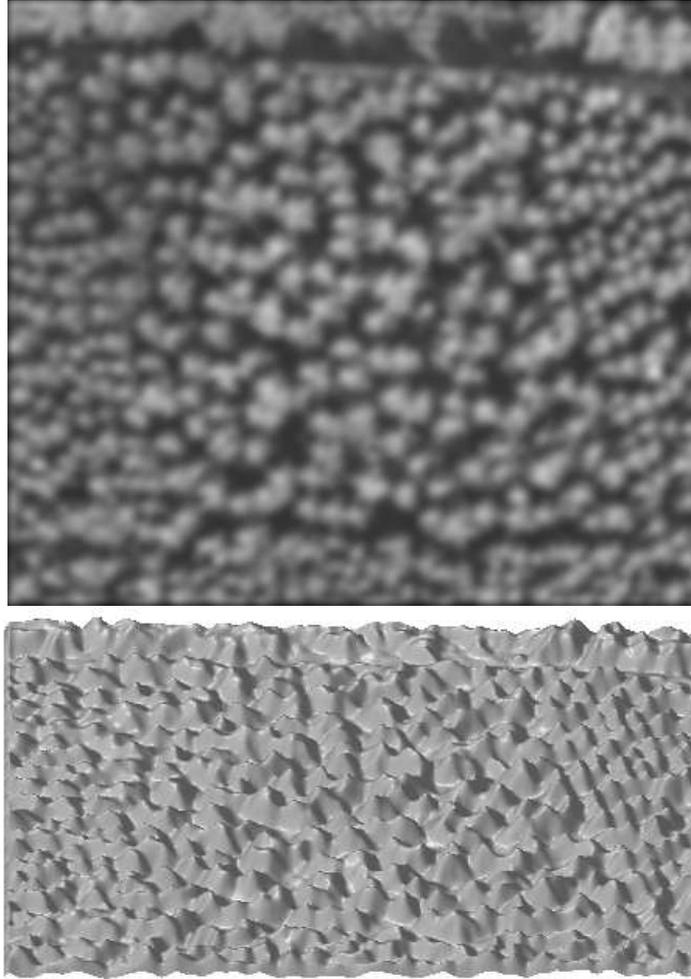


Figure 16: Upper part: Smoothed version of the image in Figure 2 by use of circular 2D Gaussian filter with $\sigma = 4.5$ pixel-widths. Lower part: The same image viewn in perspective as a 3D surface with light intensity as the vertical coordinate.

Example 1.8. *Aerial photographs of a thinning experiment. Continuation.*

Let us smooth the image in Figure 2 by use of a circular 2D Gaussian filter with a suitably chosen parameter σ to see if we can estimate the locations of the trees as 'whiteness' maxima in the smoothed image. With $\sigma = 4.5$ we find the image in Figure 16.

From Figure 16 and Figure 2 we see that maxima in the smoothed image seem to correspond well to the location of the trees. This is also indicated by Figure 17 which shows the locations of the maxima of the smoothed image (Here we have only included maxima which have a distance from the nearest edge which exceeds 3σ .)

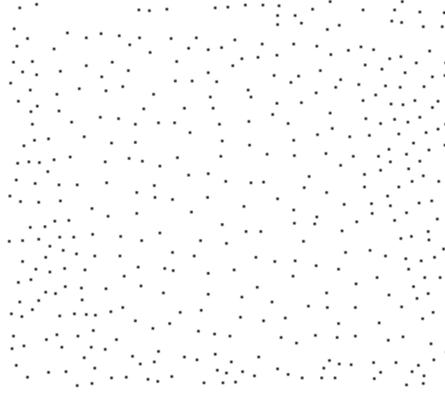


Figure 17: Location of maxima in Figure 16.

□

1.3 Histograms, thresholding and segmentation

An important characteristic of an image is its histogram. For a grey scale image, $f = (f_{ij}) = (f_{ij}, i = 1, \dots, m, j = 1, \dots, n)$, where $f_{ij} \in V$ with V as a set of real numbers, the histogram is defined as

$$h_k = \text{card}(\{(i, j) : f_{ij} \in I_k\}), \quad k = 1, \dots, K, \quad (8)$$

where $\text{card}(A)$ denotes the number of elements in the set A and $\{I_1, \dots, I_K\}$ is a set of disjoint intervals with V as their union.

If an image consists of two parts with grey levels that do not overlap too much the histogram can be used to find a threshold level t which enables us to divide the image into two segments corresponding to these parts. Thus we can define a binary image $b = (b_{ij})$ with two levels, 0 and 1, by putting

$$b_{ij} = \begin{cases} 0 & \text{if } f_{ij} \leq t \\ 1 & \text{if } f_{ij} > t. \end{cases} \quad (9)$$

Segmentation by use of a threshold level found by inspection of the histogram of an image is illustrated in the following example.

Example 1.9. *Weed seeds. Continuation.*

In the upper part Figure 18 we see one of the seeds from Figure 5, actually the seed in the lower left corner rotated 90 degrees. In the lower part of the figure we see the corresponding histogram.

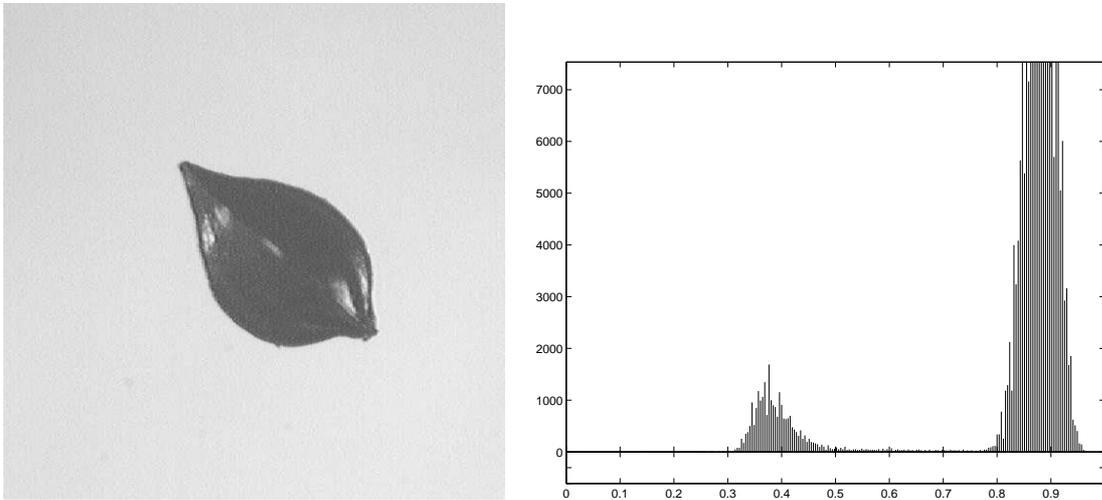


Figure 18: Above an image of a *Rumex crispus* seed and below the corresponding histogram.

It seems clear that a threshold level somewhere between $t = 0.5$ and $t = 0.8$ would be suitable. In Figure 19 we see segmentations with the levels $t = 0.5$, upper left, $t = 0.8$, upper right, and $t = 0.65$, lower left. In the lower right part of the image we see a segmentation obtained from the lower left image by filling out the white “holes”, an operation that can be performed in several ways.

□

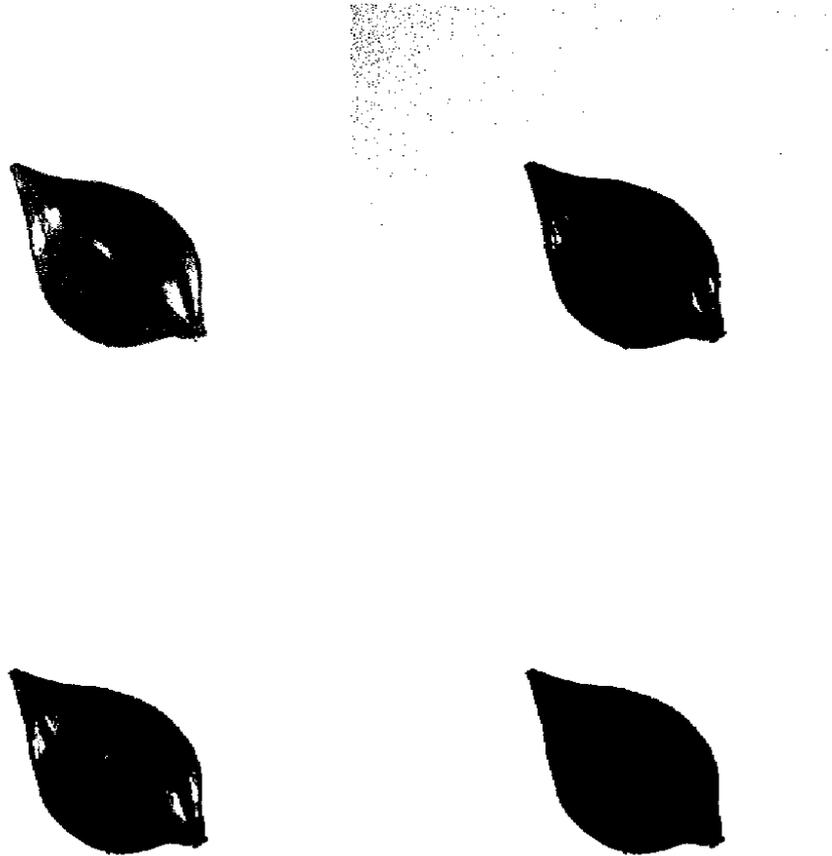


Figure 19: Binary images obtained by thresholding of the image in Figure 18 with the levels $t = 0.5$ (upper left), $t = 0.8$ (upper right), and $t = 0.65$ (lower left). The lower right image is obtained from the lower left image by filling out holes.

1.3.1 Segmentation by a normal mixture model

In many cases, as in Example 1.9 with a bimodal histogram it is fairly easy to separate components in a mixture. We will now describe a normal mixture model which can be used to get a precise threshold value and which also can be used in cases where there are not two modes in the histogram but one component only shows up as a prolonged tail. We suppose that the sets I_k in (8) consist of consecutive intervals with midpoints x_k and equal lengths Δ . Let $\phi(x) = (1/\sqrt{2\pi}) \exp(-x^2/2)$ and put

$$f(x; p_1, \mu_1, \sigma_1, \mu_2, \sigma_2) = \frac{p_1}{\sigma_1} \phi((x - \mu_1)/\sigma_1) + \frac{(1 - p_1)}{\sigma_2} \phi((x - \mu_2)/\sigma_2). \quad (10)$$

We note that $f(\cdot; p_1, \mu_1, \sigma_1, \mu_2, \sigma_2)$ integrates to one, and if the interval length Δ is small we should have

$$\Delta \sum_k f(x_k; p_1, \mu_1, \sigma_1, \mu_2, \sigma_2) \approx 1. \quad (11)$$

Let $n = \sum_k h_k$ denote the total number of pixels and assume that

$$h_k \approx n \Delta f(x_k; p_1, \mu_1, \sigma_1, \mu_2, \sigma_2). \quad (12)$$

To estimate the parameters $p_1, \mu_1, \sigma_1, \mu_2, \sigma_2$ we minimize

$$Q(p_1, \mu_1, \sigma_1, \mu_2, \sigma_2) = \sum_k (h_k - n \Delta f(x_k; p_1, \mu_1, \sigma_1, \mu_2, \sigma_2))^2. \quad (13)$$

Example 1.10. *Weed plants at an early stage. Continuation*

In the upper left part of Figure 20 we see the grey level image of a weed plant. The original a image is colour a image with three channels, blue, green and red. To separate plant pixels from soil pixels we first regard the green channel which is shown in the upper right part of Figure 20. To improve the separation of plant and soil pixels we consider the normalized green colour, which for pixel (i, j) has the pixel value

$$g_{ij} = \text{Round}(255G_{ij} / (B_{ij} + G_{ij} + R_{ij}) + 1), \quad (14)$$

where B_{ij} , G_{ij} and R_{ij} are the blue, green and red channel values for the colour image, and $\text{Round}(\cdot)$ denotes rounding to the nearest integer. The normalized green image is shown in the lower left part of Figure 20. The histogram for the normalized green channel is shown in the left part of Figure 21. Can you suggest why it is useful to normalize the green channel before computing the histogram? Now we fit the normal mixture model given by (10) and (12) for the normalized green channel by minimizing Q in (13) with the restriction $\mu_1 > \mu_2$. Thus the first component should correspond to plant pixels. Let $\hat{p}_1, \hat{\mu}_1, \hat{\sigma}_1, \hat{\mu}_2, \hat{\sigma}_2$ denote the estimated parameters. In Figure 21 we show the histogram and the fitted normal components.

To segment an images we could then choose to consider a pixel (i, j) as a plant pixel if $g_{ij} > T$, where the threshold \hat{T} is obtained by solving the equation

$$\frac{\hat{p}_1}{\hat{\sigma}_1} \phi((\hat{T} - \hat{\mu}_1)/\hat{\sigma}_1) = \frac{(1 - \hat{p}_1)}{\hat{\sigma}_2} \phi((\hat{T} - \hat{\mu}_2)/\hat{\sigma}_2) \quad (15)$$

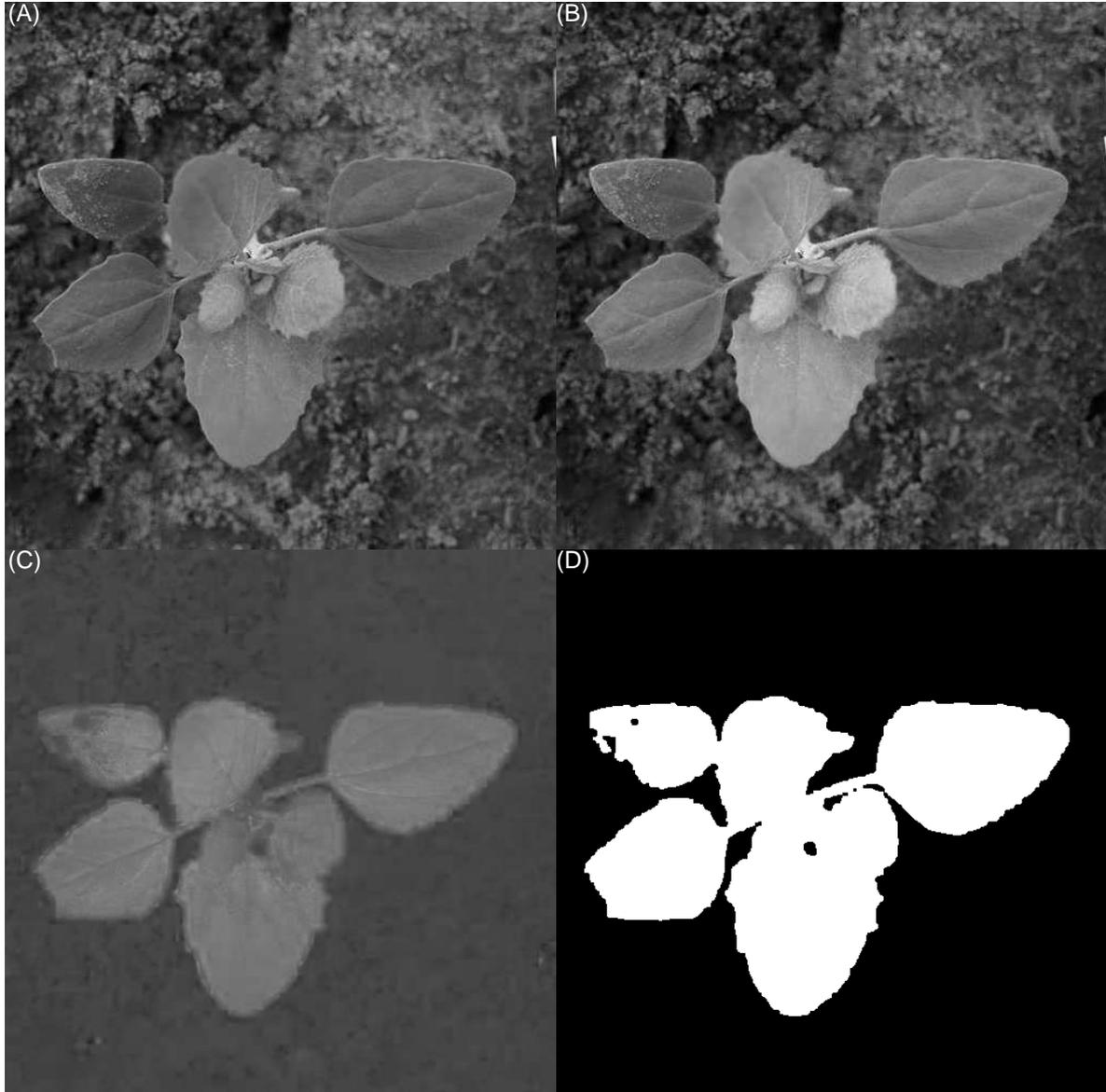


Figure 20: Images of a weed plant, lamb's quarter *Chenopodium album*, L., (A) grey scale image, (B) green channel image, (C) normalized green channel image, and (D) binary black and white image after thresholding.

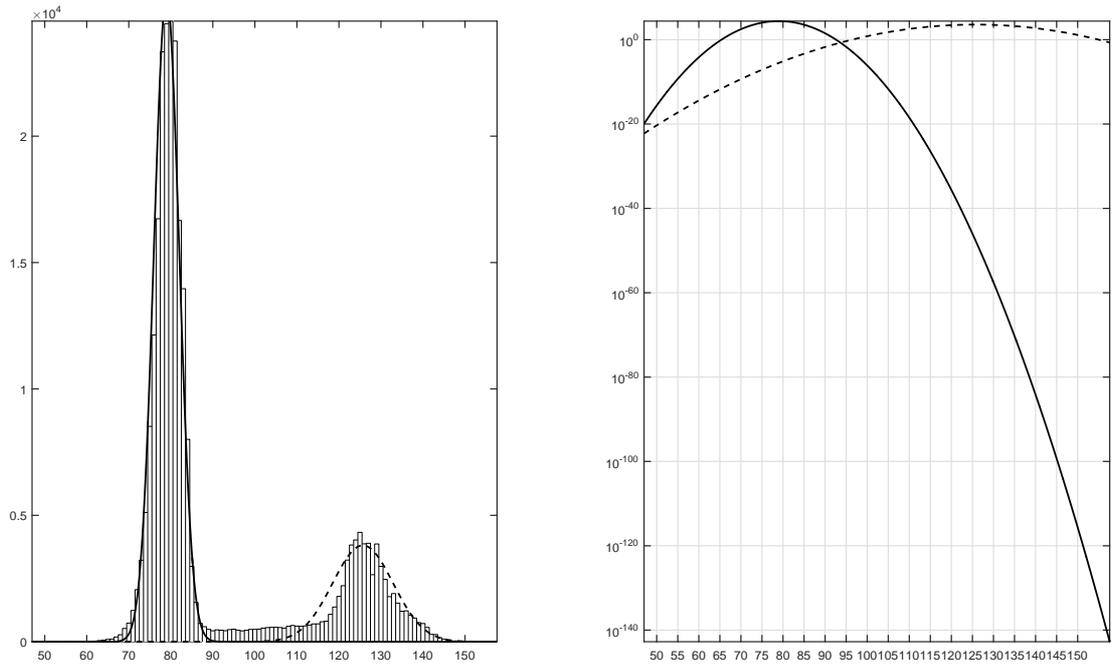


Figure 21: Left: histogram for the normalized green channel shown in the lower left part of Figure 20 and the two components shown as fully drawn and dashed curves. Right: the two components shown with a log scale on the vertical axis; here the threshold where the two curves cross can be seen.

and otherwise as a soil pixel. In the lower right part of Figure 20 we show the resulting binary black and white image obtained by thresholding the normalized green channel. For the image shown in Figure 20 we find the following parameter estimates for the two component normal mixture model

$$\hat{p}_1 = 0.263, \hat{\mu}_1 = 126, \hat{\sigma}_1 = 7.22, \hat{\mu}_2 = 79.0, \hat{\sigma}_2 = 3.02, \hat{T} = 93.6. \quad (16)$$

1.4 The Hough transform

Often one tries to find curves of specific types in images, for instance circles, ellipses or lines. A useful method to find such curves is the Hough transform. We shall here only look at the use of the Hough transform to find straight lines.

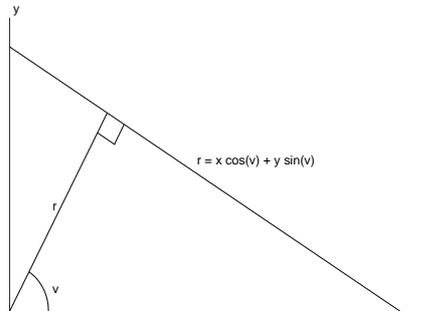


Figure 22: Representation of line in terms of angle and distance to origo.

Suppose that we have found a set S of points in an image, such as the set of tree tops in Figure 17. We are interested in finding out whether some of these points lie on lines. It is here convenient to use a representation of a line in terms of the distance r to the origin and the angle that the normal from the origin to the line forms with the horizontal axis,

$$r = x \cos(v) + y \sin(v), \quad (17)$$

see Figure 22. A point (x, y) in the original image corresponds now to a curve in the (r, v) -plane obtained by regarding r as a function of v in (17) for fixed (x, y) . In practice we discretize the (r, v) -plane into pixels regarding it as an image H and start by assigning zero to all the pixels in H . Then for each point $(x, y) \in S$ we add one to all pixels in H which the curve (17) passes through.

For the set S of maxima in Figure 17 the corresponding Hough transform for finding lines is shown in Figure 23. In particular one finds in Figure 23 three maxima in the upper left part all corresponding to the angle v equal to 16 degrees (a corresponding tick mark is placed on the horizontal axis) and three distances r (marked with three tick marks on the vertical axis close to the maximal distance r_{max}). The corresponding three lines are shown in Figure 24.

The three lines found in Figure 24 correspond actually to three lines in plot R in Figure 1 with “Heavy row thinning”, that is from the original planting in rows thinning is performed by eliminating totally some rows keeping, say, only every third row. See also Figure 2 where the rows are clearly seen in the right part of the image.

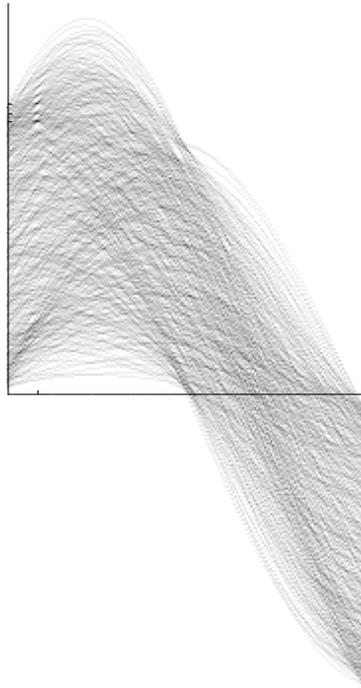


Figure 23: Hough transform for Figure 17 with angle v on the horizontal axis extending from 0 to 180 degrees and distance r on the vertical axis extending from $-r_{max}$ to r_{max} , where r_{max} is the length of the diagonal in Figure 2.

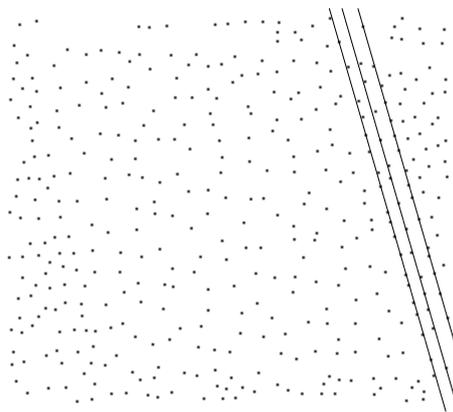


Figure 24: Location of maxima in Figure 16 together with three lines found by the Hough transform.

1.5 Morphological operations

Morphological operations can be used to regularize or clean binary images. Here we will only describe some of the most basic operations such as erosion, dilation, opening and closing. These operations are defined by a structure element S consisting of a small number of pixels with one specific pixel called reference pixel. We can, for instance, choose S as a 3×3 set of pixels with the centre pixel as reference. Let $S_{i,j}$ denote the structure element moved with reference pixel to (i, j) . Let A be a set of pixels such as the set consisting of black pixels in one of the four images in Figure 19.

The *erosion* of A , denoted $A \ominus S$, is defined by

$$A \ominus S = \{(i, j) : S_{i,j} \subseteq A\} \quad (18)$$

The *dilation* of A , denoted $A \oplus S$, is defined by

$$A \oplus S = (A^c \ominus S)^c, \quad (19)$$

where A^c is the complement of A , that is the set of pixels outside A .

The operations *opening* and *closing*, denoted $\psi_S(A)$ and $\phi_S(A)$, are defined by

$$\psi_S(A) = (A \ominus S) \oplus S', \quad (20)$$

where S' denotes the structure element rotated 180° around the reference pixel, and

$$\phi_S(A) = (A \oplus S) \ominus S'. \quad (21)$$

Thus an opening consists of an erosion followed by a dilation.

1.6 Object feature measurements

In connection with pattern recognition as mentioned in examples 1.2 and 1.3 we seek features of the objects, in the examples seeds and plants, which would enable us to distinguish between different classes of objects. Examples of such features are areas and perimeters of objects. Consider a set A of pixels as in the previous section on morphological operations. The area of A is typically defined as the number of pixels in A , possibly with some regularization operation first applied to A .

To define the perimeter we need the concept of neighbouring pixels. Typically we consider neighbourhoods consisting of either four or eight neighbours. The 4-neighbourhood of pixel (i, j) consists of the four pixels $(i-1, j)$, $(i+1, j)$, $(i, j-1)$ and $(i, j+1)$. The 8-neighbourhood of pixel (i, j) consists of the aforementioned pixels and in addition the pixels $(i-1, j-1)$, $(i-1, j+1)$, $(i+1, j-1)$ and $(i+1, j+1)$.

Edge pixels of a set A may be defined as those pixels of A that have at least one neighbour from A^c , the complement of A . Let $N(A)$ denote the number of edge pixels of A with at least one 4-neighbour in A^c . Then one can show that

$$\text{perimeter}(A) = N(A)/k_4, \quad (22)$$

where

$$k_4 = \frac{4}{\pi} \int_0^{\pi/4} \cos \theta \, d\theta = \frac{4}{\pi/\sqrt{2}} \approx 0.900, \quad (23)$$

is an approximately unbiased estimate of the perimeter of A provided that all orientations of the perimeter are approximately equally common. The unit of the perimeter estimate (22) is pixel width. As with the area, it may be useful to regularize A in some way before evaluating the perimeter. For more accurate perimeter estimates, see Glasbey and Horgan (1995), pp 165–168, and further references given there.

A feature often used is the compactness of an object defined to be

$$\text{compactness}(A) = 4\pi \frac{\text{area}(A)}{(\text{perimeter}(A))^2}. \quad (24)$$

Sometimes it is useful to compare a set A of pixels with the convex hull of A , that is the smallest convex set containing A . Some care has to be taken in defining convexity for a set of pixels; one possibility is to define convexity for the point set of pixel centres. The convex perimeter of a set A is then defined to be the perimeter of the convex hull of A . One useful feature is the convexity of A defined by

$$\text{convexity}(A) = \frac{\text{convex_perimeter}(A)}{\text{perimeter}(A)}. \quad (25)$$

1.7 Moment features

Consider a grey level or binary image $f = (f_{ij}) = (f_{ij})$, and let $A \subseteq \{1, \dots, m\} \times \{1, \dots, n\}$ be a subset of pixels, typically corresponding to an object but sometimes the whole image. The moment of order (p, q) in A is defined as

$$m_{pq} = m_{pq}(A) = \sum_{(i,j) \in A} i^p j^q f_{ij}, \quad p = 0, 1, \dots, \quad q = 0, 1, \dots, \quad (26)$$

and the *centroid* is defined as

$$\text{centroid} = \text{centroid}(A) = \left(\frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}} \right). \quad (27)$$

We also consider central moments (with respect to the centroid)

$$\mu_{pq} = \mu_{pq}(A) = \sum_{(i,j) \in A} \left(i - \frac{m_{10}}{m_{00}} \right)^p \left(j - \frac{m_{01}}{m_{00}} \right)^q f_{ij}, \quad p + q > 0. \quad (28)$$

One could note that central moments are invariant with respect to translation of objects. It is possible to construct moments that are also invariant with respect to rotations. Two such second order moments are

$$\mu_{20} + \mu_{20} \quad \text{and} \quad (\mu_{20} - \mu_{20})^2 + 4\mu_{11}^2. \quad (29)$$

An informative discussion of different types of moments with literature references can be found in Glasbey & Horgan (1994), pages 156–161.

In Example 1.10 we saw how we could discriminate between plant and soil pixels quite well by use of a suitable feature, the normalized green colour. To discriminate between classes of objects we can as will be seen in detail in the next chapter on pattern recognition use a number of suitable chosen feature variables. In the following example we will consider two feature variables and a suitable plotting technique.

Example 1.11. *Handwritten digits. Continuation*

In this example we will consider discrimination between digits “one” and “two” by use of two second order moments. We use digits “one” and “two” among the first 400 digits in MNIST. Plotting moment μ_{11} on the vertical axis versus moment μ_{20} on the horizontal axis we get the plot shown in Figure 25. Try to draw by free hand first a straight line

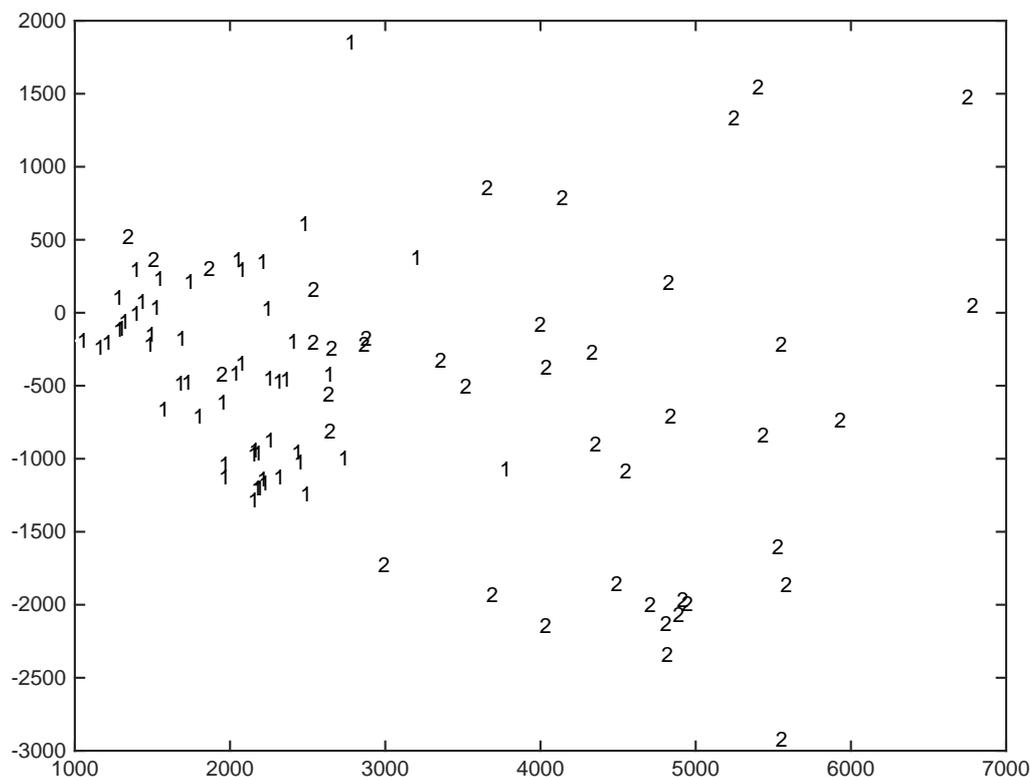


Figure 25: Plot of μ_{11} versus moment μ_{20} for handwritten digits digits 1 and 2 among the the first 400 digits in the MNIST data base.

and then an ellipse that gives as good a discrimination as possible between the “one” and “two” digits. In the next chapter we shall describe systematic methods to draw such boundaries.

1.8 Exercises

The images used in the exercises below may be found at <http://www.math.chalmers.se/~rudemo/images.html>

Exercise 1.1. Let R, G and B denote the values in the red, green and blue channels for one of the images from Example 1.3. Get the grey-level image corresponding to normalized

green,

$$g = \frac{G}{R + G + B}.$$

Exercise 1.2. Find the histogram for the image of Exercise 1.1. Try to segment the image by use of the histogram.

Exercise 1.3. Compute area, perimeter and compactness for the green segment for the image of the two previous exercises.

Exercise 1.4. Get one of the seed images from Example 1.2. Note that one has to resample the image to get the correct form of the seed. How can that be done? After resampling, reduce the number of columns to get a square image.

Exercise 1.5. Apply the averaging filter (2), the median filter (5) and the edge emphasizing filters (6) and (7) to the image of the previous exercise.

Exercise 1.6. Consider the image from Exercise 1.4. Compute the histogram and transform to a binary image. Zoom in to see the individual pixels at the object edge. Apply the operations erosion, dilation, opening and closing. What is the effect of these operations? What happens when one iterates these operations?

1.9 Literature on image analysis

There is a wealth of books on digital image processing. Glasbey and Horgan (1995) give an excellent treatment from a statistical point of view focussing on examples from biology. A mathematically oriented text is Rosenfeld and Kak (1982), which is now a bit old but still quite useful. A comprehensive introduction to image processing and analysis may be found in Sonka, Hlavac and Boyle (1999).