

# Spatial statistics and image analysis. Lecture 9

Mats Rudemo

May 4, 2020

Today's lecture will cover

Repetition of nonlinear warping

Two-dimensional electrophoresis

Point processes observed with noise. First application: Finding tree top and tree base positions from aerial photos.

Point processes observed with noise. Second application: Finding tree tops from aerial photos acquired from different angles.

## Repetition: Warping with bilinear and perspective transformations

We have a reference image  $Y = Y(x)$  and another image  $Y'$  that we want to warp (transform) into  $Y$  as closely as possible such that  $Y'(x')$  is close to  $Y(x)$ .

A simple nonlinear warping is the *bilinear transformation*

$$\begin{aligned}x'_1 &= a_{11}x_1 + a_{12}x_2 + c_1x_1x_2 + b_1 \\x'_2 &= a_{21}x_1 + a_{22}x_2 + c_2x_1x_2 + b_2.\end{aligned}\tag{1}$$

Note that for fixed  $x_2$  the bilinear transformation (1) is linear in  $x_1$  with slope and intercept depending on  $x_2$

An axes-parallel rectangle in the  $x_1x_2$ -plane is transformed into a polygon with four sides and four corners in the  $x'_1x'_2$ -plane (but generally not with pairwise parallel sides).

Another nonlinear warping function is the *perspective transformation*

$$\begin{aligned}x'_1 &= (a_{11}x_1 + a_{12}x_2 + b_1)/(c_{11}x_1 + c_{12}x_2 + 1) \\x'_2 &= (a_{21}x_1 + a_{22}x_2 + b_2)/(c_{21}x_1 + c_{22}x_2 + 1).\end{aligned}\tag{2}$$

To choose parameters of a warping transformation  $x' = f(x) = (f_1(x_1, x_2), f_2(x_1, x_2))$  consider minimization of a distortion-weighted least squares criterion function

$$\ell(Y', Y, f) = \sum_x (Y'(x') - Y(x))^2 + \lambda D(f), \quad (3)$$

where  $D(f)$  is a distortion measure of the warping function  $f$ , and  $\lambda$  is a non-negative weighting constant

For normally distributed variables least squares minimization corresponds to log-likelihood maximization, and a method where we use a distortion measure as in (3) is often called a *penalized log-likelihood method*.

The distortion measure could measure the deviation from linearity of the warping function, and could be a sum of squared second derivatives of  $f$  integrated over the region regarded,

$$D(f) = \sum_{i=1}^2 \sum_{j=1}^2 \sum_{k=1}^2 \int \left( \frac{\partial f_i}{\partial x_j \partial x_k} \right)^2 dx_1 dx_2 \quad (4)$$

Partial derivatives are approximated by finite differences.

Integrals are approximated by sums over pixels.

A useful type of warping consists of a grid of local bilinear transformations, which can be used to match and average handwritten digits and to match 2D gels electrophoresis images

To find the average handwritten 5-digit we first average over 958 5-digits. Then we warp all 958 digits separately with the average as target by use of a  $7 \times 7$  grid of bilinear transformations. Then we average the warped 5-digits, warp into the new average and proceed iteratively until changes are sufficiently small. After a few iterations we obtain the average shown in Figure 2.

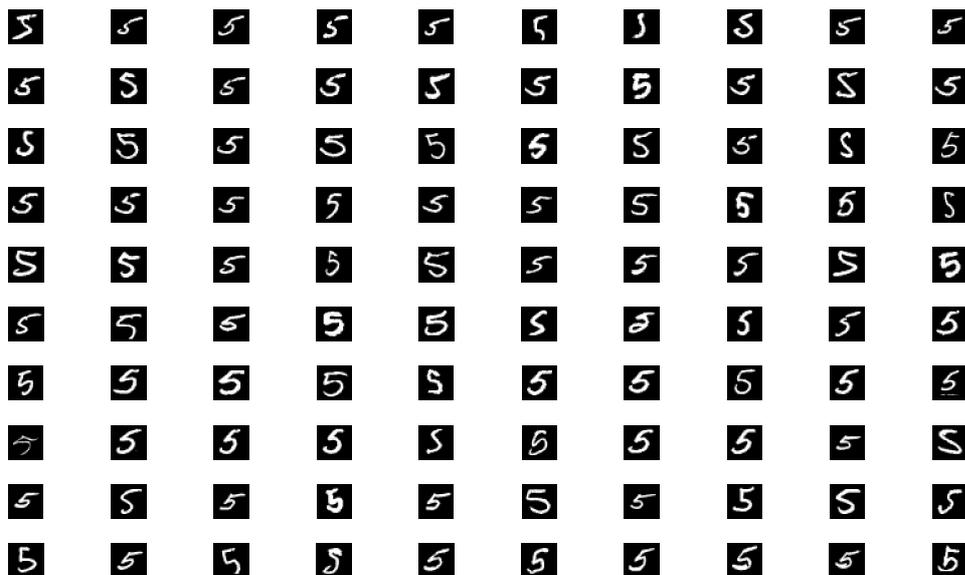


Figure 1: First 100 digits “5” in the MNIST database.

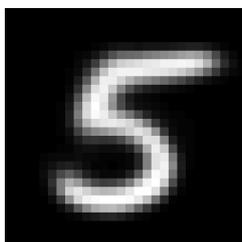


Figure 2: Average handwritten digit “5” obtained by sequential warping and averaging.

## **Two-dimensional electrophoresis**

An experimental technique to measure the expression of up to several thousands of proteins

Our image data consist of five images from 2D gel electrophoresis of baker's yeast grown in a standard solution

and five images from 2D gel electrophoresis of baker's yeast grown under stress in a solution with salt added

Electrophoresis images are obtained in two steps:

(i) Protein molecules move horizontally along a string to a position determined by the protein isoelectric point  $pI$  (plus random noise)

(ii) A polyacrylamide gel is cast between two glass plates separated from each other by thin plastic spacers and placed vertically in a bath.

The protein string is placed horizontally on the top of the polyacrylamide gel. A voltage is applied between the upper and the lower boundaries of the plates and the proteins perform a Brownian motion with downwards vertical drift in the bath.

The vertical distances traveled by the protein molecules are determined (except for random noise) by the protein mass.

During the second step there may be current leakage sideways

Warping consists of two steps.

In the first warping step a partial differential equation is solved with suitable boundary conditions taking care of current leakage.

The result of the first warping step is illustrated in Figure 3

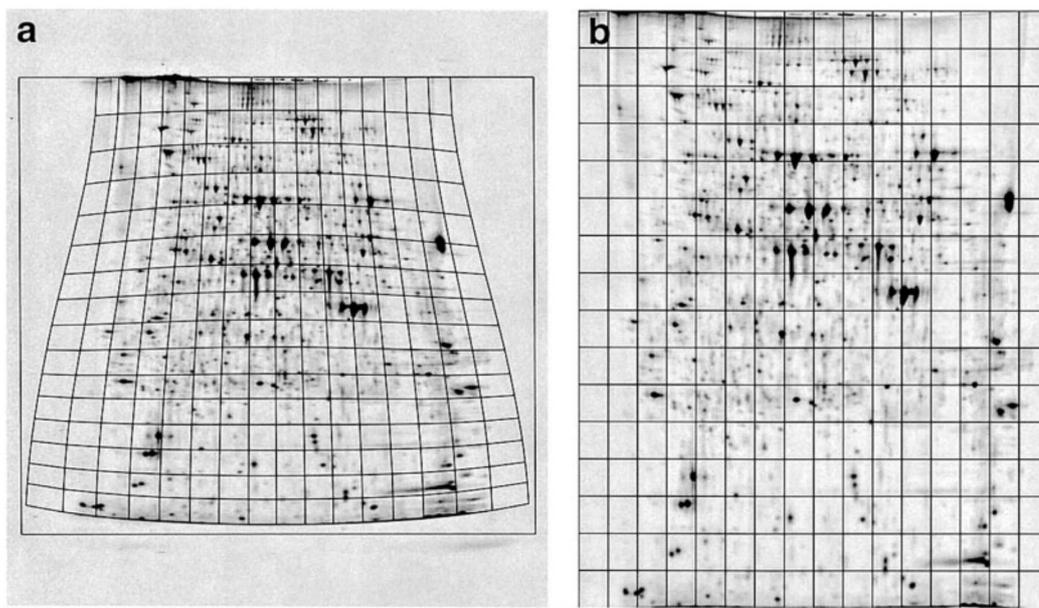


Figure 3: Illustration of warping step I with correction for current leakage sideways through the left and right boundaries during the second-dimensional gel electrophoresis. Part **a** of the figure shows the original image and part **b** shows the warped current-leakage corrected image.

After the first warping step two image transformations are applied. Firstly, to compensate for large scale trends in the background level, a top-hat transformation is applied, Secondly, a logarithmic transformation of pixel values is applied.

In the second warping step images are transformed by use of a  $12 \times 8$  grid of bilinear transformations similar to the warping of handwritten digits

The result of such a warping is shown in Figure 4.

One of the five images for yeast grown under standard conditions is used as a reference image, and the other nine images are warped onto this reference image.

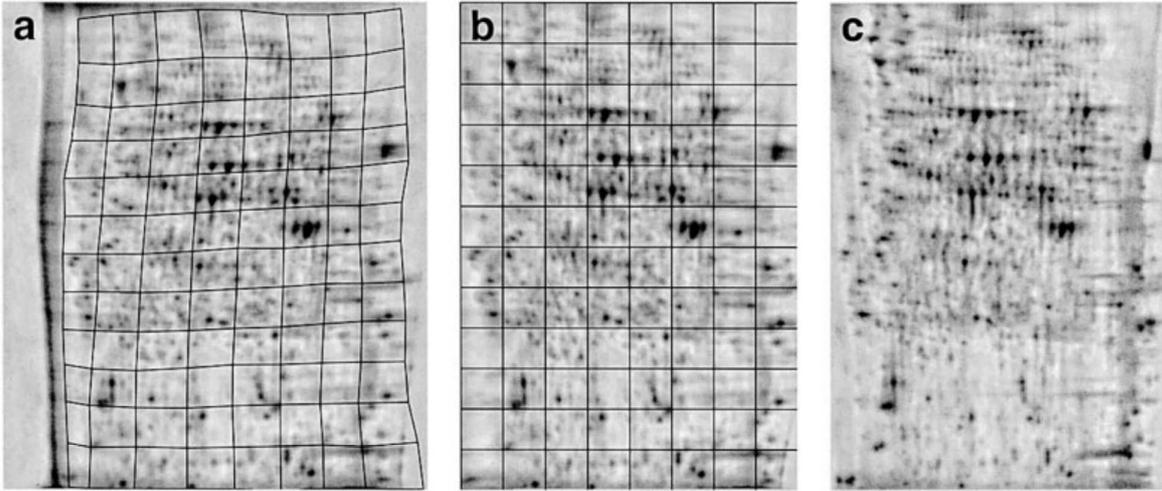


Figure 4: Illustration of warping step II. The image in **a** is warped onto the reference image in **c** by use of the grid shown in **a** warped to the grid in **b**.

We use a penalized log-likelihood method and minimize a criterion function such as (3) with  $D(f)$  given by (4). Thus we minimize with respect to  $f$  the criterion function

$$\ell(Y', Y, f) = \sum_x (Y'(x') - Y(x))^2 + \lambda \sum_{i=1}^2 \sum_{j=1}^2 \sum_{k=1}^2 \int \left( \frac{\partial f_i}{\partial x_j \partial x_k} \right)^2 dx_1 dx_2, \quad (5)$$

with  $x' = f(x)$  and where we sum over pixels  $x$ . The partial derivatives in computations are approximated by finite differences, and the integrals are approximated by sums over pixels.

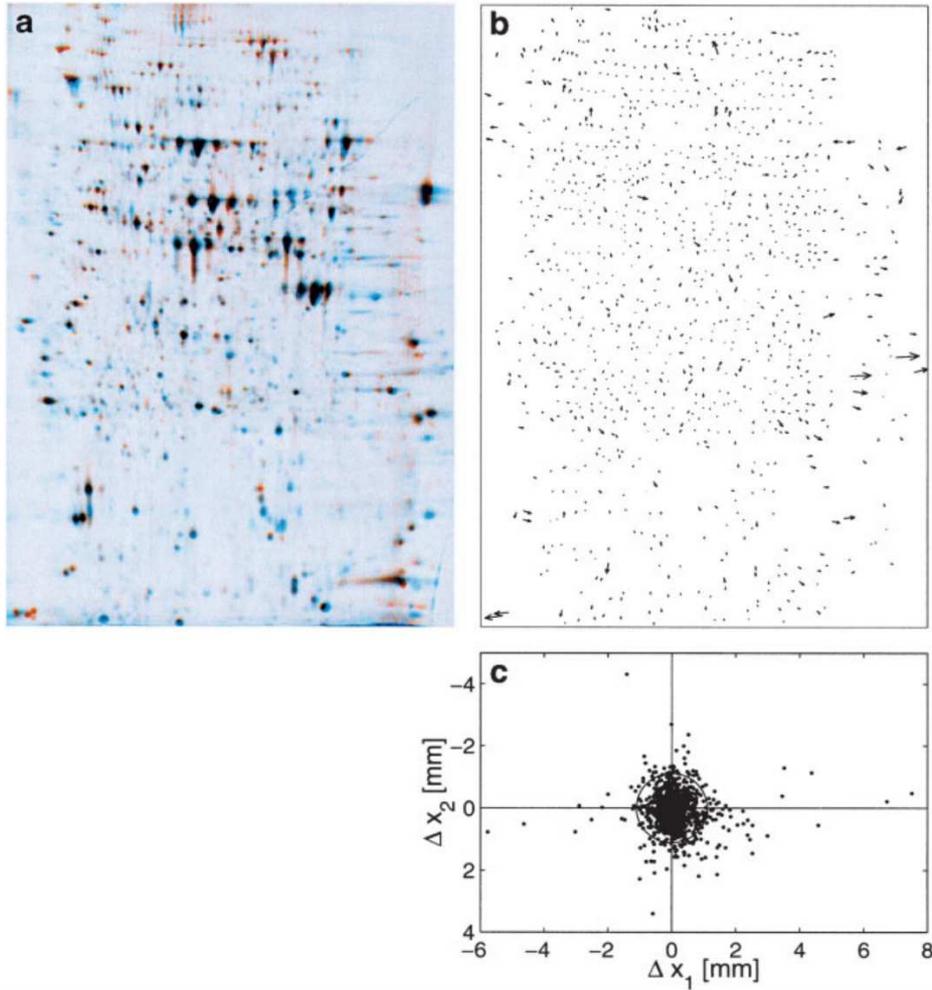


Figure 5: Further illustration of warping step II. In part **a** the reference image coloured red and the warped image coloured blue are superimposed. Displacement vectors for spots are shown in part **b**, and also in part **c**, here as relocated vectors starting at the origin and ending at dots. In **c** we also show a criterion for adjacency of spot pairs: adjacent spot pairs have dots within the circle shown.

The second warping step is further illustrated in Figure 5. We show in part **a** of the figure a superposition of the reference image coloured red and the warped image coloured blue.

For protein spots that are equally expressed in both images we should then ideally get black spots.

If the warping is less perfect we expect adjacent spots coloured red and blue.

In part **b** of Figure 5 spot displacement vectors are shown, and for more clear illustration arrow heads are large for large displacements.

Large displacements mainly occur close to the borders.

Spot displacement vectors are also shown in part **c** of the figure, with displacement vectors are relocated so that they start in the origin and end in positions shown as dots.

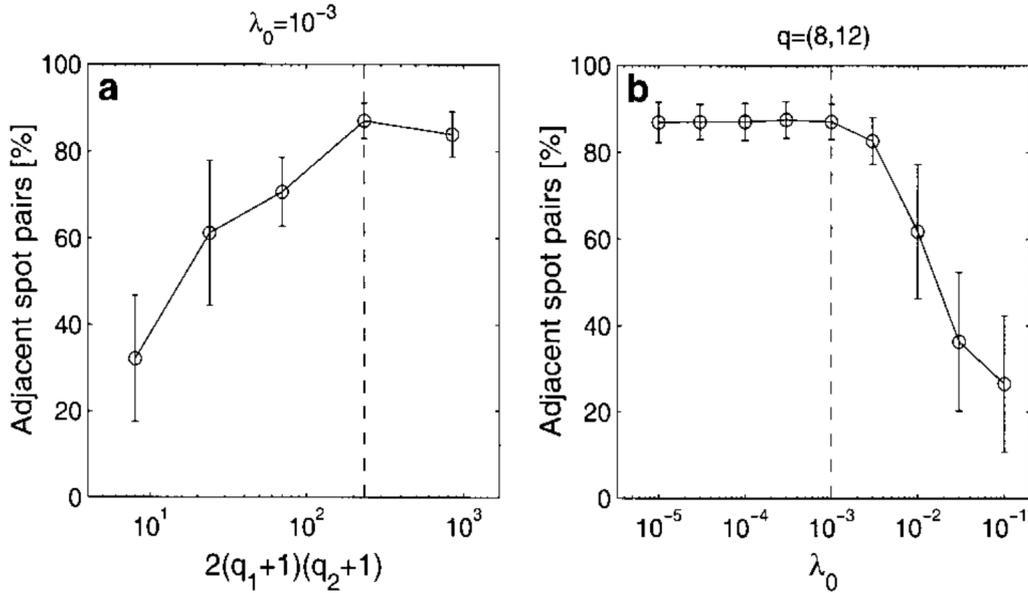


Figure 6: Illustration of spot pattern similarity in aligned images. The left part **a** shows the effect of changing grid size for the particular  $\lambda$ -value  $10^{-3}$ . The graph shows the percentage of adjacent spot pairs as a function of the number of grid size parameters. The right part **b** shows the effect of changing the log-likelihood penalizing parameter  $\lambda$  for the particular grid  $q = (8, 12)$ , and the graph shows the percentage of adjacent spot pairs as a function of  $\lambda$ . Circles show mean values and error bars show standard deviations for the nine images aligned to the reference image. Vertical dashed lines show the finally chosen grid size and likelihood penalty weight.

Two crucial issues:

- (i) choice of how fine the grid in the bilinear transformation net should be
- (ii) the size of the non-negative parameter  $\lambda$  in the penalization of the likelihood in (5).

If we start with a coarse net and steadily refine it we can expect the fit to improve but to level off at a certain degree of fineness.

Similarly if we start with a large  $\lambda$ -value and then decrease  $\lambda$  we can expect an improvement in fit but a leveling off at some point.

A measure of fit: the percentage of spot pairs with dots inside the circle in **c** of Figure 5.

Specify the net grid by  $q = (q_1, q_2)$ , where  $q_1$  and  $q_2$  are the number of rectangles in the horizontal and the vertical directions.

In Figure 4 we have  $q = (8, 12)$ .

the number of parameters in a grid specified by  $q = (q_1, q_2)$  is  $2(q_1 + 1)(q_2 + 1)$ .

Use a sequence of grids with  $q$  equal to:  $(1, 1)$ ,  $(2, 3)$ ,  $(4, 6)$ ,  $(8, 12)$  and  $(16, 24)$ .

Similarly use the following sequence of  $\lambda$ -values:  $30\lambda_0$ ,  $10\lambda_0$ ,  $3\lambda_0$ ,  $\lambda_0$  and  $0.3\lambda_0$ , with  $\lambda_0 = 10^{-3}$ .

Results from some computations with different grid sizes and different  $\lambda$  parameters are shown in Figure 6. The chosen grid size is  $q = (8, 12)$ , and the chosen  $\lambda$ -value is  $\lambda_0 = 10^{-3}$ .

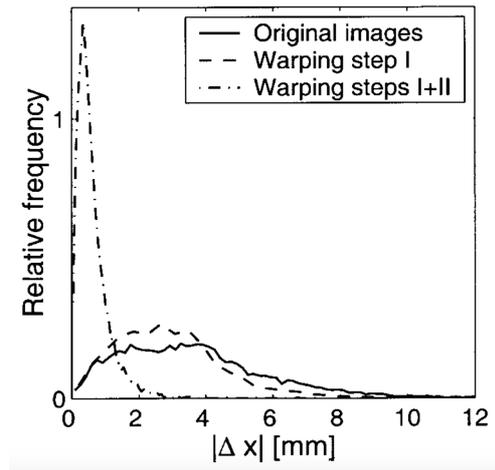


Figure 7: Length distribution of spot displacement vectors for the original data (solid line), after the current leakage warping step (dashed line) and after both warping steps (dash-dot line).

The two warping steps are compared in Figure 7, which shows the length distribution of spot displacement vectors for three sets of images: the original images, the current leakage corrected images (only warping step I) and the current leakage corrected and aligned images (warping steps I and II).

The warping step I gives some improvement, but the large improvement is obtained with the combination of both warping steps.

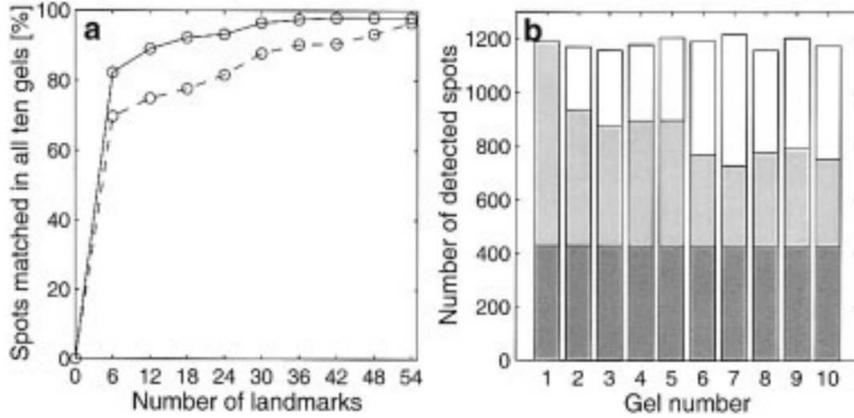


Figure 8: Efficiency profiles in the left part **a** showing the number of automatically matched spots in all ten gels (with gel images two-step warped) by the software PDQuest as a function of an initial manual matching of a number of spots (in the image called landmarks) both for the original set of images (dashed line) and for the set of warped images (solid line). The right part **b** of the figure shows the number of detected spots in the ten gels for the warped gel images. The spots detected in all gels are shown dark grey, the spots found additionally in common with the reference gel 1 is shown for each gel in light grey, while detected spots not in common with the reference gel are shown in white.

Figure 8 illustrates the improvement in matching efficiency when the warped images are used together with the PDQuest software

In the figure the reference image is divided into 54 subrectangles and in each subrectangle the most intense spot is chosen. The chosen spots are ordered according to intensity and an increasing number of the spots are manually matched.

Based on this manual matching the software PDQuest then automatically matches other spots. The left part **a** of the figure shows the global matching efficiency as the number of automatically matched spots found in all ten gel images as a function of the number of manually found spot pairs.

The dashed line shows the efficiency profile for the original images and the solid line shows the efficiency profile with two-step warping. A clear improvement using warping can be seen

In part **b** of Figure 8 we see bars showing the number of spots detected in the ten gels. Here gels 1–5 are gels with yeast grown in standard solution (including the reference gel 1) and gels 6–10 are gels grown with salt added.

The mean number of gels detected in all ten gels is 1194, and the average number of detected spots in common with the reference gel (for gels 2–9) is 826, while the number of spots detected in all ten gels is 430.

Point processes observed with noise, two examples with aerial photographs of forests

### **Estimation of tree top and tree base positions from aerial photos**

Consider estimation of tree base positions from Image 148, shown Figure 9. Start by estimating the tree top positions using a Gaussian filter.

Let  $X = (x_1, \dots, x_n)$  denote the tree base positions, and let  $Y = (y_1, \dots, y_m)$  denote the positions of maxima after the Gaussian filtering.

The object is to estimate  $X$  from  $Y$  and also to find the correspondence between  $y$ - and  $x$ -points.

Start by going the other way and estimate  $Y$  from  $X$ . If we know the positions of tree stems at ground level we can estimate the positions of tree tops as indicated in Figure 9.

The white linear segments in the figure show the expected positions of tree trunks projected on ground as seen from the aeroplane with tree ground position at the segment end point closest to the nadir point (the point vertically below the aeroplane) and the tree top position at the segment end point furthest away from the nadir point.

The model used here for the tree top positions is based on field measurements of the tree ground positions and the diameter at breast height (1.3 m), and additionally a regression of tree height from breast height diameter, which is a well-known method of estimating tree height in forestry.

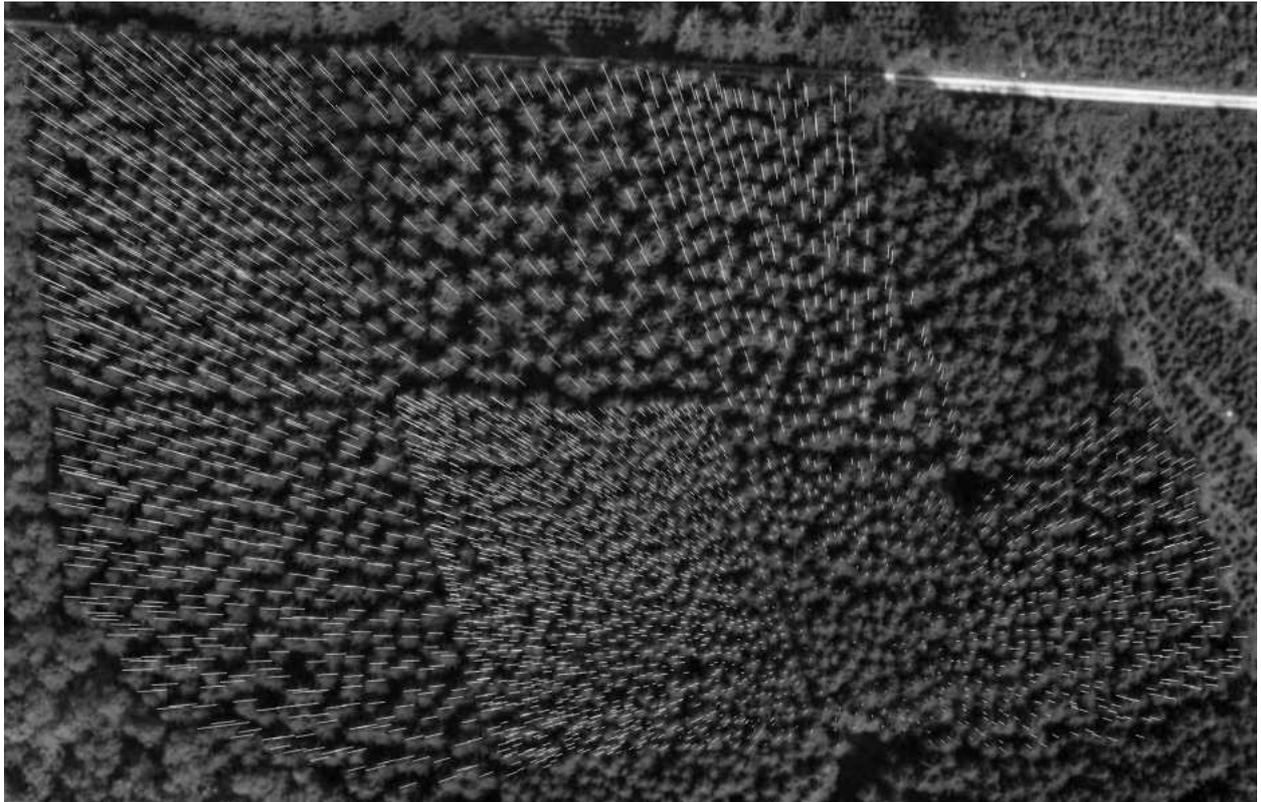


Figure 9: The Image 148 with tree projections superimposed, as seen from the camera position. The tree projections were computed from the tree ground level location measurements and the tree heights estimated from height-diameter regression and breast diameter field measurements, see the text for further details.

In practice it is tedious to measure tree ground positions and the object is to use aeroplane photographs to estimate first tree top positions and then tree ground positions. The model used is based on the assumption of three sources of distortion:

1. some trees are lost (errors of omission),
2. remaining trees become displaced as a consequence of image geometry and lighting conditions; displacement of a point  $x_i = (x_{i1}, x_{i2})$  is composed of a systematic displacement from  $x_i$  to  $x'_i$  and a random displacement from  $x'_i$  to  $x'_i + z_i$ ,
3. some spurious maxima that do not correspond to real trees are generated (ghost trees, errors of commission).

In the model we will make the simplifying assumption that these three mechanisms are mutually independent, and further that within each of these three categories the trees behave independently of each other. More specifically, we assume:

1. For each tree there is a probability  $\theta_0$ , depending on the thinning treatment, that the tree gives rise to a maximum. Thus the probability of an error of omission is  $1 - \theta_0$  for each tree, and the events that different trees are omitted are independent.
2. The systematic displacement to  $x'_i$ , see Figure 10, of the base location  $x_i$  of a tree is obtained by two displacements in the horizontal plane, or, equivalently, in the image plane. Move first along the projection of the tree a distance  $\theta_1 p_i$ , where  $p_i$  is the projection length, and move then orthogonally in the horizontal plane (to the same side of the tree projection as the sun) a distance  $\theta_2 h_i \sin \alpha_i$ . Here  $h_i$  is the height of a tree and  $\alpha_i$  is the angle between the horizontal projection of

the tree and a line which is the intersection of the horizontal plane and a vertical plane containing both  $x_i$  and the sun. The subsequent random motion  $z_i = (z_{i1}, z_{i2})$  in the local coordinate system with one axis parallel to the tree projection and one axis orthogonal to it has a two-dimensional normal distribution with means zero, standard deviations  $\sigma_1$  and  $\sigma_2$  and correlation coefficient  $\rho$ .

3. Spurious maxima are generated by a Poisson process with the intensity  $\lambda$  maxima per hectare.

For the systematic displacement from  $x_i$  to  $x'_i$  and for a corresponding maximum  $y_{j(i)}$  in the smoothed image we thus assume

$$x'_i = x_i + \theta_1 p_i e_{i1} + \theta_2 h_i \sin \alpha_i e_{i2} \quad (6)$$

and

$$y_{j(i)} = x'_i + z_i = x'_i + z_{i1} e_{i1} + z_{i2} e_{i2}, \quad (7)$$

where  $z_{i1}$  and  $z_{i2}$  are random errors and  $e_{i1}$  and  $e_{i2}$  are unit vectors, see Figure 10.

We assume that the parameters  $\theta_0$ ,  $\theta_1$ ,  $\theta_2$ ,  $\sigma_1$ ,  $\sigma_2$ ,  $\rho$  and  $\lambda$  are constant within subplots. In the discussion in the text close to Table 1 below we will see that some of the parameters vary between subplots in a way that may be interpreted in terms of thinning treatments and the geometry at image acquisition.

The parameters are estimated iteratively. For trees in a polygonal area  $A$  we compute a displaced area  $A_d$ , compare Figure 11, by moving each border polygon corner point according to the transformation (6) as if the border point was the ground location of a tree with a height computed from the height-diameter regression when the tree diameter corresponds to mean tree diameter of the subplot. Here the current estimates of  $\theta_1$  and  $\theta_2$  are used.

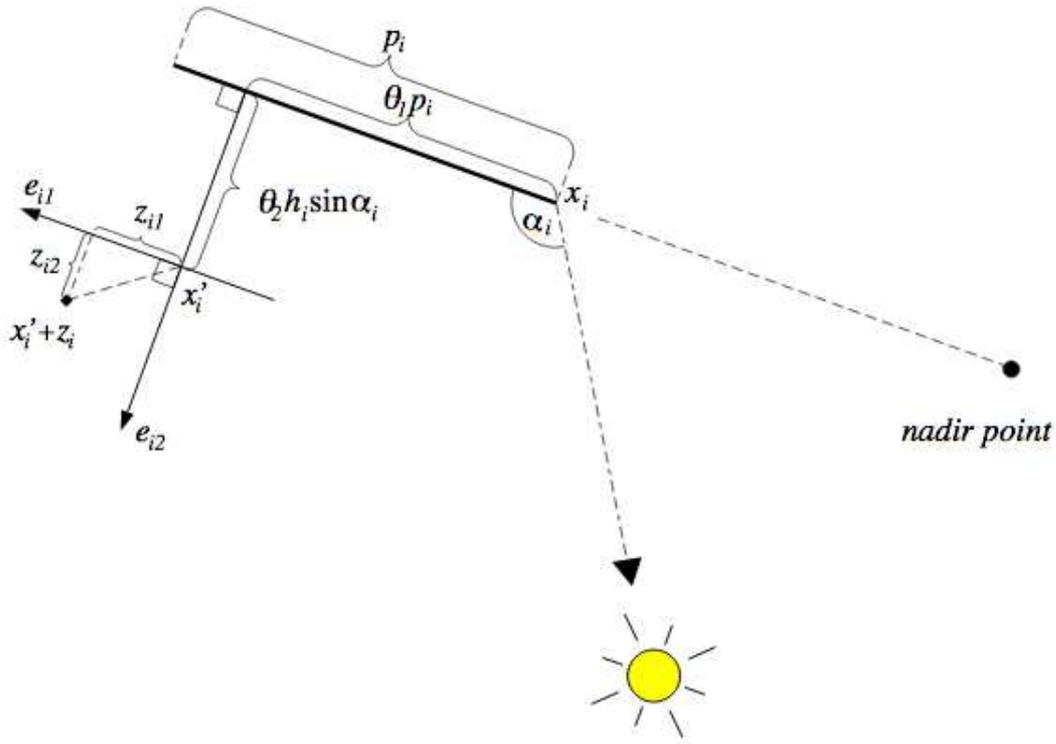


Figure 10: Displacement model for the positioning of trees. The full-drawn thick line represents the  $i$ th tree stem projected, as seen from the camera, onto the image with the base position  $x_i$  nearest to the nadir point. The length (in pixel units) of the projection is denoted  $p_i$ , and the height of the tree (also in pixel units) is  $h_i$ . The systematic displacement takes  $x_i$  to the expected position  $x'_i$  for the grey-level maximum and an additional random displacement gives the observed location  $x'_i + z_i$  of a corresponding maximum. The coordinates of  $z_i$  are assumed to have a two-dimensional normal distribution with zero means.

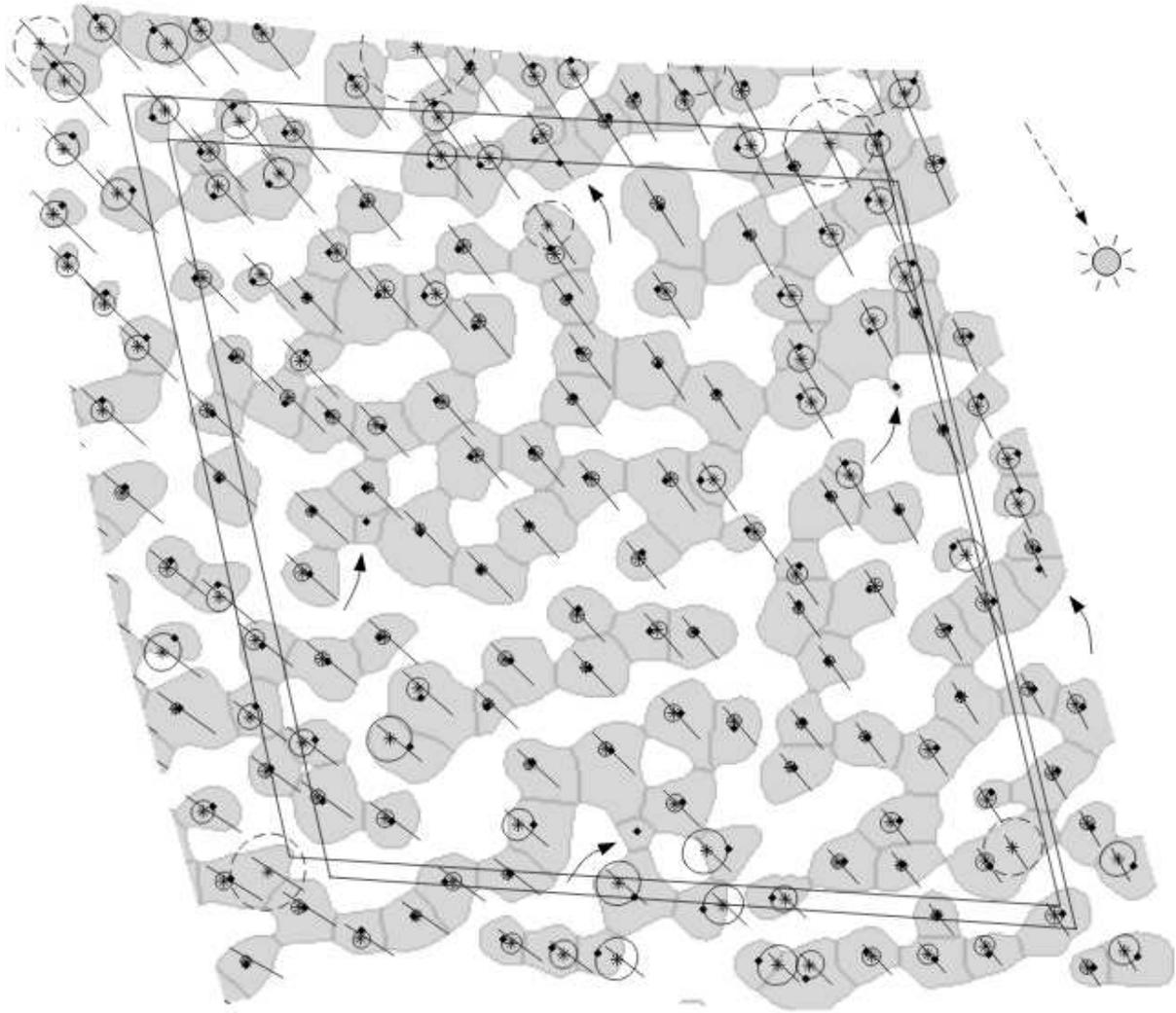


Figure 11: Subplot D in Image 148 with the net subplot borders (lower right quadrilateral) and the corresponding displaced area where maxima are expected (upper left quadrilateral). The local maxima after smoothing with the optimal bandwidth are shown as small black squares (diamonds), and for each local maximum the corresponding “watershed” segment above median grey level is shown in light grey colour with borders between segments in slightly darker grey colour. Tree projections, as seen from the camera, based on ground measurements are shown as line segments and expected positions for local maxima, according to the model indicated in Figure 10, as stars. From each star an ellipse is grown until it hits a local maximum. The ellipse is dashed if this maximum has already been hit by a smaller ellipse from another star. Thus stars with a dashed ellipse represent errors of omission, while small squares not hit by an ellipse (these squares have pointers to them) represent errors of commission. The sun azimuth is marked in the upper right corner of the image.

For a tree with base  $x_i$  we compute the position  $x'_i$  according to (6) and around this point we let an ellipse with constant probability density according to the model (7) grow. Here the current estimates of  $\sigma_1$  and  $\sigma_2$  are used but with  $\rho = 0$ . Ellipses grow simultaneously at the same rate around all points  $x'_i, i = 1, \dots, n$ , for the trees in the regarded area, or, actually, in a slightly larger area as shown in Figure 11. As soon as an ellipse starting from  $x'_i$  catches a maximum in the smoothed image the growth of that ellipse is stopped. If this maximum has not been caught from another point earlier the maximum (at  $y_{j(i)}$ ) is associated with the  $i$ th tree with base location  $x_i$ . (If the maximum has already been caught from another tree, no maximum is associated with the  $i$ th tree; it is considered lost.)

Let  $x_i, i \in I'$ , be the set of base locations in  $A$  for trees that catch maxima, and let  $n' \leq n$  be the number of elements in this set. As an estimate for the probability  $\theta_0$  that a tree gives rise to a maximum (one minus the probability of omission), put

$$\hat{\theta}_0 = n'/n. \quad (8)$$

Using (6) and (7) estimates for  $\theta_1$  and  $\theta_2$  are obtained by coordinate-wise linear regression analyses along the  $e_{i1}$ - and  $e_{i2}$ -axes, respectively, for  $i \in I'$ . Corresponding estimates for  $\sigma_1^2, \sigma_2^2$  and  $\rho$  are obtained as the sample variances and the sample correlation for the set of the  $n'$  two-dimensional residuals  $y_{j(i)} - x'_i, i \in I'$ .

An estimate for the density of spurious maxima (errors of commission) is

$$\hat{\lambda} = (m - m')/|A|, \quad (9)$$

where  $m$  is the number of maxima in  $A_d$ ,  $m' = n'$  is the number of maxima in  $A_d$  that are caught by trees, and  $|A|$  is the area of  $A$ .

The estimation procedure is performed for each of the subplots D, R, DB and B and also in groups 'all except B' and 'all', and the results are shown in Table 1, where also the estimate of the root-mean-square random displacement in metres

$$\sigma = 0.15(\sigma_1^2 + \sigma_2^2)^{1/2} \quad (10)$$

is given.

From Table 1 we see that for medium and heavy thinning, around 95% of the trees are found with a root-mean-square residual error in the displacement model of about 60 cm or less, and for light thinning around 85% of trees are found and positioned with an error of about 75 cm. The unthinned control was not investigated here because this treatment gives an exceptionally dense population, and a large number of trees are suppressed which are not possible to see from above.

Table 1: Parameter estimates for subplots in Image 148 with five different treatments.  $N$  is the true stem number per hectare;  $\theta_0$  is the probability that a tree gives rise to a maximum (and  $\hat{\theta}_0$  the corresponding parameter estimate) ;  $\theta_1$  and  $\theta_2$  specify the systematic displacement from the base location  $x_i$  to  $x'_i$  at which the corresponding intensity peak is expected (Figure 10);  $\sigma_1$  and  $\sigma_2$  (in pixel units corresponding to 15 cm at ground level) and  $\rho$  are parameters in a two-dimensional normal distribution for the random displacement  $z_i$  from the expected to the observed location (Figure 10);  $\lambda$  is the expected number of spurious maxima per hectare;  $\sigma$  is the root-mean-square random displacement in metres.

Subplot	$N$	$\hat{\theta}_0$	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\sigma}_1$	$\hat{\sigma}_2$	$\hat{\rho}$	$\hat{\lambda}$	$\hat{\sigma}$
D	367	0.970	0.651	0.028	2.74	2.94	0.370	15	0.60
C	625	0.971	0.731	0.056	2.48	1.69	0.088	37	0.45
R	746	0.980	0.634	0.082	3.20	2.12	-0.313	15	0.58
DB	824	0.956	0.767	0.006	2.69	2.19	-0.219	40	0.52
B	1257	0.843	0.871	0.045	4.29	2.65	-0.035	168	0.76
All except B		0.969	0.730	0.046	3.23	2.76	-0.096	26	0.64
All		0.925	0.734	0.045	3.61	2.75	-0.071	55	0.68

One could try to use a maximum likelihood method corresponding to our statistical model (6) - (7) for the present data set with both the image and the ground truth available. However,

a straightforward computation that takes all possible correspondences between the set of maxima and the set of trees is prohibitive, as the number of such correspondences is astronomical. One may concentrate on a small number of 'probable' correspondences. Here we have been even more reductionistic, considering only one such correspondence. After establishing the correspondence, the subsequent parameter estimation is straightforward, particularly if we assume that the correlation between errors along the tree projection and orthogonal to it is zero.

## Optimal templates for finding tree tops in aerial photos from different angles

In the previous section we analyzed aerial photos acquired essentially vertically above the area studied. We shall here study photos obtained from varying angles, in particular three specific cases where the trees are sidelighted, backlighted and frontlighted.

The tree is modelled as a generalised ellipsoid that in  $(x, y, z)$  coordinates has the surface

$$\frac{(z^2)^{n/2}}{a^n} + \frac{(x^2 + y^2)^{n/2}}{b^n} = 1, \quad (11)$$

where  $z$  is the vertical coordinate, the “centre” of the tree crown is at the origin,  $a$  is half the length of the ellipsoid,  $b$  is half the width and  $n$  is a shape parameter; here we use  $a = 17.7$  m,  $b = 2.84$  m and  $n = 1.6$ . The tree model is shown in Figure 12, where we also show how light from the sun is scattered by single reflection into the camera on board the aeroplane.

We will study an algorithm for automatically selecting tree tops from images and compare them with manually obtained “ground truth” data. Let us first consider manually estimated “ground truth” tree top positions. We can from known tree base positions estimate the tree top positions according to (6). The tree top positions thus obtained were then manually corrected by inspection of the photos to compensate for errors in tree height estimates, variations due to wind, and imprecision in image rectification.

In the upper right part of Figure 13 we see an ellipse template placed close to the top of the optical model for sidelighted trees. There are three template parameters, size  $r$ , which is the radius of a circle with the same area as the ellipse, shape  $s$ , the width to

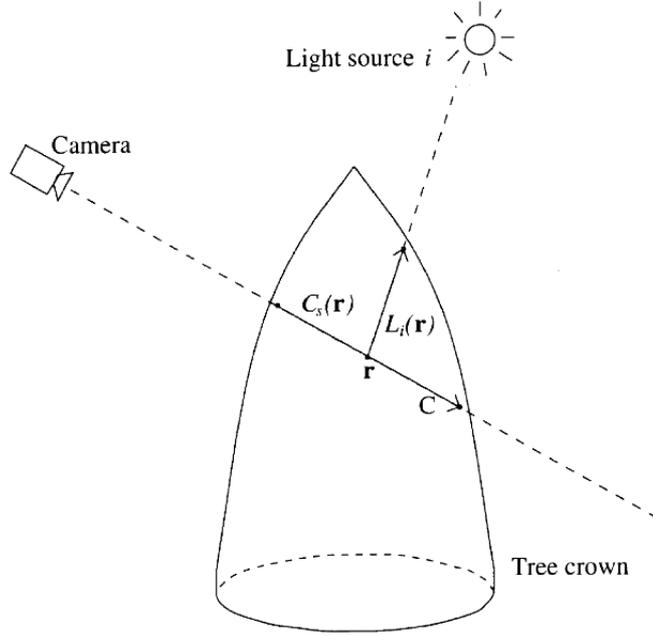


Figure 12: Tree crown model from Equation (11). The figure also shows single reflection of light from the sun into the airborne camera.

length ratio of the ellipse, and  $t$ , the translation in  $r$ -units of the ellipse centre along the tree trunk such that  $tr$  is the downwards translation.

The following procedure was used to match the set tree top candidates with “ground truth” positions. Pairs of positions from the two sets were found in order of increasing error distance such that each position in each set was used at most once. The procedure was stopped when the error distance in the next match exceeded  $d_{max} = 1$  m, and trees not matched at this stage were declared “unmatched”. As penalty measure used for comparing a set of tree top candidates with “ground truth” the following modified standard error measure was used

$$SE^* = \sqrt{\frac{\sum_{i \in \text{matched}} |\mathbf{x}_i - \bar{\mathbf{x}}|^2 + n_{\text{unmatched}} d_{max}^2}{n_{\text{matched}} + n_{\text{unmatched}}}}, \quad (12)$$

where the sum is taken over all matched tree tops,  $\mathbf{x}_i$  is the error vector for tree top  $i$ ,  $|\cdot|$  is the Euclidean distance,  $\bar{\mathbf{x}}$  is the average error vector for all matched trees, while  $n_{\text{matched}}$  and  $n_{\text{unmatched}}$  are the number of matched and unmatched tree tops.

In the search for optimal parameters  $r$ ,  $s$  and  $t$  with criterion function  $SE^*$  an iterated grid search was used. The search was performed separately for the three image with sidelighted, backlighted and frontlighted trees as seen in Figures 13 – 15. Similar parameter values were obtained with averages radius  $r = 1.5$  m, width/length ration 0.9 and translation factor  $t = 0.2$ . The number of found and missed trees for the three images are shown in Table 2.

Table 2: Results for the three images with sidelighted trees, image 120, backlighted trees, image 124, and frontlighted trees 144.

	Image		
	120	124	144
Total number of trees	171	171	171
Number of missed trees	15	6	3
Percent of trees found	91	96	98
Standard error in cm (matched only)	27	24	28
Modified standard error (12) in cm	39	30	31

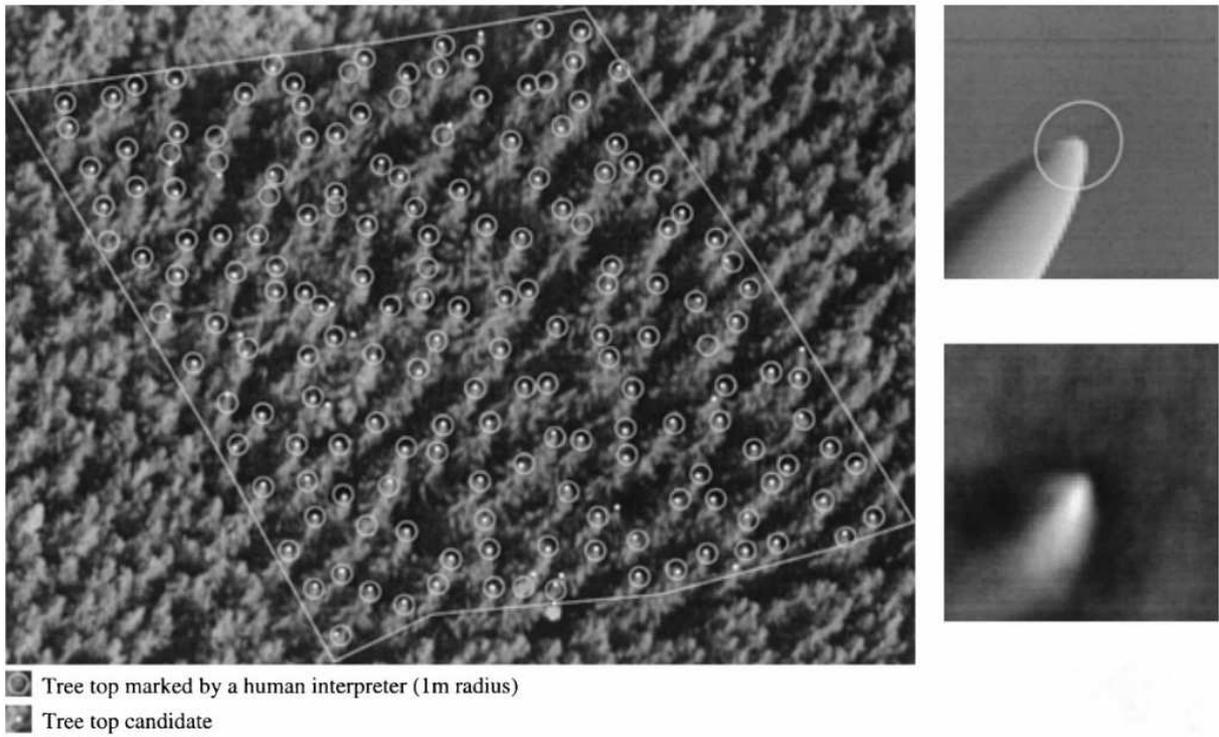


Figure 13: The sidelighted image “120” with 171 tree tops manually marked (circles) and automatically estimated (dots), and to the right the corresponding single reflection optical model with optimal boundary (upper right) and the empirical average (lower right).

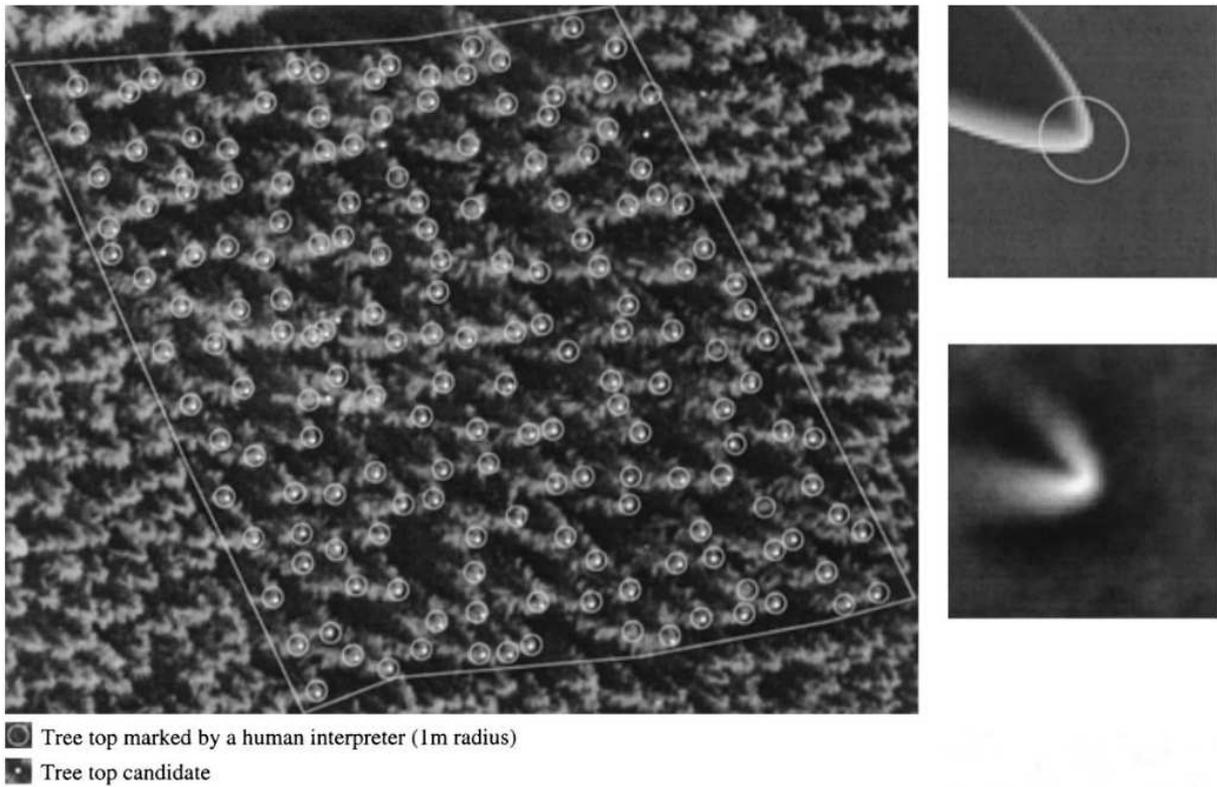


Figure 14: The backlighted image “124” with 171 tree tops manually marked (circles) and automatically estimated (dots), and to the right the corresponding single reflection optical model with optimal boundary (upper right) and the empirical average (lower right).

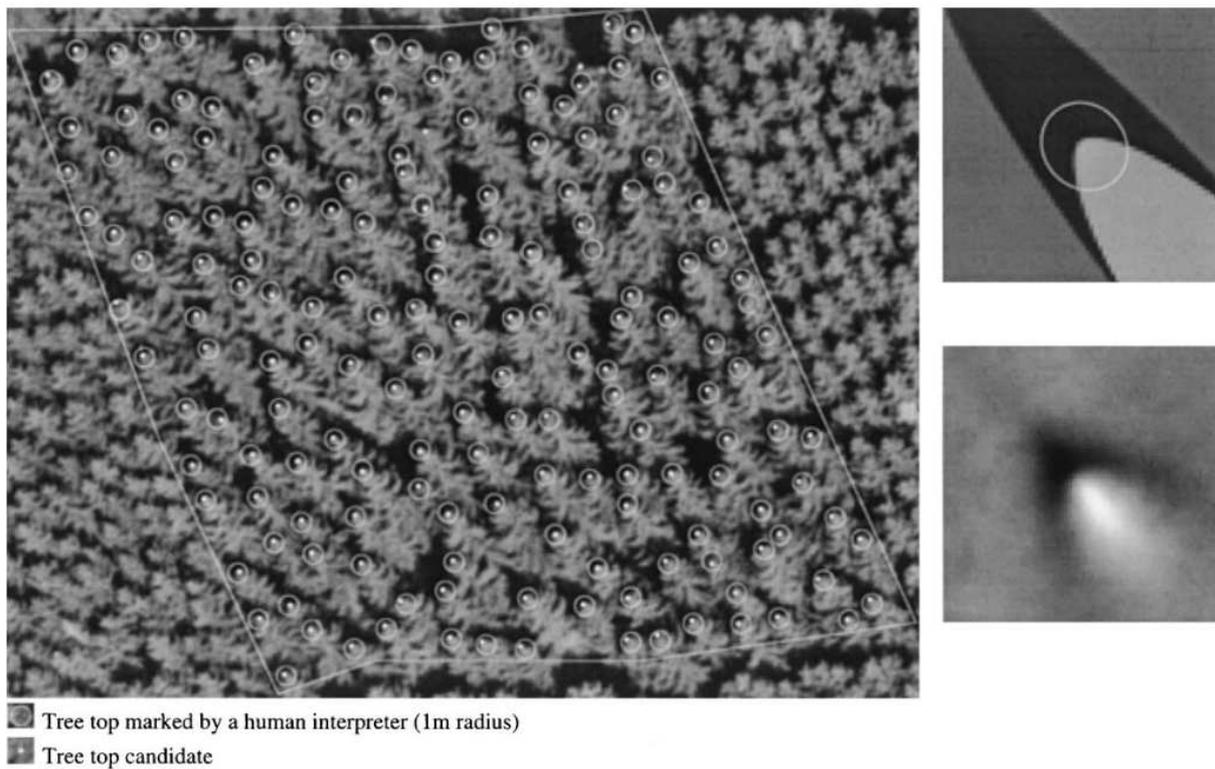


Figure 15: The frontlighted image “144” with 171 tree tops manually marked (circles) and automatically estimated (dots), and to the right the corresponding single reflection optical model with optimal boundary (upper right) and the empirical average (lower right).