

# A STOCHASTIC SHAPE AND ORIENTATION MODEL FOR FIBRES WITH AN APPLICATION TO CARBON NANOTUBES

SALME KÄRKKÄINEN<sup>✉,1</sup>, ARTTU MIETTINEN<sup>2</sup>, TUOMAS TURPEINEN<sup>2</sup>, JUKKA NYBLOM<sup>1</sup>, PETRA PÖTSCHKE<sup>3</sup> AND JUSSI TIMONEN<sup>2</sup>

<sup>1</sup>Department of Mathematics and Statistics, P.O. Box 35 (MaD), FI-40014 University of Jyväskylä, Finland;

<sup>2</sup>Department of Physics, P.O. Box 35 (YFL), FI-40014 University of Jyväskylä, Finland; <sup>3</sup>Leibniz Institute of Polymer Research Dresden, Department of Polymer Reactions and Blends, Hohe Strasse 6, D-01069 Dresden, Germany

e-mail: salme.m.karkkainen@jyu.fi, arttu.i.miettinen@jyu.fi, tuomas.turpeinen@jyu.fi, jukka.o.nyblom@jyu.fi, poe@ipfdd.de, jussi.t.timonen@jyu.fi

(Received June 6, 2011; revised November 30, 2011; accepted December 7, 2011)

## ABSTRACT

Methods are introduced for analysing the shape and orientation of planar fibres from greyscale images of fibrous systems. The sequence of image processing techniques needed for segmentation of fibres is described. The identified fibres were interpreted as deformed line segments for which two shape and two orientation parameters are estimated by the maximum likelihood method. The methods introduced are shown to perform quite well for simulated systems of deformed line segments with known properties. They were applied to TEM images of carbon nanotubes embedded in polycarbonate.

Keywords: binarization, carbon nanotubes, deformed line segments, multivariate von Mises distribution, 2D fibre identification.

## INTRODUCTION

Fibrous structures are common in biological tissues, like the actin network as a supporting structure in eucaryotic cells, and in industrial materials, like the network of wood fibres in paper. Less known examples include networks of fibres such as, for instance, those formed by electrically conducting fibres, and needed in applications like reinforcement elements, smart clothing, electromagnetic shields or armors, and textile based sensors. Among such fibres the exceptional mechanical and electrical properties of carbon nanotubes make them particularly interesting (Fig. 1). Here they were embedded in a polymer matrix and melt spun to form electrically conducting fibres (Pötschke *et al.*, 2005).

The fibrous shape of carbon nanotubes with very large aspect ratios (length to diameter ratio), even as high as 1000–10,000, allows already at very low volume contents the formation of percolated networks needed for electrical conductivity. Above the percolation threshold the network provides mechanical rigidity and connected pathways for conduction electrons.

The properties of percolating structures of *e.g.*, carbon nanotubes are influenced by their aspect ratio and spatial distribution. In Pegel *et al.* (2009) clustering of dispersed nanotubes by secondary agglomeration was found to enhance the electrical

conductivity of the system. For getting a low percolation threshold, straight and randomly oriented nanotubes are expected to be preferred, but a small degree of orientation in a nanotube system has however been found to be favourable (Du *et al.*, 2005). Even though nanotubes are typically considered as rigid rods, they rather are very flexible, and appear in networks as curved and wavy, which complicates the determination of their orientation. Thus, new methods are needed to properly determine their possible orientation, and thereby the effect of orientation on *e.g.*, the conductivity of the system.

To this end we introduce a stochastic model for the shape and orientation of fibres applied here to carbon nanotubes. In addition, we strive to produce a method for analyzing stochastic fibre systems (Mecke and Stoyan, 1980) in such cases in which the “rose of directions” as a measure of orientational anisotropy does not necessarily provide sufficient means for characterizing the system.

As a stochastic model for individual fibres we used *deformed line segments* in a similar manner as was used in Grenander and Manbeck (1993) for shape analysis. Further, the collection of fibres was modelled by a Boolean model (Matheron, 1975). To begin with, the model system was composed of a random collection of line segments. For the orientation of line segments we chose a typical model for circular data, *i.e.*, a von Mises density (Mardia and Jupp,

2000) described by two orientation parameters, strength and direction (see an illustration in Fig. 4). When deforming the line segments, each segment was divided into short segments whose orientations were thereafter ‘shaked’ stochastically such that the distribution of angles between adjacent short segments satisfied a multivariate von Mises distribution with two shape parameters. One parameter described the affinity of a segment angle to that of the previous segment, and the other parameter its affinity to the orientation angle of the original undeformed segment. The resulting ‘composite’ (deformed) segments were then the fibres of the system and they had stochastically varying shape and orientation (see an illustration in Fig. 5).

A given (two dimensional) image of a fibrous system could then be fitted by the corresponding Boolean model of fibres using the maximum likelihood method by which the parameters of the fibre model, two shape parameters and two orientation parameters, were estimated. This work is an extension to a previous model for individual fibres introduced in Kärkkäinen *et al.* (2009).

Estimation of these four parameters from the grey-scale data of the image required various image processing phases, *e.g.*, segmentation and identification of individual fibres. In this case the fibres were carbon nanotubes, and we needed to know the pixels that belonged to each nanotube, branching of fibres was not allowed as it is not physically realistic. Therefore, we developed a method for connecting segments of intersecting nanotubes where the resulting structures were physically possible and visually realistic. The method introduced is similar to that in Rizvandi *et al.* (2008), but we also used the integrated curvature of the fibres and the second derivative of their orientation to properly describe their shape near the intersection areas, in addition to the first derivative already used in Kärkkäinen *et al.* (2009).

The methods developed were tested against simulated data, and then real data were analyzed, which were taken from carbon nanotubes embedded in a polycarbonate matrix and melt spun to form *electrically conducting fibres* in microscale (Pötschke *et al.*, 2005). Images of this system were acquired by transmission electron microscopy (TEM) (Fig. 1). For this, ultrathin cuts of about 60 nm thickness were prepared from melt spun fibres of polycarbonate including 2 wt% multiwalled carbon nanotubes, and for the cutting fibres were embedded into an epoxy resin. Cutting was done along the axis of the fibrous material, which was from south-east to north-west in Fig. 1. TEM thus represents a specific 2D projection of the 3D material.

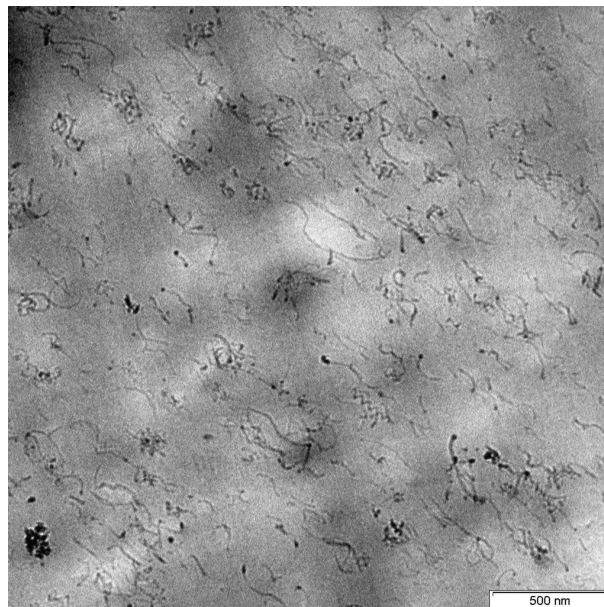


Fig. 1. A TEM image of a system of 2 wt% carbon nanotubes embedded in a polycarbonate matrix (the image is reproduced from Pötschke *et al.*, 2005).

## IMAGE PROCESSING

TEM images of carbon nanotube systems must be transformed into those of abstract network structures before they could be analyzed by the new segmentation methods developed here. To this end we needed to apply a set of image processing methods, and these are described in this section.

TEM images typically contain imaging noise, and the carbon nanotubes to be identified appeared with an inhomogeneous background, see Fig. 1 for an example. So as to obtain adequate segmentation results, the image was first bandpass filtered (Gonzalez and Woods, 2002). Filtering parameters were chosen so that only details in the size range of carbon nanotube diameter were preserved. To further reduce the noise, a median filter was applied as in Nisslert *et al.* (2007). This procedure substantially improved the image quality and allowed us to use isodata thresholding in the segmentation of the image into binary form (Ridler and Calvard, 1978).

The binary image was processed by dilation to increase the connectivity of the isolated structures (Gonzalez and Woods, 2002). Since individual small particles do not contribute to the shape of carbon nanotubes, they were removed from the image. After this the image was skeletonized (Gonzalez and Woods, 2002).

The pixels in the skeletonized image were classified into four groups, the background points,

skeleton end points, skeleton branch-intersection points and normal skeleton points. This classification was performed using the following ordered rules: If the pixel colour was zero, the pixel was classified as a background pixel. If the 8-neighbourhood of a nonzero-valued pixel contained exactly one pixel with a nonzero value, the pixel was classified as a skeleton end point. If the 8-neighbourhood of a nonzero-valued pixel contained exactly two pixels with nonzero values, the pixel was classified as a normal skeleton point. The remaining nonzero-valued pixels were classified as skeleton intersection points.

After classification each sequence of pixels formed by normal and end points represented a nanotube or part of a nanotube in a bundle of nanotubes. Intersection points were located in the areas where two or more nonparallel nanotubes were touching or one was on top of the other.

We observed that, in the above processing, some physical intersection areas were divided into two or more detached intersection areas. This problem was dealt with such that intersection areas connected with nanotube segments of lengths less than a predefined intersection diameter were merged. The merged areas contained then links to all the remaining segments in the original areas. The short connecting segments were also removed from the image. Furthermore, intersection areas containing no nanotube segments were as well removed.

At this point the actual segmentation of carbon nanotubes could be performed as follows. All intersection areas were processed one by one in a random order. In each case a weight  $w^{ij}$  was determined for each possible pair  $ij$  of nanotube segments connected to the intersection area. The segments in the pair with the lowest weight were connected to form a single fibre and removed from the intersection. This process was repeated until there were no possible pairs left, *i.e.*, there were one or none segments left in the intersection area. Thereafter the intersection area was considered as segmented and was discarded from further segmentation.

For determination of the above weight, segment directions and shapes near the intersection were quantified. To this end a predefined number of pixels was chosen from each fibre segment starting at the intersection area. Let us mark this set of points for segment  $i$  by  $\{p_1^i, p_2^i, \dots, p_n^i\}$ . For these points, the centroid and principal components were determined. The largest principal component provided a first approximation for the segment-end direction,  $\bar{v}_1^i$ . This set of points was then transformed so that the first principal component became the  $x$ -axis, and a parabola

$y^i(x) = a^i x^2 + b^i x + c^i$  was fitted to the transformed set of points. A second approximation for the segment-end direction was provided by the normalized tangent vector,  $\bar{v}_2^i = \bar{v}_2^i(x)$ , of the parabola at  $x_1^i$  such that

$$x_1^i = \arg \min_x (\|y^i(x) - p_1^i\|). \quad (1)$$

See Fig. 2 for an illustration of the relevant parameters in segment-end orientation.

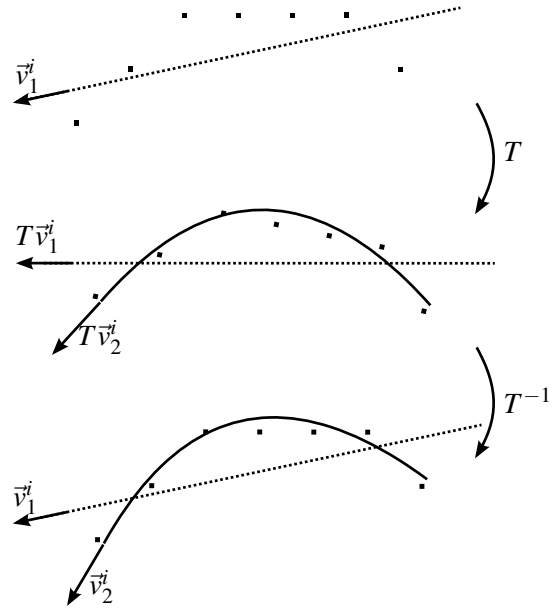


Fig. 2. Determination of the direction of segment end. Dots mark the centres of adjacent pixels that form an end of a fibre segment. On the top, the first principal component of the set of points is determined and drawn through its centroid. In the middle, the set of points is transformed ( $T$ ) so that the first principal component becomes the  $x$ -axis. A parabola is fitted to the transformed points. At the bottom, the set of points and the parabola are transformed back ( $T^{-1}$ ).

Note that a parabola could have been fitted also to the nontransformed set of points as a parametric curve. The convergence speed and robustness of the transformation – fit – inverse transformation sequence was however superior to that without transformation, and the former algorithm was thus much better suited for a completely automatic fitting method.

The shape of segment end was described by the integrated curvature  $C$  of the parabola over an interval  $[x_1^i, x_2^i]$ ,

$$\begin{aligned} C^i &= \int_{x_1^i}^{x_2^i} \frac{\partial^2 y^i(x)}{(1 + (\partial y^i(x))^2)^{\frac{3}{2}}} dx \\ &= \int_{x_1^i}^{x_2^i} \frac{2a^i}{(1 + (2a^i x + b^i)^2)^{\frac{3}{2}}} dx, \end{aligned} \quad (2)$$

with

$$x_2^i = \arg \min_x (\|y^i(x) - p_n^i\|) . \quad (3)$$

The weight for segment pair  $a, b$  was then determined such that

$$w^{ab} = \lambda_1 |C^a - C^b| + \lambda_2 \frac{\vec{v}_1^a \cdot \vec{v}_1^b}{\|\vec{v}_1^a\| \|\vec{v}_1^b\|} + \lambda_3 \frac{\vec{v}_2^a \cdot \vec{v}_2^b}{\|\vec{v}_2^a\| \|\vec{v}_2^b\|}, \quad (4)$$

in which  $\lambda_1$ ,  $\lambda_2$  and  $\lambda_3$  are material-specific parameters,  $C^{a,b}$  are integrated curvatures of the two intersecting segments and  $\vec{v}_1^{a,b}$  ( $\vec{v}_2^{a,b}$ ) are first (second) approximations to the segment-end directions. The first term on the right hand side of this equation describes the difference in the mean curvatures of the two segments, the second and third terms describe the differences in their directions. The signs of these terms were chosen so that similar curvatures and different directions of the segments resulted in a low weight. This choice favoured the formation of straight fibres.

After identification of the segments connected in the fibre crossings, we could construct sequences of unequally spaced points that could be combined so as to approximate nanotubes in the source image using a linear interpolation technique. For a resulting segmented image, *see* Fig. 10.

The optimal parameters of the above method (intersection diameter, segment length,  $\lambda_1$ ,  $\lambda_2$ ,  $\lambda_3$ ) were determined by segmentation of simulated data. To this end the fibre model described in the next section was used to create artificial images for varying fibre coverage, here defined as the percentage of the total (substrate) area covered by fibres. Segmentation result was estimated by comparing segmented simulated data with the corresponding original data with labelled fibres. The number of correctly classified pixels was used as the measure of agreement. Notice that it is possible that the segmentation splits one fibre into two or more regions. In those cases it is ambiguous which of the regions is considered correct. To assure agreement with visual perception, the largest of the regions was chosen as the correct one. Optimization of the agreement with respect to parameters resulted in the optimal parameter vector (intersection diameter, segment length,  $\lambda_1$ ,  $\lambda_2$ ,  $\lambda_3$ ) = (1, 6, 1.5, 2.5, 1.2). For this set of parameters, the fibre-detection efficiency as a function of coverage is shown in Fig. 3.

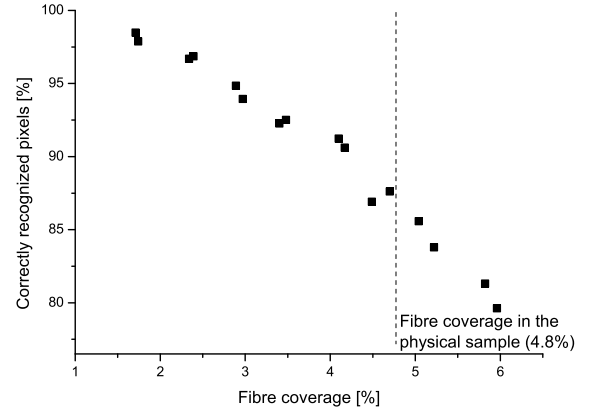


Fig. 3. *Detection efficiency as a function of coverage for simulated data. The coverage in the real carbon nanotube sample is also indicated.*

## MODELLING THE SHAPE AND ORIENTATION OF FIBRES

As a stochastic shape and orientation model for planar fibres we introduced a model of *deformed line segments* which we call a von Mises fibre. This model is a modification of the shape model used in Grenander and Manbeck (1993).

Let us consider a *single* line segment that has a fixed length and makes an angle  $\theta_0 \in [0, 2\pi)$  (the ‘‘main angle’’ in the following) with the horizontal axis. This segment was first ‘deformed’ such that it was divided into  $n$  short segments of equal length, and the orientations of these short segments  $\{\theta_i\} \in [0, 2\pi)$  were then ‘shaked’ such that they were distributed by a multivariate von Mises density in the following way. Firstly, the density of angle  $\theta_i$  for a given previous angle  $\theta_{i-1}$  and the main angle  $\theta_0$ , was given by

$$f(\theta_i | \theta_{i-1}, \theta_0, \alpha, \beta) \propto \exp\{\alpha \cos(\theta_i - \theta_0) + \beta \cos(\theta_i - \theta_{i-1})\}, \quad (5)$$

where the shape parameters,  $\alpha$  and  $\beta$ , describe the shape of the fibre such that  $\alpha$  measures the ‘affinity’ of  $\theta_i$  to  $\theta_0$  and  $\beta$  its affinity to  $\theta_{i-1}$ . In addition,  $\alpha, \beta \geq 0$  with  $\alpha + \beta > 0$ . Using standard rules of trigonometric functions and algebra, the right hand side of Eq. 5 can be expressed in the form

$$\begin{aligned} & \exp\{\alpha \cos(\theta_i - \theta_0) + \beta \cos(\theta_i - \theta_{i-1})\} \\ &= \exp\{(\alpha \cos \theta_0 + \beta \cos \theta_{i-1}) \cos \theta_i \\ & \quad + (\alpha \sin \theta_0 + \beta \sin \theta_{i-1}) \sin \theta_i\}. \end{aligned} \quad (6)$$

Express then  $(\alpha \cos \theta_0 + \beta \cos \theta_{i-1}, \alpha \sin \theta_0 + \beta \sin \theta_{i-1})$  in terms of the polar coordinates

$(\kappa_i \cos \mu_i, \kappa_i \sin \mu_i)$  with

$$\kappa_i = [\alpha^2 + \beta^2 + 2\alpha\beta \cos(\theta_{i-1} - \theta_0)]^{\frac{1}{2}} \quad (7)$$

$$\mu_i = \arctan\left(\frac{\alpha \sin \theta_0 + \beta \sin \theta_{i-1}}{\alpha \cos \theta_0 + \beta \cos \theta_{i-1}}\right), \quad (8)$$

where the branches of the arctan function must be taken appropriately. With these definitions the right hand side of Eq. 6 can be expressed as  $\exp[\kappa_i \cos(\theta_i - \mu_i)]$  which is an unscaled density of the von Mises distribution. Together with Eqs. 5 and 6 this yields

$$\begin{aligned} & f(\theta_i | \theta_{i-1}, \theta_0, \alpha, \beta) \\ &= \frac{1}{2\pi I_0(\kappa_i)} \exp\{\alpha \cos(\theta_i - \theta_0) + \beta \cos(\theta_i - \theta_{i-1})\}. \end{aligned} \quad (9)$$

Here the constant  $I_0(\kappa_i)$  is a zero-order modified Bessel function of the first kind (Mardia and Jupp, 2000). A similar type of density was used in Hughes *et al.* (2005).

Due to the Markovian way of shaking the orientation angles of subsegments, the density of angles  $(\theta_1, \dots, \theta_n)$  for a given  $\theta_0$  can be expressed in the form

$$f(\theta_1, \dots, \theta_n; \theta_0, \alpha, \beta) = \prod_{i=1}^n f(\theta_i | \theta_{i-1}, \theta_0, \alpha, \beta). \quad (10)$$

This density is called a multivariate von Mises density, and it differs from the one introduced in Grenander and Manbeck (1993) in which modelling was done of closed curves. Also, here  $\theta_0$  is regarded as a non-observable random variable in contrast with Kärkkäinen *et al.* (2009), where it was considered as an unknown parameter. Further, Breckling (1989) defined a von Mises process of angles with longer dependence structure instead of having  $\theta_0$  in the model to be estimated.

Consider now a set of von Mises fibres that have random main angles  $\theta_0^j$ . Without loss of generality we can assume that these  $\theta_0$ s follow a von Mises distribution with density

$$f(\theta_0; \kappa, \tau) = \frac{1}{2\pi I_0(\kappa)} \exp\{\kappa \cos(\theta_0 - \tau)\}, \quad (11)$$

where  $0 \leq \theta_0 < 2\pi$ ,  $0 \leq \tau < 2\pi$  and  $\kappa > 0$ , see Mardia and Jupp (2000) for this and other possible distributions. Parameter  $\tau$  stands for a ‘‘preferred value’’ of the main angles and  $\kappa$  measures the strength of concentration of the main angles around  $\tau$  such that small values lead to nearly isotropic fibre systems.

Combining the models, Eqs. 10 and 11, the density of angles of a single von Mises fibre is given by

$$\begin{aligned} & f(\theta_1, \dots, \theta_n, \theta_0; \alpha, \beta, \kappa, \tau) \\ &= f(\theta_0; \kappa, \tau) \prod_{i=1}^n f(\theta_i | \theta_{i-1}, \theta_0, \alpha, \beta). \end{aligned} \quad (12)$$

Regarding  $\theta_0$  as an unobserved random variable, we will focus in the following on the density which results from averaging over this angle,

$$\begin{aligned} & f(\theta_1, \dots, \theta_n; \alpha, \beta, \kappa, \tau) \\ &= \int_0^{2\pi} f(\theta_0; \kappa, \tau) \prod_{i=1}^n f(\theta_i | \theta_{i-1}, \theta_0, \alpha, \beta) d\theta_0, \end{aligned} \quad (13)$$

in which the integration will be performed numerically.

In order to estimate the parameters of a set of  $m$  von Mises fibres described above, we introduce the following notations: Let  $\theta_0^j$  be the main angle of the  $j$ th von Mises fibre and  $\underline{\theta}^j = (\theta_1^j, \dots, \theta_{n_j}^j)$  the angles of its segments, where  $n_j$  is the number of these segments. Combining Eqs. 9, 11 and 13, the density of the observable angles  $(\underline{\theta}^1, \dots, \underline{\theta}^m)$  of all of the von Mises fibres is given by

$$\begin{aligned} & f(\underline{\theta}^1, \dots, \underline{\theta}^m; \alpha, \beta, \kappa, \tau) \\ &= \prod_{j=1}^m f(\theta_0^j; \alpha, \beta, \kappa, \tau) \\ &= \prod_{j=1}^m \int_0^{2\pi} f(\theta_0^j; \kappa, \tau) \prod_{i=1}^{n_j} f(\theta_i^j | \theta_{i-1}^j, \theta_0^j, \alpha, \beta) d\theta_0^j. \end{aligned} \quad (14)$$

In practice, we can consider any image data of fibrous systems as a realization of a random *collection* of fibres, which can be modelled using a Boolean model (Matheron, 1975) of deformed line segments as described above. The positions of the original line segments (von Mises fibres) derive from a stationary Poisson point process with an intensity given by the mean number of points in unit area. This means that the number of fibres can be considered as a random variable. The von Mises fibres may also have a random number of segments of equal length, with the density of Eq. 12 for their angles.

## ESTIMATION OF PARAMETERS

We estimated the shape  $(\alpha, \beta)$  and orientation  $(\kappa, \tau)$  parameters (generating  $\theta_0$ ) from a measured greyscale image by fitting the density of the observable

angles by Eq. 14. This estimation required an effective use of the image processing procedures described above for obtaining the needed outlines of single fibres. Furthermore, determination of the division into segments of each outlined fibre of varying length had to be performed. A similar type of approach has been introduced in Grenander and Manbeck (1993) for modelling the shape of potatoes.

After image processing, we had a collection of pixel coordinates for the outlines of  $m$  fibres ( $j = 1, \dots, m$ ). Using these coordinates, we divided the outline of each fibre into segments as follows. For a chosen end point  $z_1^j$  of fibre  $j$ , pixel points  $z_2^j, \dots, z_{n_j+1}^j$  were searched such that the distance between points  $z_i^j$  and  $z_{i+1}^j$  along the fibre was (almost) constant ( $l$ ) for all  $i$ . We thus obtained  $n_j$  segments. Then the angles  $\theta_i^j$  of the segments were determined from the orientations of the lines connecting the end points  $z_i^j$  and  $z_{i+1}^j$  of the segments ( $i = 1, \dots, n_j$ ). For the segment length we chose  $l = 5$  pixels so as to be able to apply the continuous density of angles in  $[0, 2\pi)$  introduced above.

For the estimation of the parameters, we used the maximum likelihood method. The observable angles of the von Mises fibres follow the combined density of Eq. 14. Further, the log-likelihood of the density Eq. 14 is given by

$$\begin{aligned} l(\alpha, \beta, \kappa, \tau) &= \sum_{j=1}^m \log[f(\theta_1^j, \dots, \theta_{n_j}^j; \alpha, \beta, \kappa, \tau)] \\ &= \log\left[\prod_{j=1}^m \int_0^{2\pi} f(\theta_0^j; \kappa, \tau) \right. \\ &\quad \left. \times \prod_{i=1}^{n_j} f(\theta_i^j | \theta_{i-1}^j, \theta_0^j, \alpha, \beta) d\theta_0^j\right], \end{aligned} \quad (15)$$

containing thus only four parameters that must be estimated. This is a clear advantage over the estimation of likelihood applied in Kärkkäinen *et al.* (2009). The actual estimation was based on using the R-function *optim* (R Development Core Team, 2008).

## ANALYSIS OF SIMULATED DATA

We can also use the von Mises fibre model to produce simulated data with known properties, which can be used to evaluate the performance of the analysis methods developed.

When producing simulated data, we applied the Boolean model of von Mises fibres as follows. First, a

realization of the Boolean model of line segments was produced by numerical simulation. Each line segment had a fixed length of 85 pixels, and the main angles  $\theta_0$  of these segments followed the von Mises density Eq. 11 with  $(\tau, \kappa) = (2.2, 10)$ . The locations of the end points (one end) of the line segments, *i.e.*, the starting points of von Mises fibres, were uniformly distributed in the sampling window of size  $1024 \times 1024$  pixels with an edge length of 85 pixels. 309 starting points and related line segments were created in that window, and they are shown in Fig. 4.

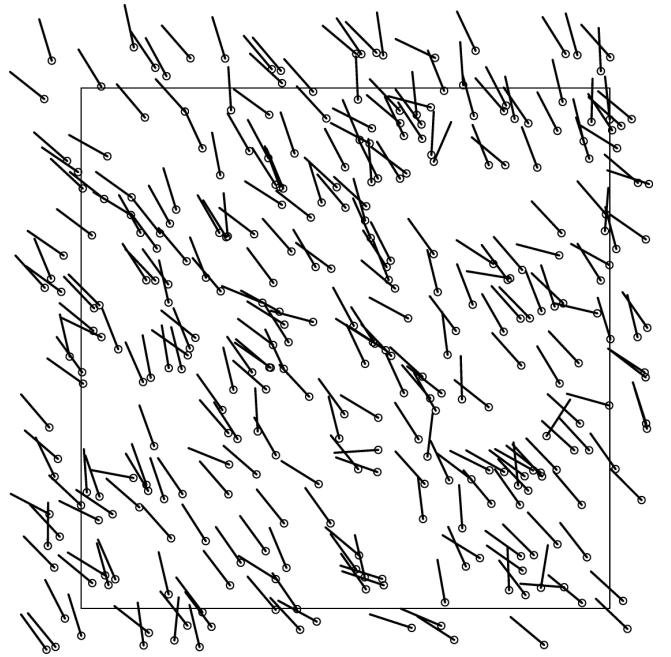


Fig. 4. Simulated line segments with a preferred main angle that indicated orientation of fibres in the north-west direction (orientation parameters  $(\tau, \kappa) = (2.2, 10)$ ). The starting points of the von Mises fibres are denoted by circles. The square is the sampling window of a size of  $1024 \times 1024$  pixels.

Then von Mises fibres were created from the above line segments. Each line segment was divided into 17 segments of a length of 5 pixels. The angles  $\{\theta_1, \dots, \theta_{17}\}$  of these segments were generated from the multivariate von Mises density Eq. 10 utilizing the von Mises densities of Eq. 9. Here we used the shape parameters  $\alpha = 0.5$  and  $\beta = 7.0$ . For an illustration of the resulting set of von Mises fibres see Fig. 5. Since in practice we most often deal with binary or greyscale images, we transformed the contents of the sampling window into a binary image of  $1024 \times 1024$  pixels (not shown).

Assessment of the performance (log-likelihood) of the analysis methods could directly be based on the simulated angles  $(\{\theta_i^j\}, i = 1, \dots, 17, j = 1, \dots, 309)$

of von Mises fibres in the whole area of Fig. 5. Using the log-likelihood of Eq. 15, we thus obtained for the estimated parameters (standard errors were calculated using the inverse of an information matrix in a standard way):  $\hat{\alpha} = 0.57$  (0.07),  $\hat{\beta} = 7.11$  (0.14),  $\hat{\kappa} = 8.60$  (0.54) and  $\hat{\tau} = 2.22$  (0.03). It is evident that the parameter values determined through the analysis tools developed are fairly close to their input values.

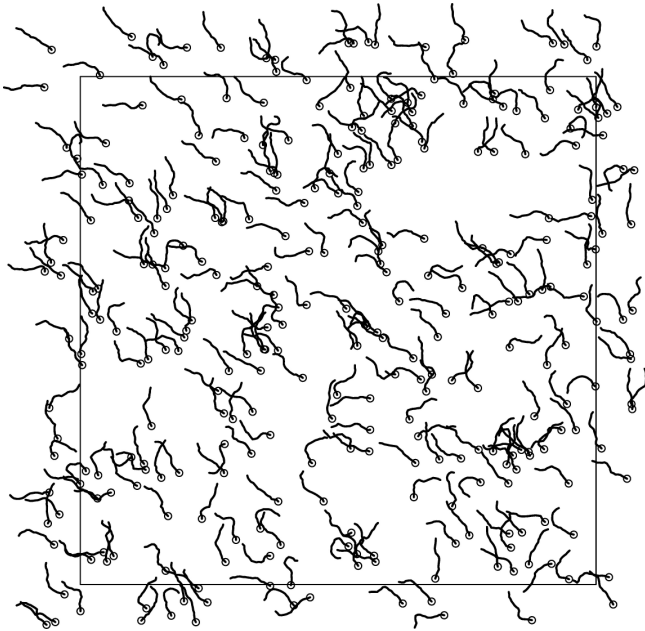


Fig. 5. A simulated set of von Mises fibres starting from the line segments of Fig. 4 such that  $(\alpha, \beta) = (0.5, 7.0)$ .

We then considered a binarized version of the simulated set of von Mises fibres. Using the image processing techniques described above, fibres of the binary image were segmented from the background, and the outlines of single fibres were determined. The starting point of a fibre was chosen such that

$$\vec{p} \cdot (\vec{x} - \vec{x}_c) \quad (16)$$

was maximized with  $\vec{p}$  the known production direction of the material, *i.e.*, the average direction of fibres;  $\vec{x}_c$  is the centroid of fibre pixels and  $\vec{x}$  the chosen starting point of the fibre, which can be at either end of the pixel sequence. As described in the previous section, division into segments was performed starting from the chosen starting point. As the result, we obtained 261 single fibres with a starting point inside the sampling window of Fig. 5. The identified fibres are shown in Fig. 7, and Fig. 6 illustrates how well individual fibres of the image were identified in one of the most challenging locations in the image.

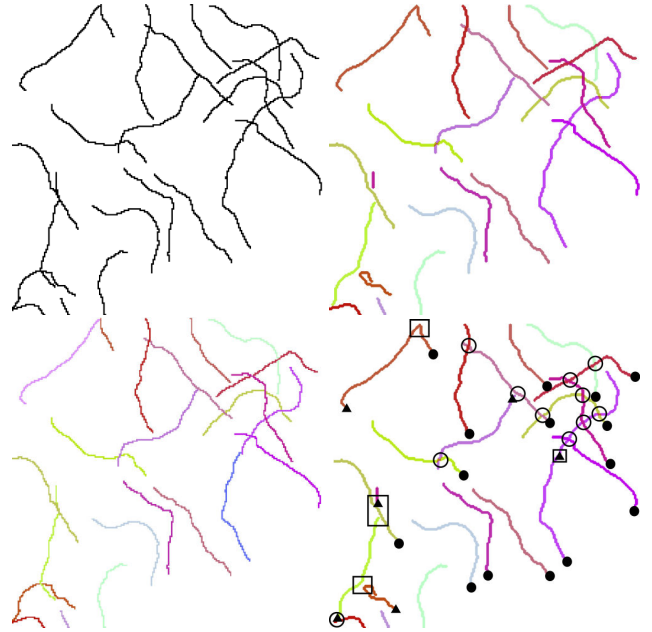


Fig. 6. Part of a binary image of simulated fibres corresponding to the upper-middle part of Fig. 5 (top-left) and the corresponding image of identified fibres coloured with random colors (top-right). The same region coloured according to true fibre paths (bottom-left) and the identified fibres with the following markers (bottom-right): circles depict correctly identified fibre intersections, rectangles depict incorrectly identified intersections, black circles depict correctly identified starting points of fibres and black triangles depict incorrectly identified starting points of fibres.

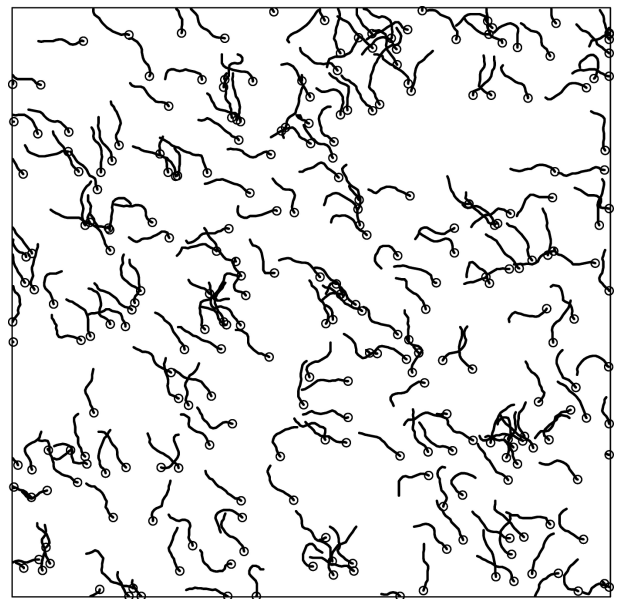


Fig. 7. The single fibres together with their starting points identified from a binary version of an image of simulated von Mises fibres. The fibres are divided into segments of a length of 5 pixels.

From the determined segments of the fibres (Fig. 7) the angles  $\{\theta_i^j\}$  were determined so as to be able to estimate the parameters. Maximizing the log-likelihood of Eq. 15, we found for their estimates with standard errors:  $\hat{\alpha} = 0.42$  (0.10),  $\hat{\beta} = 8.49$  (0.21),  $\hat{\kappa} = 3.98$  (0.42) and  $\hat{\tau} = 2.24$  (0.04). In comparison with the previous (ideal) case, the estimated values are less accurate. This was also expected as now the image was first binarized and then the single fibres together with their starting points were identified. Further, division into segments of the fibres using their outlines is expected to have had a small effect. To this end we made a small simulation study where 2 % of the fibres in Fig. 5 had wrong starting points. We found, as an example, the estimates  $\hat{\alpha} = 0.57$ ,  $\hat{\beta} = 7.15$ ,  $\hat{\kappa} = 4.73$  and  $\hat{\tau} = 2.21$ , from which we can see that  $\kappa$  is sensitive to the selection of starting points. In order to assess the decrease of  $\kappa$ , we further simulated line segments and von Mises fibres using the estimated values  $\hat{\alpha} = 0.42$ ,  $\hat{\beta} = 8.49$ ,  $\hat{\kappa} = 3.98$  and  $\hat{\tau} = 2.24$  as the parameter values. It is evident that orientation is weaker in Figs. 8 and 9 than in Figs. 4 and 5.

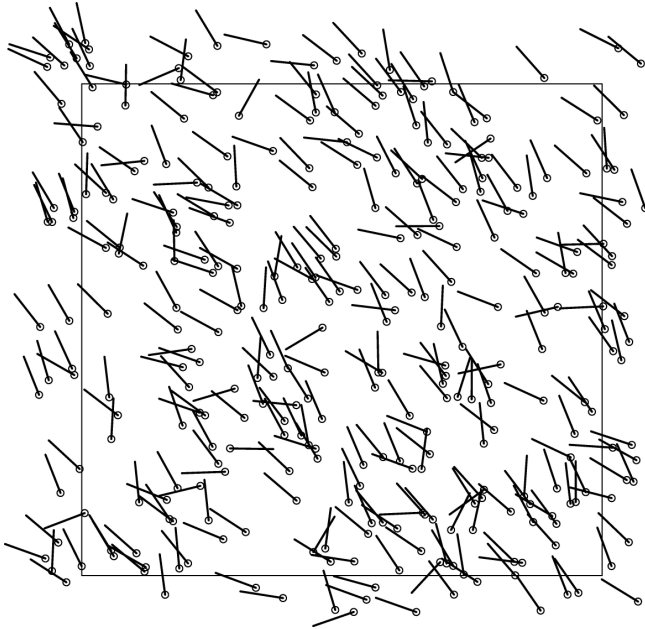


Fig. 8. Simulation result using the parameter values  $(\tau, \kappa) = (2.24, 3.98)$  estimated from the binarized image. The starting points of the von Mises fibres are denoted by circles. The square is the sampling window of  $1024 \times 1024$  pixels.

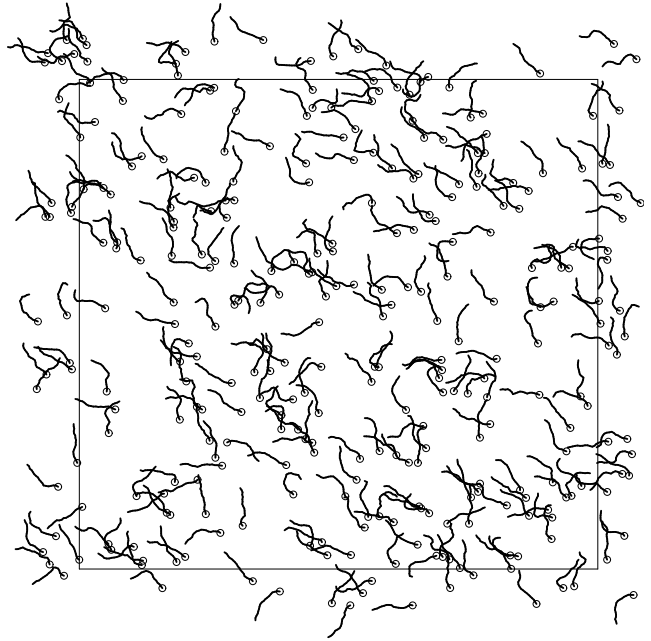


Fig. 9. A simulated set of von Mises fibres starting from the line segments of Fig. 8 such that  $(\alpha, \beta) = (0.42, 8.49)$ .

## ANALYSIS OF REAL DATA

In this analysis we used the TEM image of a system of carbon nanotubes shown in Fig. 1. The image processing techniques described above were applied to segmentation of single fibres in this image, and segmentation results are illustrated in Fig. 10. Then we proceeded to determine the corresponding von Mises model of the identified fibres. The starting point of a single fibre was chosen such that the  $\vec{p}$  of Eq. 16 described the production direction of the material evident in the image: from south-east to north-west. Having determined the starting points, the segments of the fibres and their angles  $\{\theta_i^j\}$  were determined as described above. The resulting identification of 644 von Mises fibres is shown in Fig. 11. Having now the angles  $\{\theta_i^j\}$ , the shape  $(\alpha, \beta)$  and orientation  $(\kappa, \tau)$  parameters could be estimated with standard errors by maximization the log-likelihood of Eq. 15, with the result  $\hat{\alpha} = 0$ ,  $\hat{\beta} = 5.10$  (0.15),  $\hat{\kappa} = 2.33$  (0.17) and  $\hat{\tau} = 2.28$  (0.04). Using these values, we simulated line segments and von Mises fibres shown in Figs. 12 and 13. As the lengths of line segments, we used the mean length of the identified fibres, which was about 30 pixels. Further, each von Mises fibre had 6 short segments of a length of 5 pixels. Orientation seems to be weaker in the simulated data.



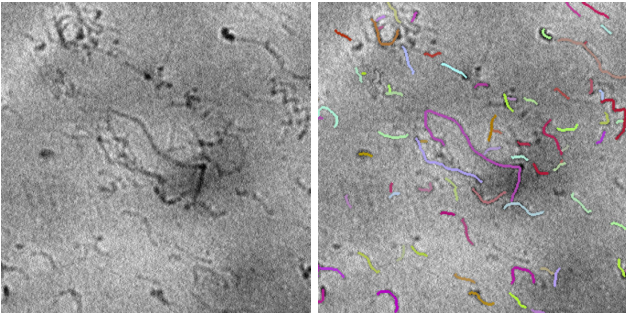


Fig. 10. Part of Fig. 1 (left) and the corresponding image in which identified fibres appear with random colouring.

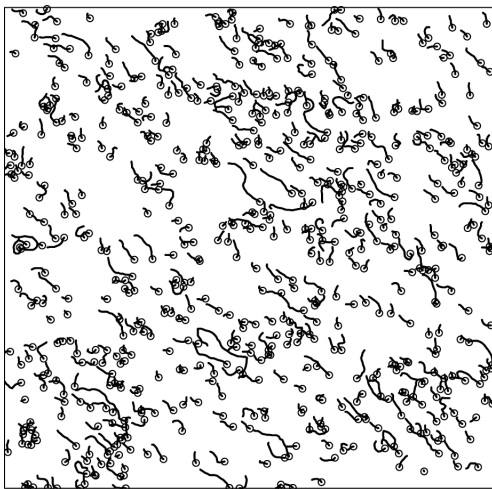


Fig. 11. The single fibres together with their starting points identified from the TEM (greyscale) image of carbon nanotubes shown in Fig. 1.

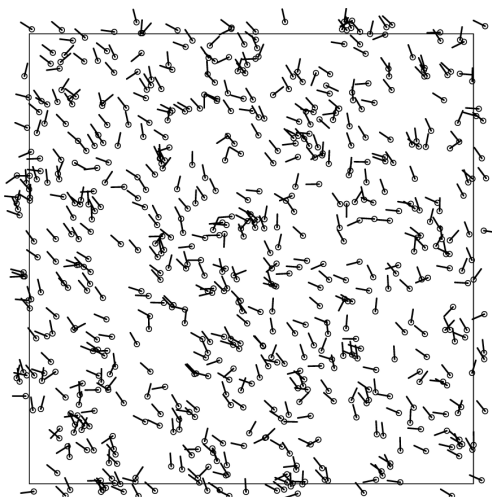


Fig. 12. Simulated line segments using the parameter values  $(\tau, \kappa) = (2.28, 2.33)$  estimated from the real data. The starting points of the von Mises fibres are denoted by circles. The square is the sampling window of  $1024 \times 1024$  pixels.

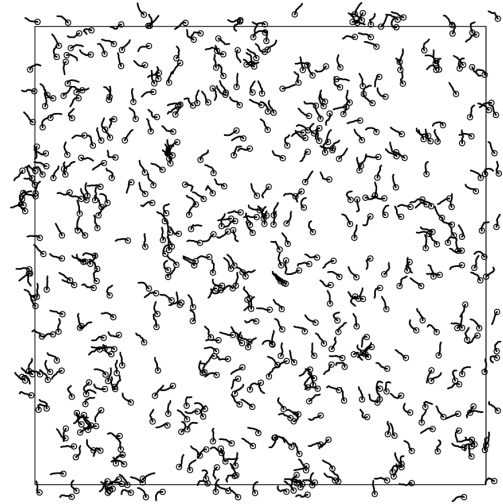


Fig. 13. A simulated set of von Mises fibres starting from the line segments of Fig. 12 such that  $(\alpha, \beta) = (0, 5.10)$ .

## DISCUSSION

New and improved methods were introduced for analysis of the shape and orientation of fibres from greyscale images of fibrous systems. Systems of von Mises fibres were generated numerically to estimate the performance of the methods developed, and as a special example a TEM image of a system of carbon nanotubes was analyzed.

Image processing techniques were introduced for handling noisy greyscale images so that they could be properly binarized, and individual fibres could be segmented from them. In segmentation the challenge was to handle fibre intersections, and to this end a rotationally invariant method was introduced. The accuracy obtained in segmentation using these techniques compared well with the present state of the art, *see e.g.*, Rizvandi *et al.* (2008). Note that the segmentation algorithm introduced does not try to guess fibre paths through intersection areas. It is evident that also these new methods work best for low density systems, and that the easiest way to improve the result of segmentation is to increase the image resolution and to decrease the inhomogeneity of the background.

To analyze the shape and orientation of the segmented fibres, they were interpreted as von Mises fibres. Thereafter their shape and orientation could be estimated based on the maximum likelihood method. The fibre model introduced is an application of the model of deformed line segments. It included two orientation and two shape parameters which were estimated. This estimation was much faster than that in Kärkkäinen *et al.* (2009), in which a somewhat

similar approach was introduced, but the number of parameters was there very large.

The performance of the shape and orientation analysis methods developed could be evaluated using simulated systems on von Mises fibres with known parameters. In the ideal case, when the angles of the simulated segments were directly used, the estimated and input values of the shape and orientation parameters were in good agreement. When various image processing techniques were used before the analysis, the results of estimation were less accurate very much as expected. When analyzing binary or greyscale images, identification of fibres and their starting points, and division into segments of fibres using their outlines, all affect the results. The most sensitive parameter to defects in the analysis appeared to be the one that describes the strength of orientation, while the orientation direction appeared to be rather insensitive to such factors. Both shape parameters behaved in a rather similar manner, and their values could be estimated with a 15-20 percent accuracy even after image processing was applied.

Using simulation studies, we showed that the selection of starting points is critical for the estimation of orientation strength. In practical applications, it will be important to utilize all possible prior information of the system as an input so as to reduce the size of the estimation problem. In the application to carbon nanotube systems considered here, the overall process direction, deduced very easily from the image, made the selection of the starting points of the fibres a much easier and more accurately performed task than otherwise would have been possible.

In the future work, a fibre model less sensitive to the selection of starting points will be sought for.

## ACKNOWLEDGEMENTS

We would like to thank Eva B. Vedel Jensen, Antti Penttinen, Aila Särkkä, Harri Högmänder and Viktor Olsbo for fruitful discussions and suggestions. The authors are also grateful to the referees for their valuable comments which helped to improve the presentation of the results. The financial support of STATCORE (a project of the University Alliance Finland) is appreciated.

## REFERENCES

- Breckling J (1989). *The analysis of directional time series: Applications to wind speed and direction*. New York: Springer.
- Du FM, Fischer JE, Winey KI (2005). Effect of nanotube alignment on percolation conductivity in carbon nanotube/polymer composites. *Phys Rev B* 72:121404.
- Gonzalez RC, Woods RE (2002). *Digital image processing*, 2nd Ed. Upper Saddle River: Prentice Hall, pp. 245–6, 523–5, 534–45.
- Grenander U, Manbeck KM (1993). A stochastic shape and color model for defect detection in potatoes. *J Comput Graph Stat* 2(2):131–151.
- Hughes G, Mardia KV, Taylor CC (2005). Circular time series with application to protein conformations. <http://www1.maths.leeds.ac.uk/Statistics/workshop/lasr2005/Proceedings/hughes.pdf>.
- Kärkkäinen S, Nyblom J, Miettinen A, Turpeinen T, Pötschke P (2009). A stochastic shape model for fibres with an application to carbon nanotubes. In: Capasso V *et al.*, eds. *Proc 10th Eur Congr Stereol Image Anal*. Bologna: Esculapio.
- Mardia KV, Jupp PE (2000). *Directional statistics*. Chichester: Wiley.
- Matheron G (1975). *Random sets and integral geometry*. New York: Wiley.
- Mecke J, Stoyan D (1980). Formulas for stationary planar fibre processes I – General theory. *Math Operationsforsch Stat Ser Stat* 11:267–79.
- Nisslert R, Kvanström M, Lorén N, Nydén M, Rudemo M (2007). Identification of the three-dimensional gel microstructure from transmission electron micrographs. *J Microsc* 225:10–21.
- Pegel S, Pötschke P, Villmow T, Stoyan D, Heinrich G (2009). Spatial statistics of carbon nanotube polymer composites. *Polymer* 50:2123–32.
- Pötschke P, Brüning H, Janke A, Fischer D, and Jehnichen D (2005). Orientation of multiwalled carbon nanotubes in composites with polycarbonate by melt spinning. *Polymer* 46:10355–63.
- Ridler TW, Calvard S (1978). Picture thresholding using an iterative selection method. *IEEE T Syst Man Cyb* 8:630–2.
- R Development Core Team (2008). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>.
- Rizvandi NB, Pizurica A, Philips W (2008). Automatic individual detection and separation of multiple overlapped nematode worms using skeleton analysis. In: Campilho A, Kamel M, eds. *Proc 5th Int Conf Image Anal Recogn. Lect Not Comput Sci* 5112:817–26.