

Image analysis of insulation mineral fibres

H. TALBOT*, T. LEE†, D. JEULIN‡, D. HANTON§ & L. W. HOBBS¶

*CSIRO, Mathematical and Information Sciences, Locked Bag 17, North Ryde, NSW 2113, Australia

†Department of Statistics, Colorado State University, Fort Collins, CO 80523-1877, U.S.A.

‡Centre de Morphologie Mathématique, ENSMP, 35 Rue Saint Honoré, F77305 Fontainebleau, France

§Centre de Recherches Industrielles de Rantigny, BP 19, 60290 Rantigny, France

¶Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge MA 02139, U.S.A.

Key words. Bias, image analysis, MMVF, quality control, sample preparation, segmentation, SEM, size distribution, statistical analysis.

Summary

We present two methods for measuring the diameter and length of man-made vitreous fibres based on the automated image analysis of scanning electron microscopy images. The fibres we want to measure are used in materials such as glass wool, which in turn are used for thermal and acoustic insulation. The measurement of the diameters and lengths of these fibres is used by the glass wool industry for quality control purposes. To obtain reliable quality estimators, the measurement of several hundred images is necessary.

These measurements are usually obtained manually by operators. Manual measurements, although reliable when performed by skilled operators, are slow due to the need for the operators to rest often to retain their ability to spot faint fibres on noisy backgrounds. Moreover, the task of measuring thousands of fibres every day, even with the help of semi-automated image analysis systems, is dull and repetitive. The need for an automated procedure which could replace manual measurements is quite real.

For each of the two methods that we propose to accomplish this task, we present the sample preparation, the microscope setting and the image analysis algorithms used for the segmentation of the fibres and for their measurement. We also show how a statistical analysis of the results can alleviate most measurement biases, and how we can estimate the true distribution of fibre lengths by diameter class by measuring only the lengths of the fibres visible in the field of view.

Introduction

Thermal and acoustic insulation materials, such as glass wool, are made of man-made vitreous fibres (MMVF). The

measurement of the diameters and lengths of these MMVF is useful for the insulation industry, and serves as a quality control tool and for process development. Important characteristics such as the thermal resistance of the insulation material can be theoretically linked using these two measurements.

Hundreds of measurements are necessary to obtain useful information about the material, but manual measurement of the diameters and lengths of so many fibres is long and tedious, and in the case of insufficiently skilled or tired operators, can be unreliable. This fact makes the automated measurement of such fibres worth pursuing. One way of automating these measurements is to use image analysis.

MMVF used in the insulation industry have a number of distinct features: the majority of fibres are elongated cylinders with parallel edges, with diameters ranging from about 0.2 μm to 20 μm ; the fibre lengths are even more variable: from about 1 μm to 1000 μm or more and there is a high incidence of irregular fibres: non-cylindrical fibres, fused fibres, broken fibres, etc.

From these characteristics, it appears that the use of electron microscopy is a necessity, and that due to the high incidence of irregularities, the automated image analysis of such fibres is not easy. The small widths of the smallest fibres implies the use of a relatively high magnification power, which is incompatible with the measurement of the true lengths of all the fibres.

Because the breadth of the spectrum of fibre diameters is smaller than that of their lengths, measuring fibre diameters is easier than measuring their actual lengths. Therefore, fibre manufacturers have been using mostly diameter measurements as their quality control tool. In this sense, the accurate measurement of the diameter of each fibre is more important. This is the reason why, during the

Correspondence: Hugues Talbot. Tel: + 61 29325 3224; fax: + 61 29325 3200; e-mail: Hugues.Talbot@cmis.csiro.au

course of this paper, we will focus mostly on how to obtain accurate, reliable and unbiased fibre diameter measurements. We will show that it is actually difficult for an automated system to obtain the true length of each visible fibre. However, we will present a way of obtaining statistically the mean length by diameter class of these fibres. In many cases, this information is almost as useful as the actual length of each individual fibre.

The paper is structured as follows: the first section presents a method that permits rapid measurements of fibre diameters only, with a relatively simple image analysis procedure, at the expense of a complex sample preparation, called the polished section method. The second section presents a method for measuring fibre diameters and estimating mean fibre lengths by diameter class, with a simple sample preparation, at the expense of complex image analysis algorithms and long image acquisition times. The third section presents the results obtained by both methods, compared with manual measurements. These results are discussed in section 4. For each method, both the sample preparation and the image analysis algorithms are presented.

1. Polished section method

It is important that MMVF manufacturers be provided with a fast and efficient quality control tool for large data processing. This tool may be of use for detecting faulty products, or during research or experiment design runs, to find quickly, among a large number of different products, which ones are of interest.

For this purpose, MMVF length measurement is not essential. We propose a method for quickly measuring a large number of MMVF diameters for a given product.

1.1. Sample preparation

This preparation was first proposed and implemented with success by Degenne (1989) on a limited number of samples. It was later enhanced and tested on a large number of products.

A tuft of binder-free MMVF is extracted from the material. Fibres present in this sample are roughly parallelized by hand, then compressed and rolled. It is cut to the desired length and labelled, then dipped and then embedded in epoxy resin in a cylindrical sample holder at room temperature. The resin is cured at 70 °C. Usually four samples can fit in a single sample holder and can be prepared in parallel. The cylinder containing the MMVF samples is cut perpendicularly to its axis, and polished to a mirror-like finish (defects smaller than 1 µm in size). This sample is coated with carbon (conductance is tested with an ohm-meter) and observed in a scanning electron microscope

(SEM) in backscattered electrons (BSE) mode. Typical images produced by such a method are shown in Fig. 1.

This apparently simple sample preparation is in fact difficult to master, because of the large number of parameters involved. Some broadly defined parameters are, for example, the way in which the MMVF are handled before the preparation even begins, the resin fluidity (which varies quickly with time), and the parameters of the polishing table. One of the most important is the pollution by non-organic materials, for example chlorine in the water used during the polishing step.

The images produced by this method are of high quality. The fibres appear as saturated white objects over a remarkably noise-free background. The use of the BSE mode eliminates all organic impurities from the images, including remaining diamond dust from the polishing phase, due to their low mean atomic number compared with the mean atomic number of MMVF. However, non-organic pollutants are generally visible on these images, which is why they must be carefully avoided at sample preparation time and during subsequent sample storage and manipulation. (Indications on the time required for the preparation are reported in Table 1.)

1.2. Image analysis algorithms

The images presented in Fig. 1 are typical. Even though the images are quite clean, their analysis is still not trivial. Indeed, Fig. 1(a) shows the wide range of fibre diameters, as well as some fused fibres and bubble defects in fibres. The smallest visible fibres are about 0.15 µm in diameter. On these images, the largest fibres are about 10 µm in diameter. Figure 1(b) shows more fused fibres, unusual fibre structures, broken fibres, etc. In particular, two or more fused fibres must not be mistaken for a single, elongated one, nor the other way around.

The analysis of these images was described by Talbot & Terol Villalobos (1992) and by Talbot (1993) in more detail. We shall only briefly describe here the main points of this method.

1.2.1. Pre-processing. The pre-processing of these images is quite simple: the grey-level histogram is obtained. The two principal peaks of the histogram are sought automatically to produce a threshold between them. Circular holes ('bubbles') in the resulting binary image are filled.

1.2.2. Segmentation using weighted skeletons. Fused fibres must be measured separately. In our case, image segmentation is the process that logically separates two or more objects fused together.

Mathematical morphology (MM) (Serra, 1982) is an image analysis theory and set of techniques well suited to the segmentation of images such as those in Fig. 1.

Table 1. Time and throughput comparison. Times are given on average per sample. A sample represents about 600 individual measurements. Throughput is estimated based on optimal utilization of available resources. The presence of a single human operator is assumed, both for sample preparation and for manual measurement.

| | Mean sample preparation time (min) | Mean image analysis time (h) | Mean manual counting time (h) | Automated throughput (samples/day) | Manual throughput (samples/day) |
|---------------|------------------------------------|------------------------------|-------------------------------|------------------------------------|---------------------------------|
| Cross-section | 30 | 1 | 2 | 16 (night) | 2–3 (day only) |
| Flat bed | 15 | 3 | 4 | 4 (night) or 8 (day + night) | 2 (day only) |

The main steps of the method are as follows (see Talbot, 1993, for more complete details), and are illustrated by Fig. 2.

Standard MM techniques for separating fused objects, such as the use of filtered ultimate erosions (Buecher & Vincent, 1990) do not work well in our case, because of the presence of fibres cut almost horizontally in the polished section, which look like elongated ellipses, on the one hand, and on the other hand, because of the presence of small fibres fused with large ones (see Fig. 2c).

A new technique was developed to deal precisely with these two problems. Talbot & Terol Villalobas (1992) proposed using their weighted skeleton to obtain an initial segmentation. This involves obtaining the skeleton (Blum, 1961) of all the binary objects in the image, and then considering the distance transform (Rosenfeld & Pfaltz, 1968) along this skeleton. The skeleton of an object is its geometric 'median axis'. The distance transform is a function applied to a binary object. The function value at each point within the object is the distance from this point

to the closest point on the boundary of the object. Intuitively, when two fibres are fused, this looks on the polished section like two fused disks or ellipses with a 'neck' between them. Looking at how the distance transform varies along the medial axis of the object (the distance transform will be smaller at the neck between the two disks than close to either centres) makes it possible to decide with precision whether or not to separate a given object into two or more particles (see Figs 2(d) and (e)).

The result of this first segmentation is further enhanced by the use of the bisector function (Talbot & Vincent, 1992). This powerful technique is used to separate fibres that are so fused together that no neck exists between them, but which should still be separated (see Figs 2(e) and (f)).

1.3. Measurements and stereological bias correction

Once the image segmentation has been performed, individual fibre measurement is straightforward: for each

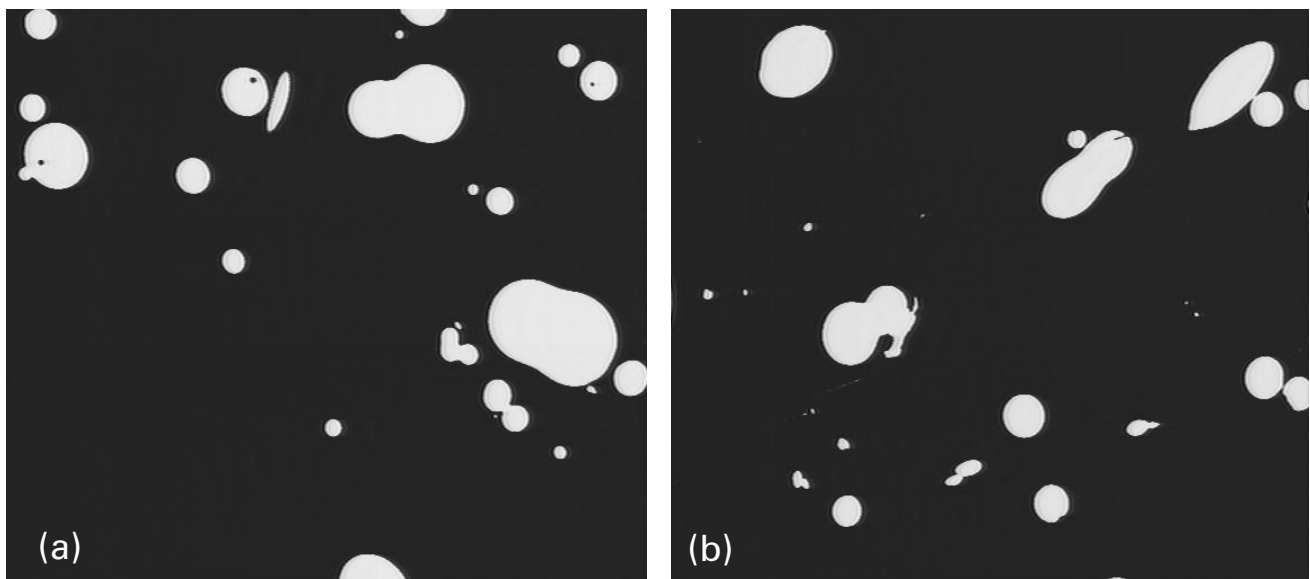


Fig. 1. Examples of images of fibres in polished section.

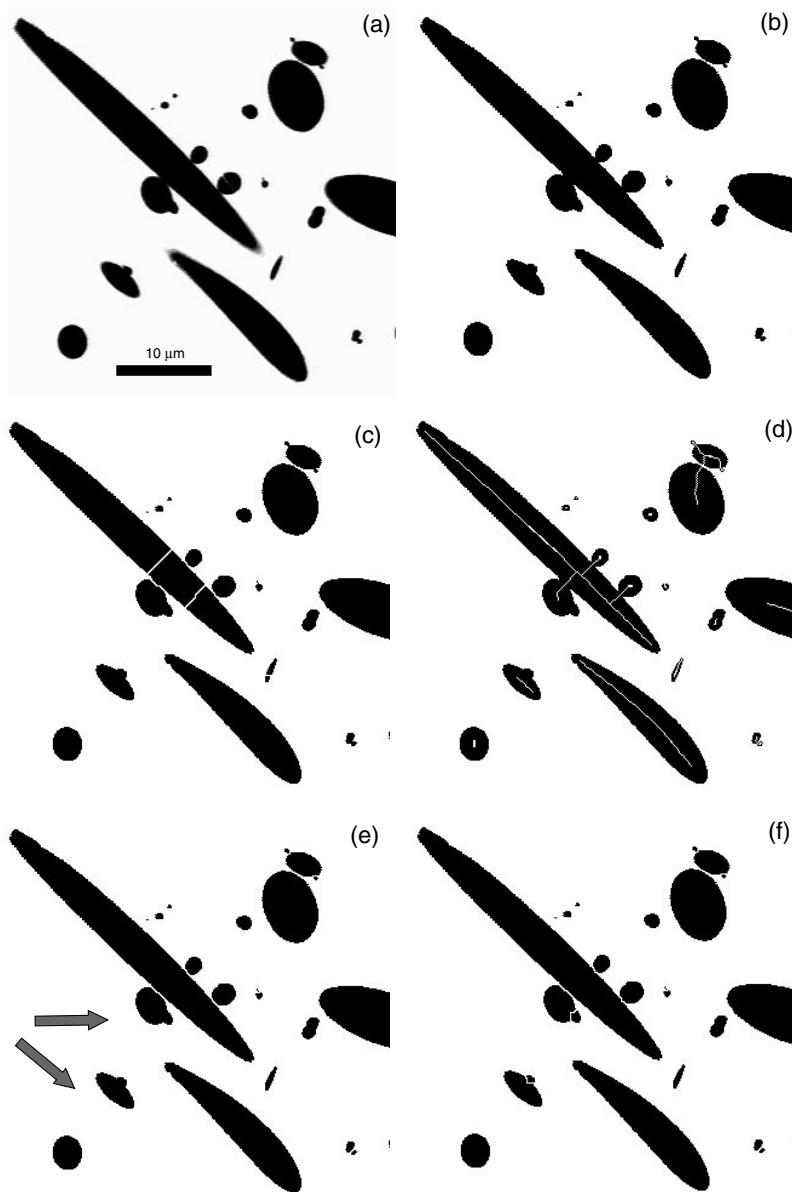


Fig. 2. Example of segmentation of MMVF in polished section: (a) is the original image; (b) is the thresholded and pre-processed image; (c) is a tentative segmentation using a filtered ultimate erosion; (d) shows the skeleton of the fibres; (e) shows the better segmentation obtained with this technique; and (f) is the best segmentation using the bisector function, which even separates the fibres shown by an arrow in (e).

individual object, the diameter of the corresponding fibre is that of the largest inscribed disk (largest disk that can entirely fit inside the object). This diameter is obtained by doubling the largest value of the distance transform found inside the object.

However, this simple measurement is biased in several ways:

- Objects cut by the border of the field of view cannot be measured in general. Larger objects stand a better chance of intersecting the border of the field of view than smaller ones. This generates a bias towards the smaller fibres that needs to be addressed. The classic way of dealing with this problem is the Miles–Lantuéjoul correction (Miles, 1974) and (Lantuéjoul, 1980).

- The fibres seen in a polished section were cut perpendicularly to their axis. This means that longer fibres have a greater chance of being visible in the polished section than shorter ones. As a consequence, the diameter histogram obtained by accumulating all the fibres diameters is statistically length-weighted. There is no guarantee that the orientation angle distribution of all the fibres is the same for each diameter class. This generates a bias which can be corrected by weighting each object by the inverse of the cosine of the orientation angle. This cosine is simply derived by computing the area of each object divided by its diameter (see Fig. 3).

With this method, 500 fibres can be measured per hour; see section 3 for actual measurement results.

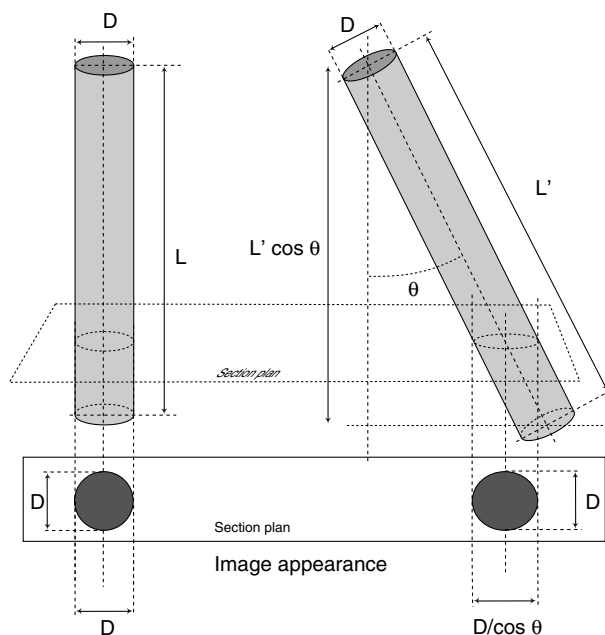


Fig. 3. The diameter histogram of the fibres is length-weighted. The orientation angle of each fibre must be taken into account. This angle is obtained easily by dividing the area of each fibre by its diameter.

2. Flat-bed fibre method

If information on the length of the fibres is desired, a new sample preparation method is necessary. Fibres are deposited on a flat surface and observed directly from above. We call this method the flat-bed method.

2.1. Sample preparation

In contrast to the previous method, the sample preparation is significantly simpler. A tuft of binder-free MMVF is extracted from the material. A small quantity is obtained with a punching tool (about 100 mg). If length-weighted measurements are sought, the sample can also be crushed. This sample is dispersed in an alkaline solution (pH 10).

A NucleporeTM filter (polycarbonate filter) with a mean hole diameter around 0.08 μm is lightly coated with sputtered gold or gold-palladium. The metal thickness should be between 10 and 30 nm. Perfect conduction at this stage is not necessary. The fibre suspension is then filtered through this filter. This filter with the fibres deposited on it is coated a second time with evaporated nickel. The filter needs to be normally conductive after this second coating. Finally, the double-coated filter is observed in a standard SEM in BSE mode.

A single, final coating with either gold or carbon is unsuitable for this kind of sample. Gold coating necessitates secondary electron imaging, which in this case is too noisy because of the presence of holes in the background. Carbon coating is insufficient because of the geometry of the fibres

and because of the lack of resolution in BSE mode (none of the elements is heavy enough), as Fig. 4(a) shows.

This unusual double coating is useful first to guarantee a good enough resolution in BSE mode (this is the role of the first heavy-metal coating) and second to avoid modifying too much the mean local atomic number of both the fibres and the filter while coating the sample correctly on all surfaces (that is the role of the nickel). BSE imaging reduces the level of noise around the holes of the filter and increases the contrast between the fibres and the filter in the background because of the chemical contrast. Figure 4(b) shows the result of such a preparation observed in a Zeiss DSM 950 SEM.

An even simpler preparation method can be used if an environmental SEM (ESEM) (Danilatos, 1988) can be used in lieu of a standard SEM. The sample is prepared in the same way with none of the coatings, and observed in an atmosphere of water vapour with the environmental secondary electron detector. The images obtained in this way are of better quality because of the resolution of this type of imaging and also because of the presence of chemical contrast unhindered by any coating. Figure 4(c) shows the resulting image of such a sample preparation observed in an Electroscan prototype ESEM at magnification $\times 1000$ (Talbot *et al.*, 1992).

An optimum mass of fibres to be deposited on the filter can be achieved experimentally by trial and error in order to minimize the number of overlapping fibres as well as the number of images.

Compared with the sample preparation of the polished section method, experience shows that a single sample can be prepared in minutes rather than hours and this preparation is much less demanding to the operator, making it more reproducible.

2.2. Image analysis algorithms

In contrast to the polished section method, the analysis of the flat-bed images is significantly more difficult.

Fibres present on images such as those in Figs 4(b) and (c) cannot be segmented as a whole. One has to differentiate the large fibres from the small fibres. The large fibres are those which appear as plain, saturated, white, elongated, cylindrical objects on the images. The small fibres appear as thin line-like, unsaturated, white, elongated objects.

Experimentally, large fibres at working magnification of $\times 1000$ cover the diameters 0.8–20 μm , and small fibres are those smaller than about 0.8 μm , down to about 0.15 μm . The smallest fibre we have recorded across a large number of samples was about 0.08 μm in diameter; fortunately such dimensions are rare with MMVF, as they are not easy to detect even at high magnification.

2.2.1. Detection of the large fibres. The large fibres are quite simple to segment as compound objects (crossing or touching fibres will not be separated yet). A standard region-growing

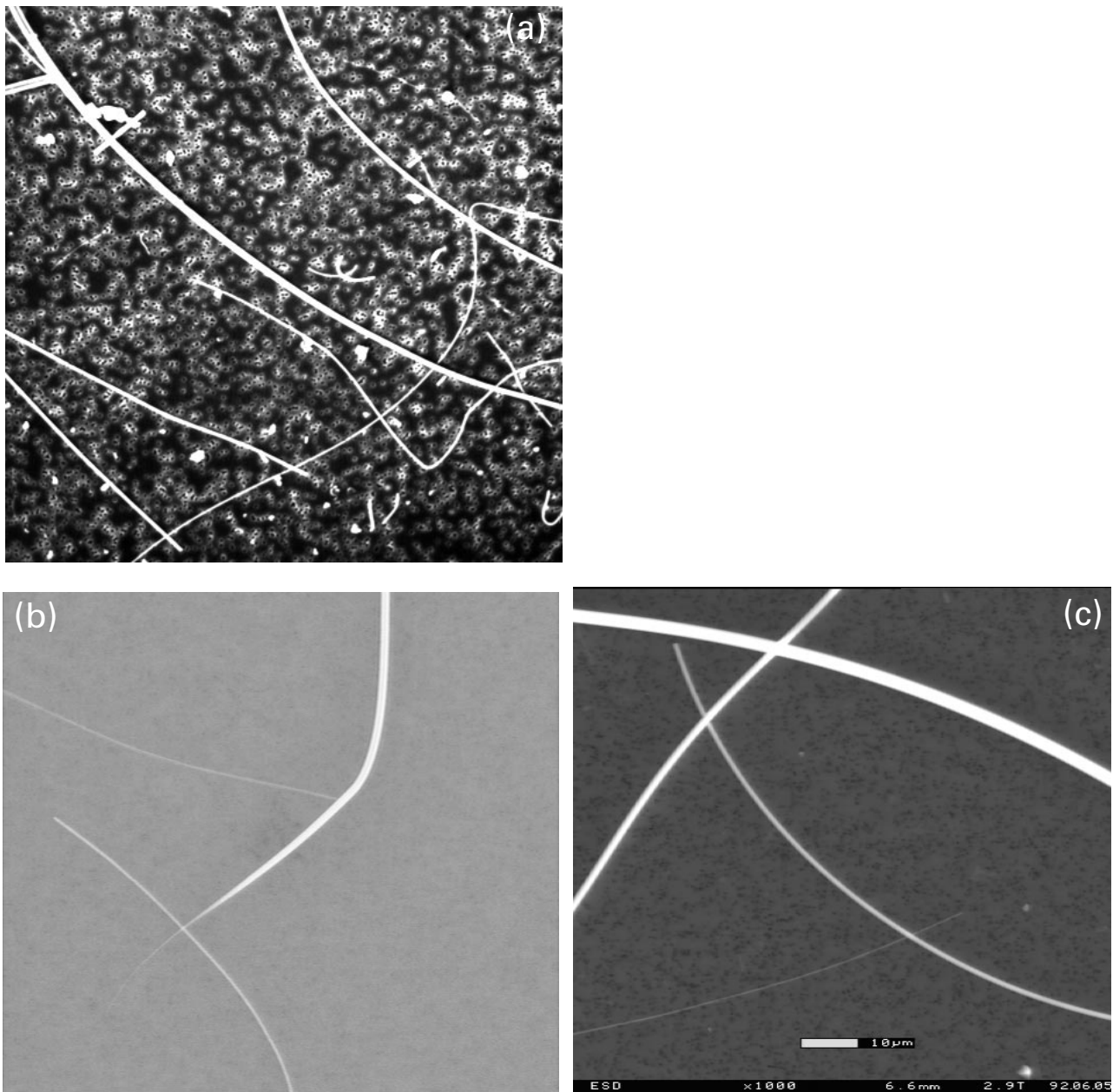


Fig. 4. Flat-bed method: sample preparation. Image quality of (a) is insufficient; noise is due to the holes in the filter in the background. Image quality of (b) and (c) is acceptable; (b) is obtained in a normal SEM, (c) can only be obtained in an ESEM.

technique such as the watershed algorithm as proposed by Meyer & Beucher (1990) can be applied easily, as these fibres and the background are sufficiently contrasted.

A marker internal to the fibres can be obtained by simple adaptive thresholding of the 10% brightest pixels present in the image above a given, fixed threshold (to account for the case when no large fibre is present). Experimentally, these will always be located inside the large fibres.

A marker external to the fibres can be found by segmenting the holes present in the background of the

filter. This can be carried out by adaptive thresholding under a fixed limit (to take care of the extreme case when the image is covered entirely with large fibres).

The Sobel gradient is used for finding the contours of the fibres. The application of the watershed algorithm yields a reliable segmentation of the large fibres, as shown in Figs 5(a) and (b).

2.2.2. Detection of the small fibres. Thin fibres are thin, white, locally straight objects and much more variable in

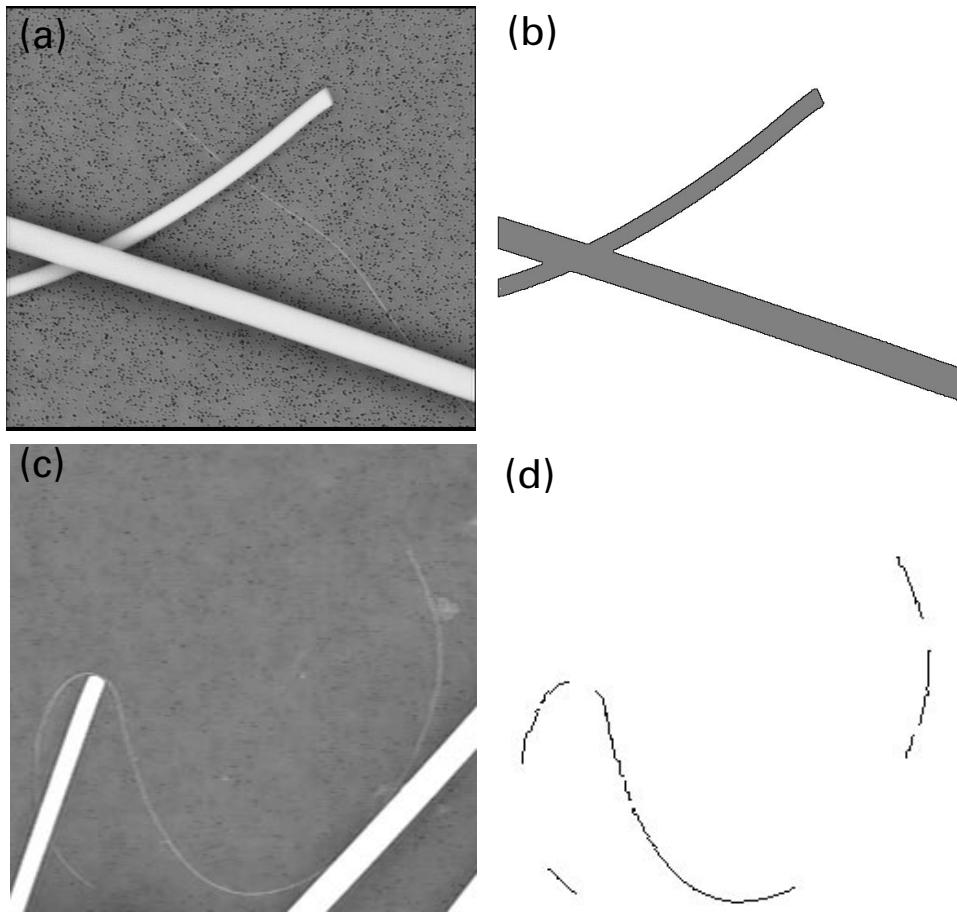


Fig. 5. Detection of large fibres ((a) and (b)) and small fibres ((c) and (d)).

appearance than large fibres. Some of these fibres are barely visible over the noise of the background of the image. For example, serendipitous alignment of pores of the filter in the background can create the illusion of small fibres.

A specific procedure proposed by Talbot (1996) was designed to segment this kind of linear feature. Its principle is to follow the ridge lines of the grey-level image in an iterative fashion. As these linear features are a succession of regional maxima and saddle points, this iterative procedure will tend to join regional maxima along the ridge line. As a result, only a small number of iterations (about 15 in this case) is sufficient to join all the relevant regional maxima, and noise features which are still detected with this method tend to be short, and can thus be deleted easily. The result of this procedure is illustrated in Figs 5(c) and (d).

2.2.3. Directional reconstruction. For the image segmentation to be completed, one still needs to deal with several important issues: the large fibre detection procedure yields compound objects; crossing and parallel fibres need to be separated; in contrast, the small fibre detection procedure

yields unconnected markers; markers belonging to a specific fibre need to be connected.

These two different problems can actually be treated in the same way, using local directional information.

General strategy for directional reconnection. The general problem that we want to solve is the following: in a 2D image, we have a finite number of features with which a number of markers are associated. We do not know *a priori* which sets of markers correspond to which features. We assume these markers to be relatively straight and elongated, and that the orientation of these markers is relevant. We want to connect together those markers belonging to individual features. Methods based on directional propagation on binary images (Kurdy & Jeulin, 1989) or on graphs (Viero & Jeulin, 1995) were proposed to connect portions of lines or of curves. A different approach is followed in the present study.

Under the above assumptions, Lee & Talbot (1995) have proposed the following method to solve the problem, also illustrated by Fig. 6.

First the markers are simplified by computing their best-

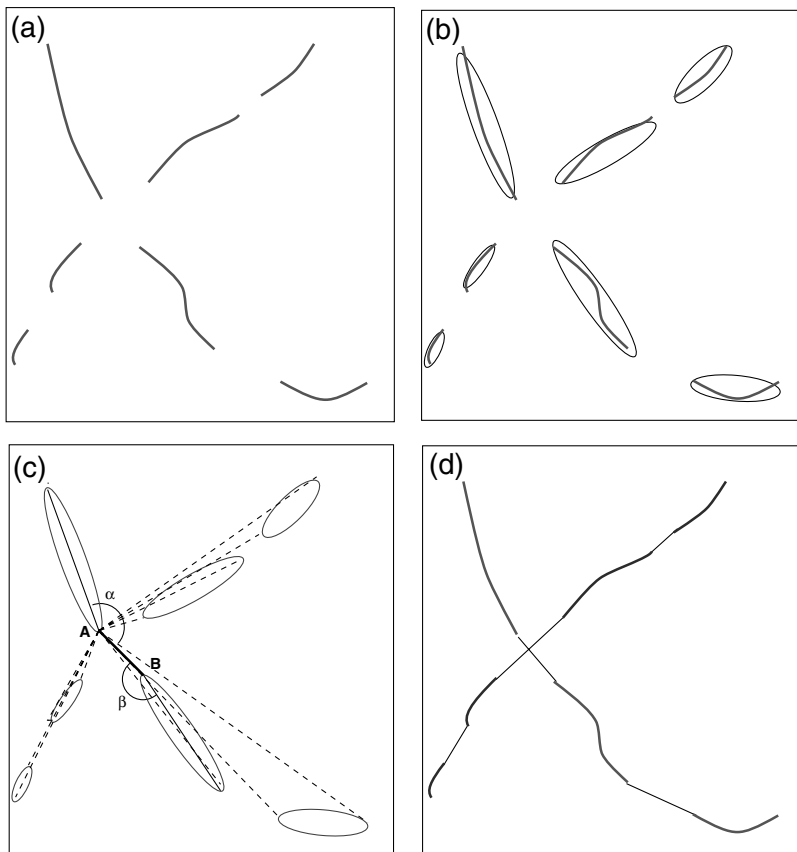


Fig. 6. General reconnection procedure. Unconnected fibre markers are shown in (a). A moment-based method is used to fit ellipses to markers (b). Distance and directional information from the ellipse parameters are used to test and rank reconnection possibilities (c). Final reconnection is shown in (d).

fitting ellipse according to their moments (Rosenfeld & Kak, 1982) (Fig. 6b). Each major axis extremity of the ellipses (starting points) are matched with all the other available extremities (ending points). Figure 6(c) depicts this procedure with point A as a starting point. Attribute a score to each match according to the following score function:

$$\text{Score} = \lambda \sqrt{\frac{2\pi - \alpha - \beta}{2\pi}} + (1 - \lambda) \frac{d}{\text{diag}}, \quad (1)$$

where α and β are the angles between the line constituting the possible reconstruction and the main axis of each of the ellipses, d is the Euclidean distance between the two extremities, diag is the length of the diagonal of the image and λ is an adjustable parameter. This function has a higher value when the distance between the two extremities decreases, and when the angle between the reconnection line and the main axis of both ellipses tends to π , i.e. when the ellipses are well aligned. The reconnection scores are ordered with best scores first. The potential reconnections are then performed in order without replacement. If no possible reconnection reaches a sufficient score, no reconnection is performed. Also, when two extremities are too close to each other, the reconnection is performed regardless of the angle.

Reconstruction of large fibres. Large fibres detected according to the procedure described in section 2.2.1 may be crossing other large fibres, in which case the detected object is a binary mask which is the union of the two or more crossing fibres.

In order to retrieve the individual fibres, it is first necessary to find if and where a crossing of two or more fibres has occurred. This is done by taking the skeleton of the binary mask and by finding its multiple points (i.e. its branch points). To perform this task, Euclidean skeletons are preferable to standard skeletons obtained, for example, by thinning procedures (Arcelli & Sanniti di Baja, 1985) because this specific skeleton is insensitive to the orientation of the fibres. Moreover, only a subset of the full skeleton is actually used, called the minimal skeleton. This skeleton is far less sensitive to boundary noise. Talbot & Vincent (1992) have published a good way of obtaining the Euclidean skeleton (both full and minimal) of binary sets.

If a crossing point has been detected, each branch of the skeleton is broken in a series of shorter segments on which ellipses can be fitted. There is indeed no guarantee that skeleton branches be straight enough for their global orientation to be meaningful. These markers are then reconnected according to the general strategy presented in this section. This procedure is illustrated in Fig. 7. The gap

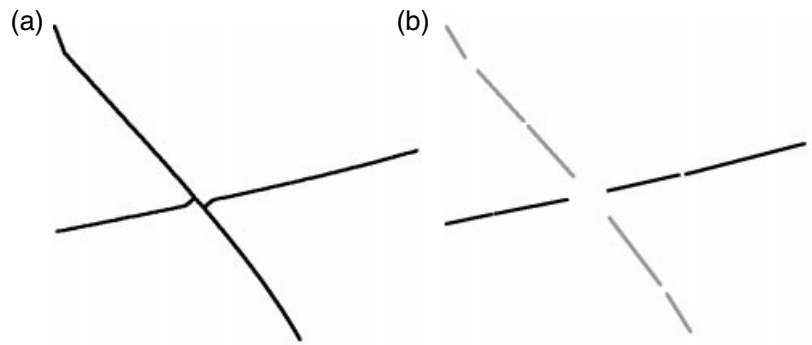


Fig. 7. Crossing management. The original image and its large fibre segmentation are shown in Figs 5(a) and (b). The Euclidean skeleton is used to extract the crossing information (multiple points of the skeleton) and the directional information of each branch (a). The fibre markers are labelled correctly (b).

between segments is smaller than across crossing points because the directional information given by the skeleton is unreliable near crossing points.

For this particular problem, the score function presented in Eq. (1) can be made more complex, for example, by incorporating information about the diameters. Obviously, the two parts of a fibre across a crossing point must have the same diameter. In practice, however, this approach can create more difficulties than it can solve, and the simple approach presented above seems to yield reasonable results already.

Reconstruction of small fibres. Small fibre markers detected by the procedure described in section 2.2.2 also need to be reconnected according to their directional information. As for large fibres, small fibre segments may need to be broken up into smaller segments.

In order to reconstruct the full length of all the small fibres, a dual marker approach is proposed: a first set of small fibre markers (set A) is found, using conservative parameters such that their probability of belonging to a small fibre is high. A second set (set B) is also found, this time using less conservative parameters, such that their probability of belonging to a small fibre (as opposed to noise) is lower. Parameters are set such that $A \subset B$. Of course, some noise is also detected in set B .

For the general reconnection procedure, starting points must belong to A , whereas intermediate or ending points can belong to both A and B . Figure 8 illustrates the proposed procedure on a particularly noisy image.

2.2.4. Detection of parallel fibres. The fibre fabrication process sometimes generates parallel fibres. Such collections of fibres must not be mistaken for single, larger fibres. Most true parallel fibres are large fibres to begin with. For physical reasons, small fibres do not tend to form long parallel structures.

The procedure for finding parallel fibres is therefore to look for elongated, thin, dark features inside large fibre binary masks as detected by the procedure described in section 2.2.1. This procedure is actually similar to the procedure used to detect small fibre markers, but on a negative image, and inside

a smaller portion of the image. These markers are called p-markers. The presence of at least one such marker inside the binary mask of a large fibre triggers a series of procedures which alter the way the diameter of this fibre is measured. This procedure is detailed in section 2.3.1.

2.2.5. Field rejection procedure. Even with the most sophisticated image analysis procedures, some images may be too difficult to segment. The range of reasons why this may happen is quite large, but include image acquisition conditions, sample preparation or just serendipitous fibre arrangements that even humans would have difficulties deciphering.

In this case, it is wise to ignore the image field. This is not as easy to achieve as it seems, precisely because of the large number of possible cases. However, a number of heuristics can be used to decide when a field should not be considered for measurements. The grey-level histogram of the image should only present at most three modes (for background, large fibres and possibly small fibres). Detecting too many objects (depending on the sample preparation) on the image indicates that something is wrong with the field under study. Consistency checks are also a good idea, such as checking that some small fibres were indeed detected if a third mode is present in the grey-level histogram.

Nevertheless, no set of heuristics can be complete enough to guarantee reliable measurements in all cases. We therefore propose to keep the measurement of all detected objects, whether or not the automated procedure indicates it was obtained in good conditions, and to keep 'reliability' information with all the measurements. This way, at the end of the segmentation/measurement run, one can construct the summary statistics (histograms, or mean values) both for reliable fields only and for all fields (reliable + non-reliable). If the two statistics are significantly different, then unreliable fields can be discarded.

2.3. Measurements and statistical analysis

From the micrography set, we intend to obtain two histograms: first a diameter histogram, and second a mean

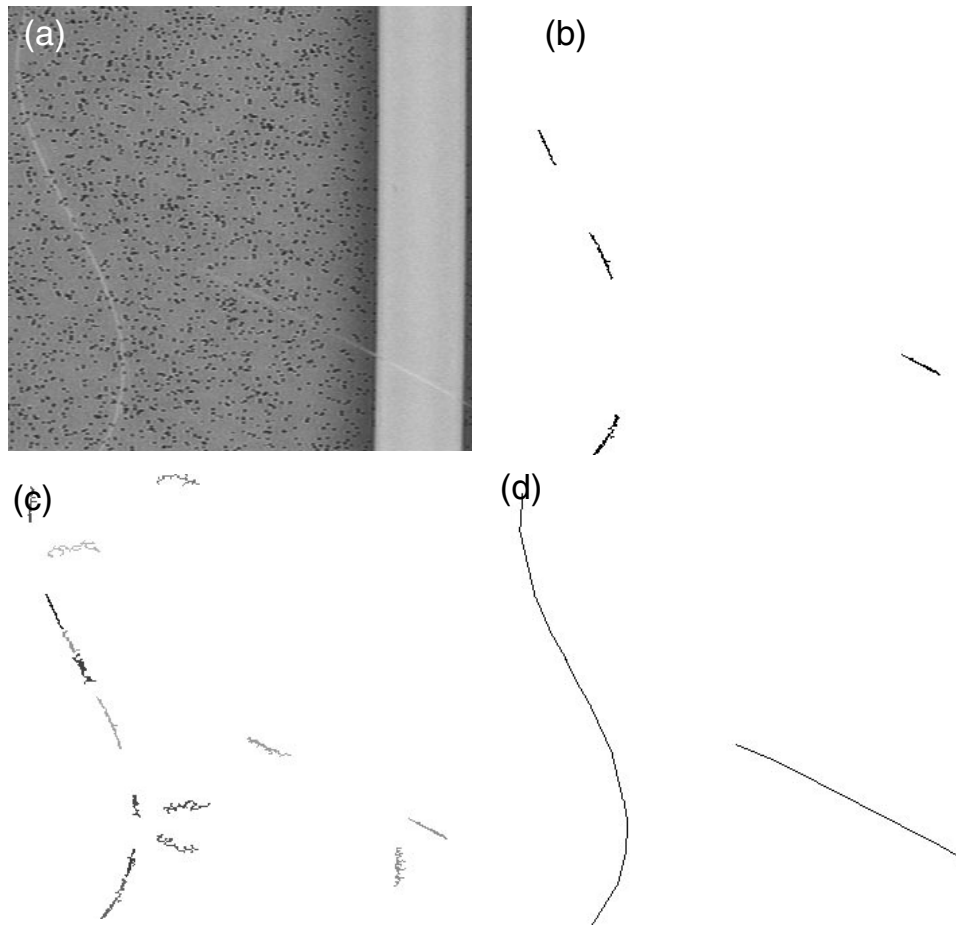


Fig. 8. General reconnection procedure for small fibres. Original image (a). Two levels of marker: reliable markers (b) and potential markers (c). Potential and reliable markers are both used to reconstruct the small fibres, using only reliable markers as starting points.

length vs. diameter histogram. Diameters can be obtained directly, but mean length cannot (most fibres are much longer than the field of view). Moreover, diameter measurements are biased. Unbiasing the diameter measurements and obtaining the mean length of fibres by diameter class can be achieved simply by measuring the length of the portion of the fibres visible in the field of view, and counting the number of visible extremities of each fibre (Talbot & Jeulin, 1993). We therefore need to make three different measurements on each fibre: its diameter, its visible length and the number of visible extremities.

2.3.1. Diameter measurement. Diameter measurement is probably the most important, as it is the one most routinely performed manually, and the one against which manual and automated measurement can be most easily compared. The highest precision is expected from this measurement.

The two most important cases are:

Thick individual fibres. These fibres are sized by taking the mean value of the Euclidean distance transform along the

minimal Euclidean skeleton. This yields the average radius of the fibre. Using this method, the radius is always a bit underestimated, due to the fact that the segmentation of the big fibres is obtained by growth of an internal marker up to a gradient maximum. Human operators tend to place diameter measurements a bit further away from the centre of the fibre. By experience, adding two or three pixel widths to the diameter obtained by the distance transform yields a measurement which is closer to the one given by an operator.

The mean diameter and the diameter variance are computed for each thick fibre marker (i.e. for each marker for which an ellipse is fitted). This information is used several times during the segmentation process. In particular, segments with a high variance are excluded from the diameter measurement process, as they most likely belong to irregular (for example broken) bits of fibres.

Thin individual fibres. The diameter of these fibres is simply assumed to be one pixel. No actual diameter measurement is made, first because it would not be precise enough, and

second because this precision is not needed at usual magnifications ($\times 1000$) and a histogram resolution of $1.0 \mu\text{m}$: all the thin fibres belong to the 0–1 μm diameter class.

Another important case is the measurement of parallel fibres. Each marker of parallel fibres (p-marker) is represented by the major axis of its best-fitting ellipse and is associated with three easy-to-compute measurements: diameter (average of the distance transform value); left-radius (average distance between the p-marker and the left fibre edge) and right-radius (same as left-radius, with the right fibre edge).

Only those radii which are of low variance are worth considering. Hence, we need a further step for filtering p-markers: either both radii are of low variance and two fibres are parallel to each other, in this case we take the larger one as the measurement; or one radius is of low variance while the other is of high variance, meaning two fibres overlap at a small angle without crossing. In this case we keep the low variance radius as the measurement.

2.3.2. Length measurements. Only the length of the visible part of each fibre can be measured. The method is the same for thick or thin fibres, but the implementation of the method is more difficult for the thick fibres than for the thin ones.

The length measurement of individual fibres presents few difficulties. The distances between best-fit ellipses are known. Their major axis lengths are known, and the distance to be added at the extremities are equal (by construction) to the radius of the fibres. The length measurement of crossing fibres is more complicated, because around the point of crossing, diameter and directional information derived from the skeleton is unusable. Therefore, this area must be removed from the skeleton, and a bigger gap than usual appears between ellipses across the crossing point. However, the diameter of the region that has been removed is known.

For thin fibres, this is all that needs to be done, as there is no significant masking effect. However, for the thick fibres, when the angle between the crossing fibres is small, the situation shown in Fig. 9 can arise. The best-fit ellipses are derived from the Euclidean skeleton information. If the crossing angle is sharp, the skeleton of the union of two fibres can exhibit two multiple points. If the angle is sharp enough, an incorrectly orientated line segment can appear. This segment is effectively eliminated by the fact that the diameter measurement along this segment will have a high variance, but this situation makes the length measurement procedure more difficult. If a reconnection does take place across the crossing point, the gap length is estimated by the Euclidean distance between the two

joined points. This tends to underestimate the true length by a small amount.

Many other difficult cases have been observed. For example, the extremity of a fibre can be masked by another fibre. In this case, a multiple point appears in the skeleton, but no reconnection is made across the crossing point for this fibre (however, it does occur for the other fibre). Only the radius of the region that has been removed around the multiple point of the skeleton is added to the length of the fibre with the invisible extremity.

2.3.3. Measurement of the number of visible extremities. We show in section 2.4 why this measurement is useful. A fibre extremity is defined as any extremity of a major axis segment of an ellipse derived from skeleton information that has not been reconnected to another major axis segment extremity. Due to the ellipse fitting strategy, all fibres have two such extremities visible in the entire field of view.

To minimize measurement bias, all diameter and length measurements are made in a subset of the field of view. This subset is different for thin and thick fibres. For thick fibres it is obtained by removing 40 pixels on each side of the image. For thin fibres, the subset is obtained by removing only seven pixels on each side. Of course, the difference in measurement area between small and large fibres must be taken into account when building the diameter and length histograms. In either situation, fibre extremities detected inside the subset are counted as true extremities, and those falling outside as false extremities.

Let f be the number of false extremities and t the number of true extremities for a fibre F . The number of true fibre extremities is obtained by:

$$t = 2 - F \quad (2)$$

For the population of fibres examined, Eq. (2) is summed over all the reconstructed objects.

2.4. Statistical analysis of the results

The direct measurement of fibre diameters is not sufficient to be able to produce unbiased diameter histograms. In the same fashion as for the cross-section method, larger fibres stand a better chance than thinner fibres of being intersected by the field of view in such a way that their diameter cannot be assessed. In addition, longer fibres are more likely to fall into a given field of view than shorter ones. If the length distribution is not the same across fibre diameters, another bias is generated. A solution to this problem was proposed by Talbot & Jeulin (1993).

2.4.1. Unbiased diameter measurements. We suppose the fibre distribution is stationary, that a background is visible and that we observe this distribution through a series of

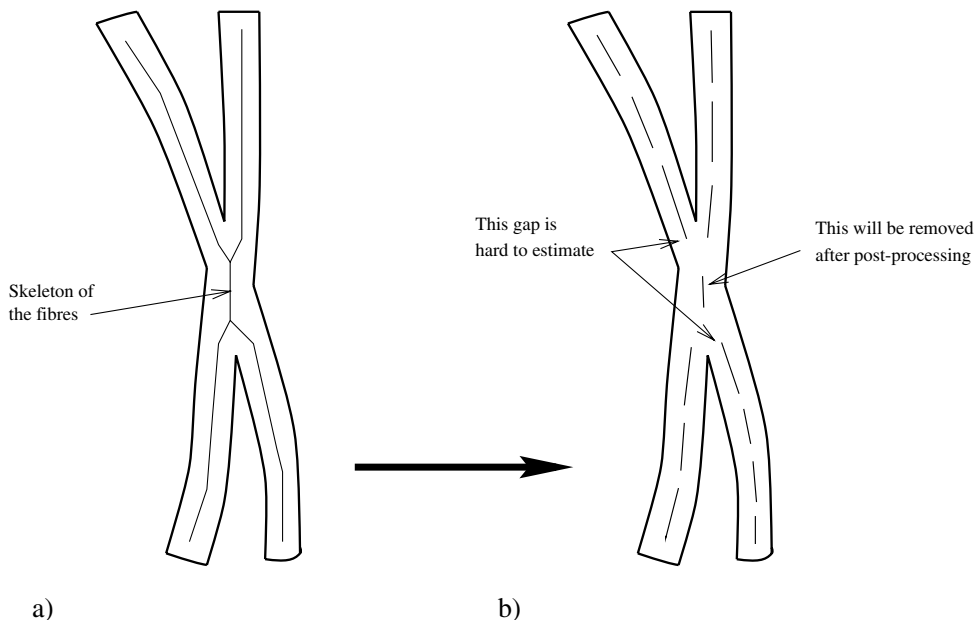


Fig. 9. Length measurement for fibres crossing at an acute angle.

fields of view. Finally, we suppose that we know the position and visible dimensions of the fibres in each field of view.

Notation: In this section, D is a diameter, L is a length, A is an area, \bar{X} is the mean value of X .

First method. Talbot *et al.* (1996) have shown that an unbiased estimate of the number of fibres by diameter class, $N(D)$, can be obtained by accumulating the number of visible extremities of fibres of diameter class D in each field of view. The histogram of $N(D)$ is an unbiased estimate of the histogram of the number of fibres by diameter class when the Miles–Lantuéjoul correction is applied.

From there, the mean fibre length by diameter class is estimated by:

$$\bar{L}(D) = \frac{2A(D)}{DN(D)} \tag{3}$$

where $A(D)$ is the cumulated visible area in each of the fields of view of the fibres belonging to the diameter class D .

This method is illustrated in Fig. 10, where a fibre is observed through nine consecutive fields of view.

Second method. Talbot *et al.* (1996) have also shown that the fibre diameter distribution can be unbiased by working out the probability that a given fibre intersects a field of view of known dimensions. An unbiased estimate of the diameter distribution can then be derived according to the following equation:

$$\theta f(D) = \frac{N^*(D)}{A(Z) + D\bar{L}(D) + \frac{1}{\pi}L(Z)[D + \bar{L}(D)]} \tag{4}$$

where θ is the fibre density, $f(D)$ is the frequency of fibres of diameter D , $N^*(D)$ is the biased diameter frequency obtained

by accumulating all the fibres visible in each field of view of diameter D regardless of the number of visible extremities, Z is the field of view, $L(Z)$ is its perimeter and $A(Z)$ its area (both assumed constant for each class of diameter). Finally, $\bar{L}(D)$ is the mean fibre length by diameter class, as estimated above. θ is given by normalization.

These two methods are easy to implement. Implementing both of them allows a consistency check to be performed.

2.4.2. Obtaining weighted diameter histograms. Once the unbiased histogram of fibre diameters and the mean length by diameter class have been obtained, several interesting and useful histograms can be derived from these equations:

$$\theta f_l(D) = \frac{\bar{L}(D)}{\bar{L}} N(D) \tag{5}$$

$$\theta f_A(D) = \frac{\pi/2D^2 + D\bar{L}(D)}{\sum_d (\pi/2d^2 + d\bar{L}(d))} N(D) \tag{6}$$

$$\theta f_V(D) = \frac{D^2\bar{L}(D)}{\sum_d d^2\bar{L}(d)} N(D) \tag{7}$$

where θ is the fibre density, $\bar{L}(D)$ is the mean length for fibres of diameter D , \bar{L} is the mean length of all the fibres, d is the diameter classes over which we sum in Eq. (5), $f_l(D)$ the length-weighted fibre frequency of diameter D . In Eq. (6), $f_A(D)$ is the area-weighted fibre frequency of diameter D . Finally, in Eq. (7), $f_V(D)$ is the volume-weighted fibre frequency of diameter D .

We do not need to compute \bar{L} or θ , as their product can be

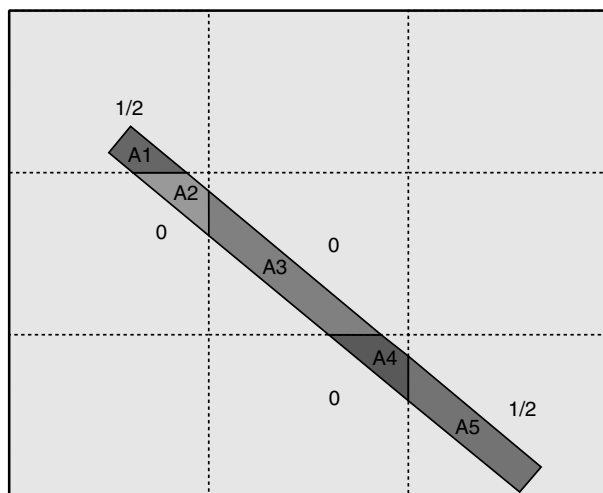


Fig. 10. The length of a fibre of diameter D can be found by accumulating visible extremities and visible fibre area over contiguous fields of view. In fields where a fibre extremity is visible, half of a fibre is counted and the visible area of this fibre is accumulated. In fields where a fibre is visible but no extremity is found, no fibre is counted, but the visible area for this fibre is accumulated. In this figure, the total number of fibres is therefore 1 and the area of the fibre is $A = \sum_i A_i$. If fields of view are not contiguous, the mean length can still be estimated, and the number of extremities encountered yield an unbiased estimate of the frequency of fibres of the measured diameter.

given by normalization. The length-weighted histogram is useful as it is far less sensitive to the way the fibres are handled prior to and during the sample preparation. Indeed, MMVF are fragile and can easily be broken into shorter segments. This histogram will not be affected by this phenomenon. It should be therefore more reproducible than the unweighted histogram.

The area-weighted histogram can be useful in studies of specific area of materials made of MMVF, and the volume-weighted histogram is the same as the mass-weighted histogram. It shows what kinds of fibres constitute the bulk of MMVF-based materials.

3. Results

The two automated methods were tested on different products for comparison with manual measurements, and for reproducibility issues.

A typical run on a given product for either method consists of analysing enough images to accumulate measurements on 1200 fibres in two batches of 600 measurements made on two different samples. A χ^2 goodness-of-fit test is conducted between the two distributions and if the difference between the two batches is not found to be significant at the 95% level of confidence, a third run of 600 measurements on a third sample is conducted.

3.1. Comparison with raw manual measurements

In this section, the automated counting and the manual counting were done on precisely the same fields of view. The manual counting was performed slowly and carefully by a trained operator, so that we can consider it as a ground truth. Although it is not practical to compare individual measurements between manual counting and automated counting, it is possible to compare the histograms.

3.1.1. Polished section method. Figure 11 shows a large (1600 fibres) diameter measurement run made both manually and automatically on the same image data (a product designated as IBR). The raw, biased data were used in this survey as it is only used for comparison purposes.

As is apparent on this figure, the difference between manual and automated counting is small. The χ^2 score (19.54) is well below the 95% critical value (24.99). For significance, all fibres greater than or equal to $16 \mu\text{m}$ in diameter were bundled together.

3.1.2. Flat-bed method. Figure 12 depicts a typical (600 fibres) diameter measurement run made both manually and automatically on the same image data (a product designated as MIG). Only the raw, biased data are shown in this figure as it is only used for comparison purposes. Figure 13 shows the histogram of the mean visible length per diameter class both for manual and automated measurements made under the same conditions. The measurements are remarkably similar. For both diameter and length measurements, the χ^2 values (6.49 and 9.73, respectively) are well below the critical value (16.92 for both), indicating that the differences between manual and automated measurements are not statistically significant. For this score, all fibres greater than or equal to $10 \mu\text{m}$ in diameter were bundled together (for significance).

The high correlation between manual and automated measurement illustrated in Fig. 14. To create this figure, three times 600 diameter measurements were performed on three different products with different kinds of MMVF: an IBR product, an MIG product and an unusual CM25 product with very fine fibres (mean diameter around $0.5 \mu\text{m}$), not normally used for insulation purposes. This last product was brought in to test the limits of the methods. For this figure, all measurements were made both manually and automatically on the same image data. This plot shows the diameter class frequencies measured automatically vs. the frequencies of the same diameter classes measured manually. As this diagram shows, the correlation is high, except for two data points corresponding to very small fibres. For these two data points, the chosen image resolution was insufficient to measure the corresponding fibres with enough accuracy, either manually or automatically.

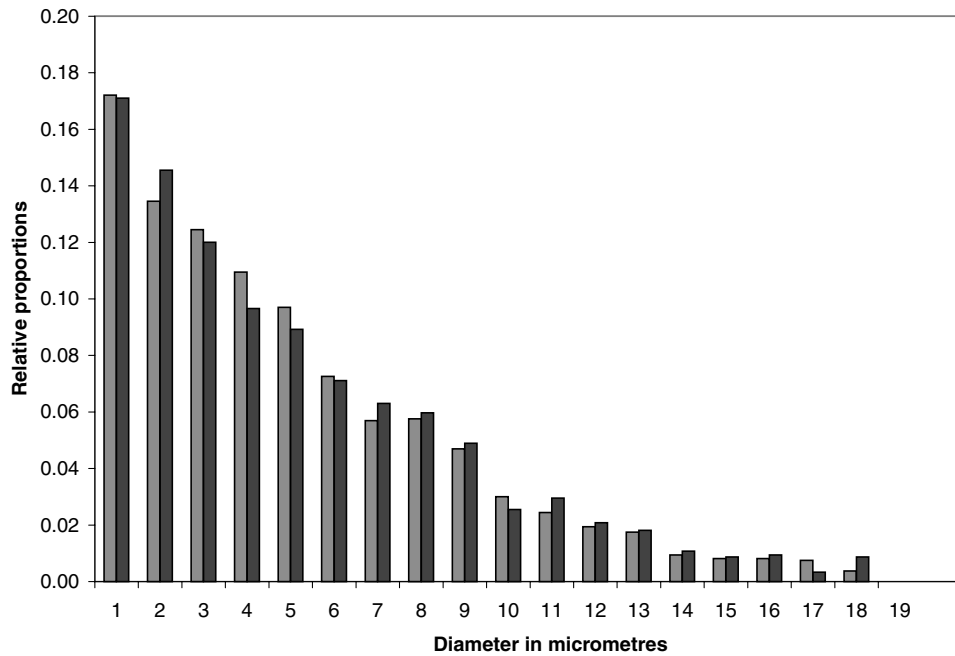


Fig. 11. Typical polished section method raw diameter measurement histogram; manual vs. automated measurements. See Table A.1 for numerical values.

3.1.3. Comparison between automated and manual measurements. It is also interesting to compare both automated methods, as well as testing the unbiasing techniques described in previous sections. Unfortunately, it is impossible to compare

these results to a ground truth. The best that can be done is to provide unbiased manual measurements using standard stereological techniques, which cannot be made on the same samples, contrary to the previous cases. On the other hand, this

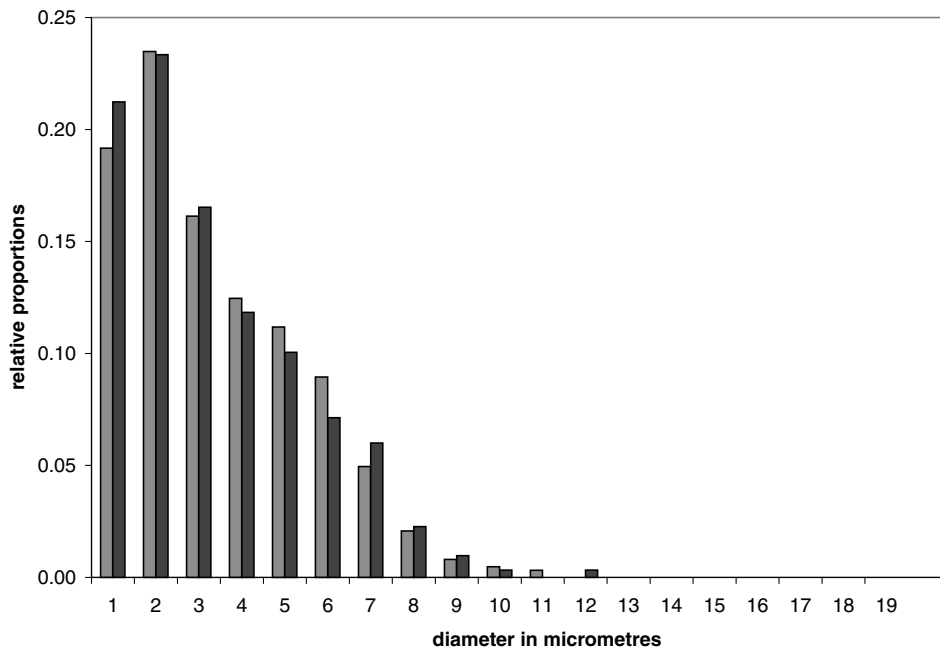


Fig. 12. Typical flat-bed method raw diameter measurement histogram; manual vs. automated measurements. See Table A.1 for numerical values.

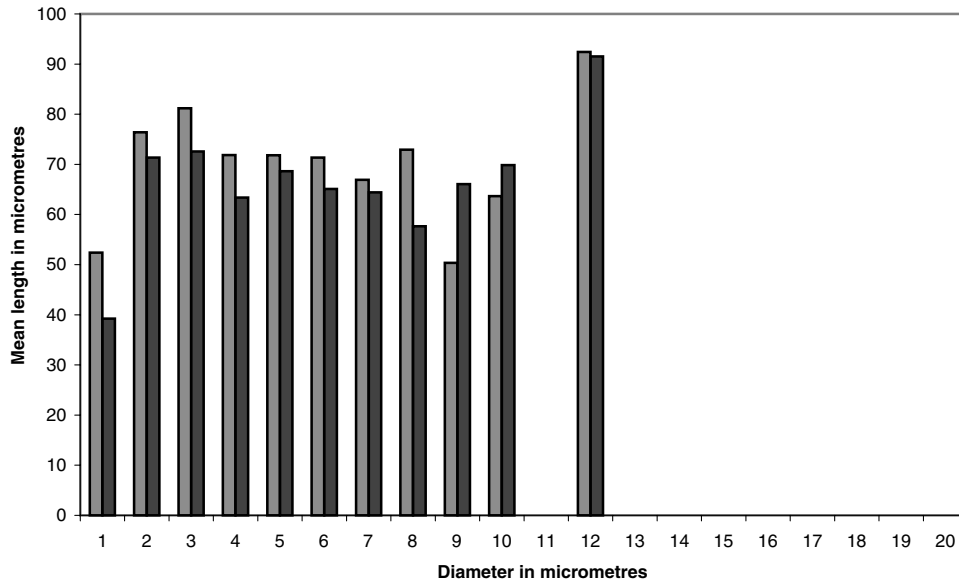


Fig. 13. Visible length measurements; manual vs. automated. See Table A.1 for numerical values.

comparison in some sense tests the whole chain from sample preparation to the unbiasing techniques via image analysis.

The comparison has to be made using Eq. (5). An unbiased manual length-weighted histogram can be obtained using the so-called TIMA method (Koenig *et al.*, 1993), which involves counting only fibres that touch or cross a horizontal line across the middle of the field of view, on a flat-bed fibre layout.

The three independent distributions are shown on Fig. 15. They look similar, but we need to test that they belong to the same underlying distribution. For this, one can use the contingency matrix χ^2 test (Fienberg, 1983). As before, the counts for $D \geq 10 \mu\text{m}$ are all put together. The result of this test is shown in Table 2. In this table, we show the three-way test as well as the three two-way tests. All the tests are passed, and the manual vs. flat-bed comparison is the best in this instance.

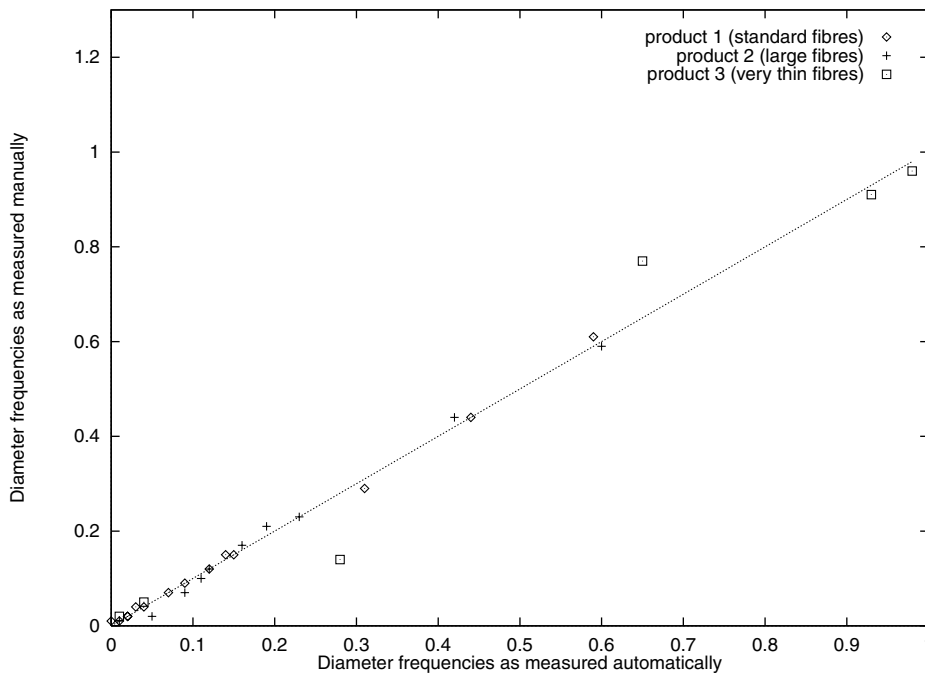


Fig. 14. Plot of the diameter class frequencies; automatic vs. manual.

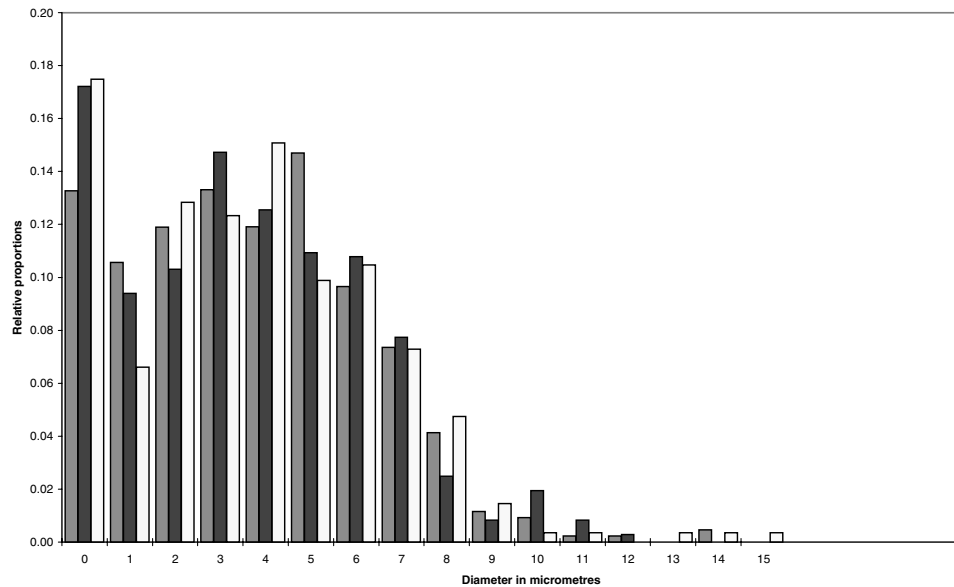


Fig. 15. Three-way comparison between manual counting, the flat-bed and the cross-section methods. All three histograms are length-weighted. The manual counting does not provide a ground truth in this case. See Table A.1 for numerical values.

4. Discussion

In terms of precision, both automated methods compare well with manual methods for mainstream insulation MMVF diameter and length measurements. For smaller fibres, a change of magnification would be necessary, together with some parameter adjustments, for the automated methods to work reliably.

The polished section method is effective in terms of the number of fibres that can be measured in a given time period. However, the diameter histogram produced by this method can only be length-weighted, and produces little additional information on the MMVF product. Finally, the sample preparation is rather long and complex, and it is difficult to ensure its reproducibility.

On the other hand, the flat-bed method is much more difficult to implement in terms of image analysis algorithms and methods, because the images are much more variable and subject to noise. However, it does produce much more information about the product under study. In particular, the length information is interesting, because it is difficult and time-consuming to obtain manually. In addition, the sample preparation is simple and fast. Table 1 provides an estimate of sample preparation and image analysis times for both methods, as well as a throughput estimate, again compared with manual methods.

The assumptions behind Table 1 are: for the polished cross-section method, a single trained operator prepares the samples, parallelizing as many tasks as possible. Sixteen samples are prepared in parallel and analysed overnight in the SEM. Conversely, for the flat-bed method, four samples can be prepared in an hour, but the image analysis is much slower, mainly because a flat-bed image field contains

significantly fewer fibres than a cross-section one. Therefore, only four samples can be analysed in a given night or up to eight samples if the SEM is also used during the day.

From the software point of view, both methods analyse a field of view in less than a minute, which is the time it takes our SEM (Zeiss DSM 950) to move to a new field of view and to auto-focus reliably on it.

For the manual method, a single operator is assumed both for sample preparation and actual counting.

The cross-section method is probably the method of choice for high-throughput quality control needs. It does require a lot of qualified manpower to produce the samples, but its throughput is higher. The flat-bed method is probably best reserved for in-depth studies about a smaller number of products, for example in the case of research and development of production methods.

It is important to note that to make an image analysis approach successful, everything in the whole measurement process, from sample preparation to microscope control, must be automated and made as reliable as possible, otherwise potentially little may be gained in terms of both data quality and data quantity. The human factor is also still important: it is far better to have people operate an instrument than have them mindlessly feed mass-produced samples to an unknowable machine. Finally, it is worthwhile to stress the importance of sound stereological practices for such endeavours.

5. Conclusion

We have presented two innovative image analysis-based, automated methods for measuring large numbers of MMVF

Table 2. Tests for the three-way comparison between manual length-weighted counting, flat bed length-weighted automated analysis and polished cross-section automated analysis.

| | χ^2 value | 95% critical value |
|----------------------------|----------------|--------------------|
| Three-way test | 25.99 | 28.87 (18 d.f) |
| Manual vs. flat-bed | 9.78 | 16.92 (9 d.f) |
| Manual vs. cross-section | 16.08 | 16.92 (9 d.f) |
| Flat-bed vs. cross-section | 13.33 | 16.92 (9 d.f) |

dimensions from scanning electron microscope data. The first method measures MMVF diameters from cross-sections observed in BSE mode. The second method measures the MMVF diameter, length and number of extremities, which enables the unbiased estimation of diameter histograms and of the mean length of fibres by diameter class, which in turn provides estimates of the length, area or volume-weighted histograms.

The proposed approach gives a reliable and reproducible technique for estimating the diameter and length distributions of insulation MMVF. Both automated methods are as precise as, and on average between two and eight times faster than manual measurements, when all steps of the process are taken into account. From overnight runs, it is possible to use the SEM dedicated to this task in an optimal way. The number of measurement fields can be increased to obtain more accurate histograms. This technique can be used to collect data on the production plant in order to optimize or detect any drift in the process. The method presented in this paper can be used for other purposes, for example environment monitoring, and extended to other types of fibrous media, such as reinforcement glass fibres, paper or asbestos.

The production implementation of both methods has been carried out in an MMVF pilot plant. The next step will be, after long periods of validation, to contribute to the definition of new standards, based on the image analysis of fibres.

References

- Arcelli, C. & Sanniti di Baja, G. (1985) A width-independent fast thinning algorithm. *IEEE Trans. Pattern Anal. Machine Intell.* **7**(4), 463–474.
- Beucher, S. & Vincent, L. (1990) Introduction aux outils morphologiques de segmentation. *Traitement d'Image en Microscopie à Balayage et en Microanalyse Par Sonde Électronique*. ANRT, Paris, 1–43.
- Blum, H. (1961) An associative machine for dealing with the visual field and some of its biological implications. *Biological Prototypes and Synthetic Systems, 2nd Annual Bionics Symposium*, Cornell University (ed. by E. E. Bernard and M. R. Kare), pp. 244–260. Plenum Press, New York.
- Danilatos, G.D. (1988) Foundation of environmental scanning electron microscopy. *Adv. Electronics Electron Phys.* **77**, 109–250.
- Degenne, M. (1989) *Préparation des coupes polies de fibres de verre d'isolation en vue d'une observation au MEB*. Internal Technical report, Isover Saint Gobain, CRIR, Rantigny, France.
- Fienberg, S.E. (1983) *The Analysis of Cross-Classified Categorical Data*, 2nd edn. MIT Press, Cambridge, MA.
- Koenig, A.R., Hamilton, R.D., Laskowski, T.E., Oslon, J.R., Gordon, J.E., Christensen, V.R. & Byers, C.D. (1993) Fiber diameter measurements of bulk MMVF. *Anal. Chim. Acta*, **280**, 289–298.
- Kurdy, M. & Jeulin, D. (1989) Directional mathematical morphology operations. *Proceedings of the 5th European Congress on Stereology, Freiburg im Breisgau, Germany. Acta Stereol.* **8**(2), 473–480.
- Lantuéjoul, Ch. (1980) On the estimation of mean values in individual analysis of particles. *Microsc. Acta*, **5**, 266–273.
- Lee, T. & Talbot, H. (1995) A fast method for detecting and matching linear features in images. *Proc. DICTA, Brisbane, Australia*, 649–654.
- Meyer, F. & Beucher, S. (1990) Morphological segmentation. *Visual Comm. Image Repr.* **1**(1), 21–46.
- Miles, R. (1974). *Stochastic Geometry, On the Elimination of Edge Effects in Planar Sampling*. J. Wiley and Sons, New York.
- Rosenfeld, A. & Kak, A. (1982) *Digital Picture Processing, Vol. 2*. 2nd edn. Academic Press, London.
- Rosenfeld, A. & Pfaltz, J. (1968) Distance functions on digital pictures. *Pattern Recogn.* **1**, 33–61.
- Serra, J. (1982) *Image Analysis and Mathematical Morphology*. Academic Press, London.
- Talbot, H. (1993) Analyse morphologique de fibres minérales d'isolation. PhD Thesis, Ecole des Mines de Paris, France.
- Talbot, H. (1996) A morphological algorithm for linear segment detection. *Proceedings ISMM'96 Mathematical Morphology and its Applications to Image and Signal Processing* (ed. by P. Maragos, R. W. Schafer and M. A. Butt), pp. 219–226. Kluwer, Atlanta, GA.
- Talbot, H. & Jeulin, D. (1993) Estimation of fibre length and diameter distribution from SEM images. *Acta Stereol.* **13**(2), 357–362.
- Talbot, H., Jeulin, D. & Hanton, D. (1996) Image analysis of insulation mineral fibres. *Microsc. Microanal. Microstr.* **7**, 361–368.
- Talbot, H., Jeulin, D. & Hobbs, L.W. (1992) Scanning electron microscope image analysis of fiber glass insulation. *Proceedings of the 50th Electron Microscopy Society of America, Boston, MA*, 994–995.
- Talbot, H. & Terol Villalobos, I. (1992) Binary image segmentation using weighted skeletons. *Image Analysis and Morphological Image Processing III, San Diego, CA, SPIE*, **1769**, 147–155.
- Talbot, H. & Vincent, L. (1992) Euclidean skeleton and conditional bisectors. *Visual Communication and Image Processing, Boston, MA, SPIE*, **1818**, 862–873.
- Viero, T. & Jeulin, D. (1995) Morphological extraction of line networks from noisy low-contrast images. *Visual Comm. Image Repr.* **6**(4), 333–347.

Appendix. Numerical data**Table A.1.** Numerical values for all the histograms

| Diameter | Fig. 11 | | Fig. 12 | | Fig. 13 | | Fig. 15 | | |
|----------|---------|-----------|---------|-----------|---------|-----------|---------|----------|---------------|
| | Manual | Automated | Manual | Automated | Manual | Automated | Manual | Flat-bed | Cross-section |
| 0 | 275 | 255 | 120 | 131 | 52.39 | 39.22 | 54 | 99 | 132 |
| 1 | 215 | 217 | 147 | 144 | 76.39 | 71.35 | 43 | 54 | 50 |
| 2 | 199 | 179 | 101 | 102 | 81.2 | 72.56 | 49 | 59 | 96 |
| 3 | 175 | 144 | 78 | 73 | 71.85 | 63.36 | 54 | 85 | 93 |
| 4 | 155 | 133 | 70 | 62 | 71.81 | 68.62 | 49 | 72 | 113 |
| 5 | 116 | 106 | 56 | 44 | 71.35 | 65.1 | 60 | 63 | 74 |
| 6 | 91 | 94 | 31 | 37 | 66.93 | 64.41 | 39 | 62 | 79 |
| 7 | 92 | 89 | 13 | 14 | 72.94 | 57.65 | 30 | 45 | 55 |
| 8 | 75 | 73 | 5 | 6 | 50.34 | 66.05 | 17 | 14 | 36 |
| 9 | 48 | 38 | 3 | 2 | 63.68 | 69.84 | 5 | 5 | 11 |
| 10 | 39 | 44 | 2 | 0 | 0 | 0 | 4 | 11 | 3 |
| 11 | 31 | 31 | 0 | 2 | 92.41 | 91.52 | 1 | 5 | 3 |
| 12 | 28 | 27 | 0 | 0 | 0 | 0 | 1 | 2 | 0 |
| 13 | 15 | 16 | 0 | 0 | 0 | 0 | 0 | 0 | 3 |
| 14 | 13 | 13 | 0 | 0 | 0 | 0 | 2 | 0 | 3 |
| 15 | 13 | 14 | 0 | 0 | 0 | 0 | 0 | 0 | 3 |
| 16 | 12 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 17 | 6 | 13 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 18 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 19 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Total | 1598 | 1491 | 626 | 617 | 771.29 | 729.68 | 407 | 577 | 752 |