NBSIR 78-1442

Algorithms for Image Analysis of Wood Pulp Fibers

Russell A. Kirsch

Institute for Basic Standards
National Bureau of Standards
Washington, D.C. 20234

January 1978

Final

Paper presented at

1978 Annual Convention of Technical Association of the Pulp and Paper Industries, March, 1978

NBSIR 78-1442

ALGORITHMS FOR IMAGE ANALYSIS OF WOOD PULP FIBERS

Russel A. Kirsch

Institute for Basic Standards
National Bureau of Standards
Washington, D.C. 20234

January 1978

Final

Paper presented at
1978 Annual Convention of
Technical Association of the
Pulp and Paper Industries,
March, 1978

U.S. DEPARTMENT OF COMMERCE, Juanita M. Kreps, *Secretary*

Dr. Sidney Harman, *Under Secretary*

Jordan J. Baruch, *Assistant Secretary for Science and Technology*

NATIONAL BUREAU OF STANDARDS, Ernest Ambler, *Acting Director*

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

LIST OF FIGURES AND TABLES (cont.)

ALGORITHMS FOR IMAGE ANALYSIS OF WOOD PULP FIBERS*

Russell A. Kirsch
National Bureau of Standards
Washington, D.C. 20234

ABSTRACT

Image analysis technology can be used to measure the visible morphology of pulp fibers. But before such measurements can be accepted, it is necessary to achieve precise definition of the necessary measurements in the form of suitable algorithms that have been experimentally tested on images of actual fiber data. We present such measurement results on both semiautomatically traced fiber data and on automatically scanned images. We explore the variety of definitions possible for some simple well known properties of wood fiber morphology by applying suitable algorithms to fiber image data. Finally, we suggest that this exploratory approach to the specification of the precise image analysis measurements needed in paper manufacturing can facilitate the introduction of a technology for process control that will result in savings in paper manufacturing cost and in reduction of energy requirements.

Key words: Algorithms; artificial intelligence; image analysis; morphological analysis; paper fibers; pattern recognition; pulp characterization.

*This work was done, in part, under interagency agreement EA-77-A-01-6010 A019-IEC between the Department of Energy and the National Bureau of Standards.

INTRODUCTION

The properties of manufactured paper are largely determined by the physical and chemical properties of the wood pulp fibers used in its manufacture. To a lesser extent, so are the energy costs of manufacturing the paper. To control the quality of the pulp raw material is to control a large part of the manufacturing cost and the resulting physical properties of paper.

This control can be achieved in two very different ways: (a) by holding substantially constant those pulp properties which affect costs and paper properties, or (b) by making measurements on less well controlled pulps, thereby enabling compensation to be made during manufacture for the inevitable variation in pulp quality. Although many of the measurements conventionally made on pulps are well developed and well understood, they are necessarily of use only retrospectively in the laboratory and consequently of limited use for process control because of the time needed to perform them.

Many important measurements made in the laboratory are directed to the visible morphology of pulp fibers. Such measurements are prime candidates for the application of automatic image analysis technology. This technology, now over twenty years old,[1] has matured to the point where there are many commercial instruments available,[2] some at quite modest costs compared to the costs of performing such measurements only a few years ago. Most uses of this technology have been in fields different from fiber morphology measurement. But where image analysis has been used, there are enough similarities to warrant the investigation of the possible extension of the technology to the analysis of pulp fiber morphology. Such fields include earth satellite image analysis, biological microscopy, quantitative metallography, and a spectrum of lesser developed fields ranging from commercial to research uses.

In each of the fields where image analysis technology has been successfully used, there has been an initial exploratory phase during which the development of the particu lar types of measurements needed used general purpose instruments. More specialized and economical instrumentation has followed this exploratory phase. Today, the technology for development use is readily available. If an understanding of image analysis technology's potential for developing suitable fiber morphology measurements becomes widespread, one can anticipate a rapid acceptance of image analysis technology by the pulp and paper industry with the consequent benefits of reduced energy costs and improved quality control. It is to this challenge for new applications of image analysis to the measurement of fiber morphology that we direct our attention in this paper. In a broader context than that of this paper, we hope to show how measurement methods using image analysis and pattern recognition can be used to predict paper properties. The measurement methods will be tested, at the National Bureau of Standards Paper Laboratory. Those measurements that survive the tests by usefully predicting paper properties will be made available to the electronics industry, in the form of algorithms, to be transferred to more specialized electronic instrumentation. Although this transfer will involve substantial improvements (to be made by others), it is expected that the algorithmic basis for the measurements that will have been developed and verified experimentally, will remain substantially unchanged. Thus, our contribution is to supply tested measurement algorithms for fiber morphology. The present paper is devoted to the algorithm development. Subsequent laboratory validation will be reported elsewhere.

In determining the morphology of pulp fibers, it is useful separately to consider the two stages of morphological measurement: data acquisition, and algorithmic derivation of the desired properties from the data. The first stage consists of using suitable transducers to make measurements that are directly connected with simple geometric properties of the image being measured. Thus such properties as the coordinate locations (in some appropriate coordinate system) of points in an image are determinations that belong to the data acquisition stage. The second stage consists of performing algorithm based operations upon the results of the first stage to determine properties that are usually more macroscopic and complex. Thus length measurements belong to this algorithmic derivation stage.

The reason for distinguishing these two stages is that the data acquisition stage is usually concerned more with physical properties of the apparatus for making measurements whereas the algorithmic derivation stage is usually concerned more with the algorithms (i.e., programs, procedures) for calculating derived properties. Usually, the precision of measurements is significantly determined by how the first stage is conducted whereas the gross structure is determined by the second stage. We would thus expect, that in areas where the measurements are well defined and widely accepted as universal procedures, emphasis is placed (and properly so) upon the data acquisition stage. But where there is little agreement upon the nature of the kinds of measurements to be made, emphasis must be placed on the formulation of suitable algorithms which can serve as a basis for agreement upon which later measurements can be based.

We will see examples, below, of both stages of morphological measurement. For the first stage, we will present measurements made both with graphic stylus semiautomated techniques and with automatic scanning directly from image data. For the second stage, we will show how different algorithms, operating upon the same set of data, can be used to characterize a particular morphological property, that of fiber curl.

An additional reason for considering the second stage of algorithmic determination in a comparatively new area like pulp fiber morphology measurement is that in the development of new measurement methods this second stage probably should be considered first. It is all too easy to adopt a technology for data acquisition and then to allow those measurements which follow directly from the adopted technology to assume positions of importance in the measurement scheme. This has always been the classical approach in scientific measurement. But the existence of general purpose programmable computers as mediators in scientific measurement has opened up wide new possibilities for producing measurements from instruments where the data furnished by the instruments are by no means direct derivations from the instrument transducers. Rather, they are the results of elaborate and sometimes ill-understood computations. To have a firm algorithmic base for these derived measurements is to insure that the results of the underlying elaborate computations will not be misunderstood. It is interesting to note that in the area of image analysis technology we already have routine occurrences of investigators 'looking' at images that have never been seen before and that could not even, in principle, be 'seen' in any simple sense of the word. A brain scan in a tomograph, a false color satellite image of the earth, and even a scanning electron micrograph all present a misleadingly simple picture that requires intimate knowledge of the algorithms underlying the instrument in the first two cases, and of the transducer properties on the third, before one can hazard an interpretation of the simple image presented by the instrument. Since the application of image analysis to fiber morphology measurement is so new, we will devote a correspondingly larger part of our discussion, below, to these algorithmic derived measurements.

FIBER MORPHOLOGY IMAGE DATA ACQUISITION

There are several aspects of the morphology of pulp fibers that can be determined with suitable scanning methods. These include both optical and electron microscopy based scanning methods. We will be concerned here only with optical scanning techniques although some of the methods described below will be equally applicable to measurements that can be performed on electron microscope images. For purposes of exploring the algorithmic base for fiber measurements, a semi-automatic data acquisition method is first presented here. Subsequently we will discuss methods for direct scanning of microscope images of fibers.

The semi-automatic method uses computer graphics technology. A photomicrograph of a suitable pulp preparation is placed on a graphic tablet. This tablet is connected to a general purpose programmable computer which can accept indications, from a crosshair cursor, of the coordinate locations being pointed to by a human operator. The operator places this cursor, successively, on locations along the length of a particular fiber that he has chosen to measure in the photomicrograph. Thus the operator describes a particular fiber to the computer which represents that fiber as a sequence of (x,y) coordinate pairs. Each of the fibers in the photomicrograph is manually described in this way to the computer. For a photomicrograph, such as that shown in Fig. 1, the process of tracing all the fibers takes slightly less than 10 minutes, the time being related to the number of measurement points selected by the operator. In the case of Figure 1, this yields the

Figure 1: Photomicrograph of Southern Softwood Pulp Fibers

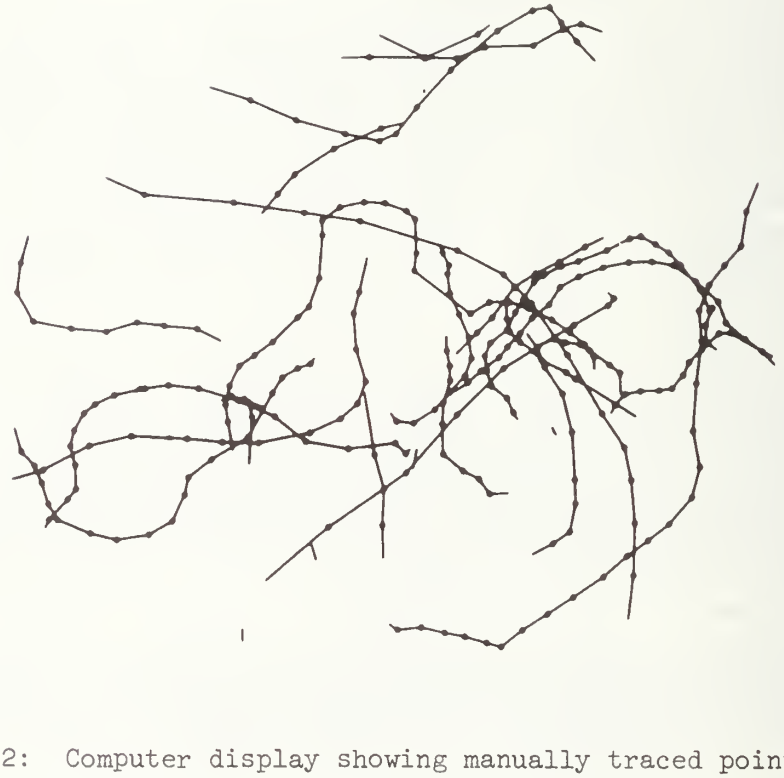

Figure 2: Computer display showing manually traced points of Figure 1

points shown marked on the computer display of the measured photomicrograph in Fig. 2. The tablet used for the present experiments is capable of resolving crosshair locations to within 0.25 mm in both x and y directions. Of course, the resolution in terms of fiber dimensions is a function of the photographic magnification used.

We see, in Fig. 3, a particular fiber selected from the data of Fig. 2 in which the coordinates of points chosen by the operator in tracing that fiber are given. The original coordinate system used is an arbitrary one imposed by the tablet. Since the actual fiber length from end to end as measured with a calibrated microscope reticle is 2.02 mm, the coordinates have been converted to units of micrometers, which are the ones appearing in Fig. 3.

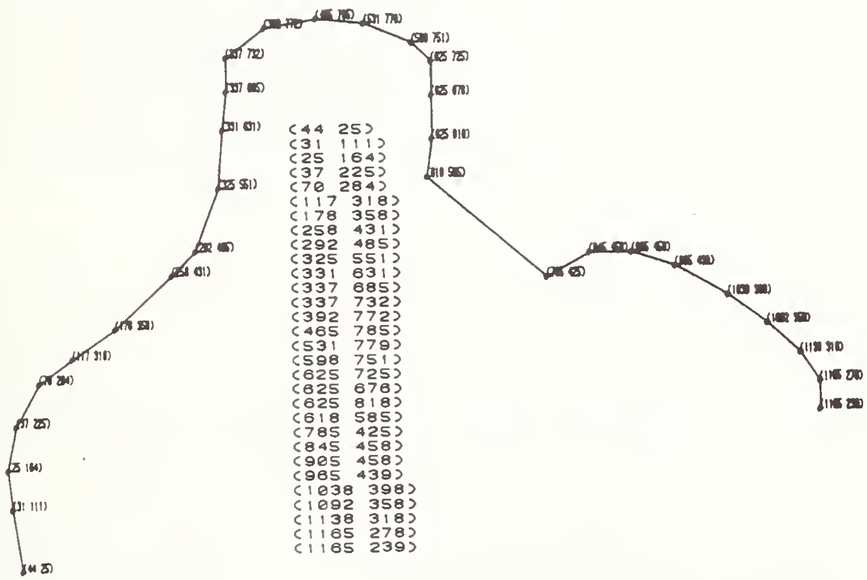

Figure 3: Coordinates (micrometers) of points traced on a single fiber

This semi-automatic method of data acquisition allows us to define a number of useful algorithm-derived measurements described below. However, the measurements made are significantly dependent upon the vagaries of manual placement of a tracing cursor and the underlying decisions made by an operator while doing the tracing. It is necessary to develop a fully automatic method for data acquisition if we are to be able to insure reproduceability of any measurements made. For this approach, automatic image scanning is necessary.

There are a number of different types of image scanners that can be used to acquire image data on pulp fibers.[2] The most direct method involves the use of a scanner looking directly through a microscope. The scanner can be either a mechanical one which directs the whole image, via galvanometer mirrors, to an aperture through which a photomultiplier views a small segment of the field, (a pixel), or a light source can be directed through the microscope optics to illuminate an individual pixel in the specimen which is then sensed by a suitably placed photodetector.

A less direct method of image scanning, which we have used for the experiments reported here, is to scan a photomicrograph of the specimen. The photomicrograph can be scanned through a microscope just as the original specimen, or it can be placed on a rotating drum type of scanner which scans successive pixels in a helix as the drum rotates. We have used such a drum scanner to scan photomicrographic transparencies of pulp preparations. The scanner resolves a photograph into pixels 25×10^{-6} meters square. Again, the resolution in terms of the fiber depends on the magnification of the microscope and the photographic process. For each pixel, the optical density is measured to 8 bits (256 parts) over the range from 0.0 Optical Density to 2.0 Optical Density. Here too, these densitometric measurements must be interpreted to include the photographic process and the characteristics of the microscope.

Once such a fiber image has been scanned, it is useful to be able to view the scanned data before further processing. We have written a display program that gives a view of the scanned data useful for discovering any gross errors that may have occurred in the scanning process. This display represents the original 256 levels of scanned density with 10 distinguishable symbols of varying darkness on a cathode ray tube display. Various mappings can be used to map the actual range of densities (typically smaller than the full potential range of 256 levels) into the 10 display symbols. Such a displayed scan is shown in Fig. 4. Actually, the full scan resolution is shown in the segment of a scanned image in Fig. 5, the image of Fig. 4 being obtained by displaying a suitable spatial average of the pixels of the original scan.[3]

The possibilities of scanning fiber images are not limited to the two types of data acquisition methods shown here. With direct microscope scanning, it is possible to change the focal plane of the microscope under computer control. With suitable narrow depth of field optics, this makes possible the successive scanning of three dimensional image arrays. Certain limited measurements that go beyond the measurements possible on two dimensional projections can thus be made. It is necessary, however, to do fairly complex computer processing of the multiple scans to separate the in- and out-of-focus portions of such images if three dimensional reconstruction is to be done.

Another possibility for scanning fiber images with light microscopy is to use monochromatic light as the illumination source. The spectral

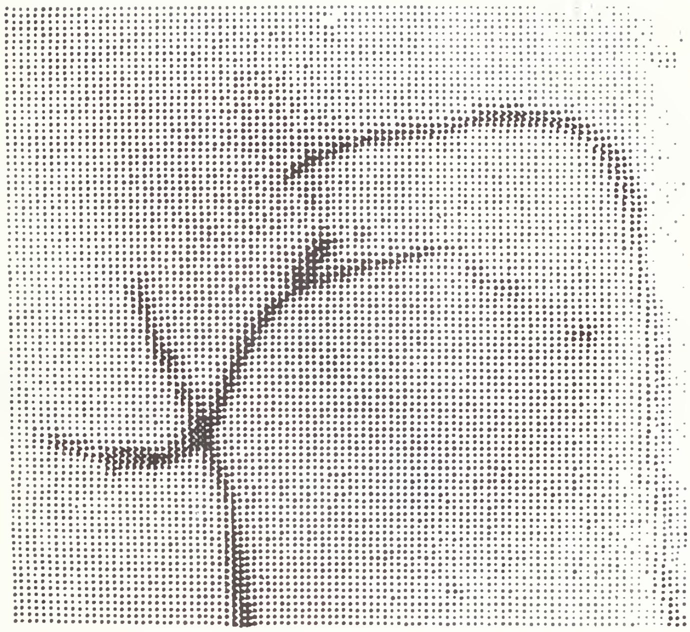

Figure 4: Computer display of scanned fiber photomicrograph showing, with 10 distinct characters, the range of 256 scan density levels

transmission and reflectance characteristic signatures of fibers thus obtained can be used to obtain fiber classification. The widespread use of this technology for viewing and analyzing satellite images of the Earth suggests this false color viewing as a fruitful area for investigation.

Also, among the various scanning modalities, mention should be made of the possibility (and indeed in some cases of the necessity) for computer control of spatial selective scanning of fiber images. In dilute preparations of fibers in suspension, a scanned image will typically contain only a small fraction of the image area subtended by fibers. To scan such an image at full resolution is to be wasteful of time, storage capacity, and processing capability. When the scanner is under direct computer control, it is possible to scan at low resolution in order to find regions requiring higher resolution scanning, and then to rescan these regions with the desired resolution. Thus one could scan an image like that shown in Fig. 4 to locate the regions requiring

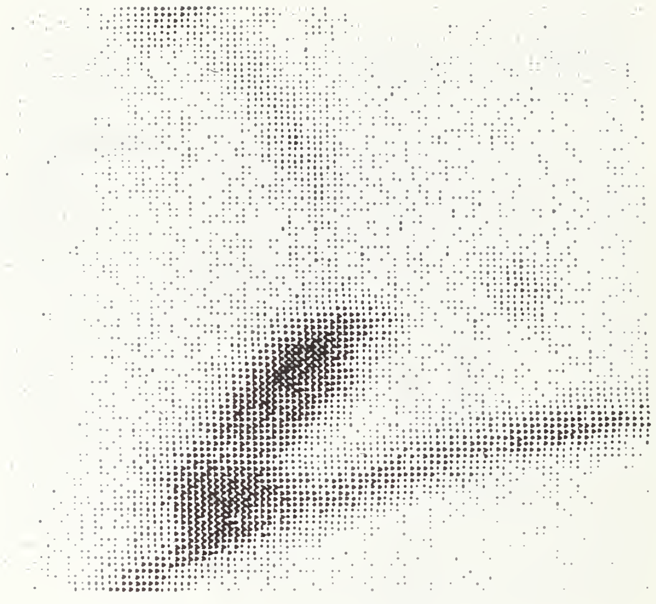

Figure 5: Computer display of full spatial resolution of scanned photomicrograph

fine structural detail to be resolved. Then one could rescan to obtain the resolution shown in Fig. 5. The first scan is, in effect, a map of the regions subsequently to be scanned at higher resolution. Sometimes, this rescanning is necessary not because of the need for higher resolution, but because analysis shows that further analysis of a previously scanned region is necessary. Unless storage capacity for the very large quantity of data scanned is available, this usually requires rescanning of the region requiring further analysis.

Finally, there is a scanning modality closely related to a feature available in existing commercial image analysis instruments. This is the dynamic range selection capability. Once an image has been coarsely scanned, it is possible to determine the actual density ranges occurring in the image. Suitable readjustment of the scanning density range to maximize contrast can then be done. This capability must be used cautiously, however, because it is possible to introduce scanning artifacts into an image at the boundaries where different density ranges have been selected.

ALGORITHMS FOR ANALYSIS OF FIBER IMAGES

Once image data for a pulp preparation has been acquired, there are many kinds of measurements that may be made on the data. These measurements can be made with precisely defined, algorithmic procedures that have the virtues of unambiguity and reproduceability. In the process of constructing the algorithms for making these measurements, it becomes evident that the level of precision and completeness required by a computer programming language demands that many decisions be made regarding different interpretations of what would otherwise be considered well defined measurement methods. To illustrate the varied ways in which algorithmic definitions can be made, we will consider a set of definitions of a property well known in fiber morphology studies to be an important determinant of finished paper characteristics, the property known as curl. These definitions will be illustrated with the semi-automatically obtained fiber data presented above.

We will not attempt to argue for any one definition of curl. Rather, we will suggest that there are many more definitions, and combinations of these, among which an intelligent choice may be made regarding the one that incorporates notions of what morphological properties are important in prediction of paper properties. Our objective is to illustrate the ease with which one can use a general purpose computer to test such definitions and to produce candidates which may then be further evaluated in the laboratory.

The data to be used in presenting each algorithm is shown in Fig. 3. For each traced point, the coordinates of that point scaled to the original fiber image, are shown in micrometers. This has been done by using a calibrated reticle in the microscope and photographic process that was used to produce the images that were traced. Thus, relative to some arbitrary origin, the fiber traced in Fig. 3 begins at a point whose coordinates, in micrometers, are (44,25), and then continues to (31,111), (25, 164), etc. The fiber can be closely approximated, for fiber length calculation purposes, by assuming that these points are joined by straight line segments. Had the tracing points been much farther separated, this would not be a good assumption. Using this assumption, we can compute the distances between tracing points as shown in Fig. 6. Each of the distances shown is rounded to the nearest micrometer. The sum of these distances is 2022 micrometers. The actual length calculated from the coordinate data is 2019 micrometers, the disparity coming from rounding errors.

We are now prepared to calculate a first proposed curl measure. This measure corresponds to one discussed by Kibblewhite.[4] For a fiber, we calculate the length and divide this by the distance between end points. If the fiber is relatively straight, this ratio will be near 1.0. If it curves very much, the ratio will be larger. In Fig. 7 we show several such ratios calculated for different segments of the same fiber. For the initial segment, which is fairly straight, this curl

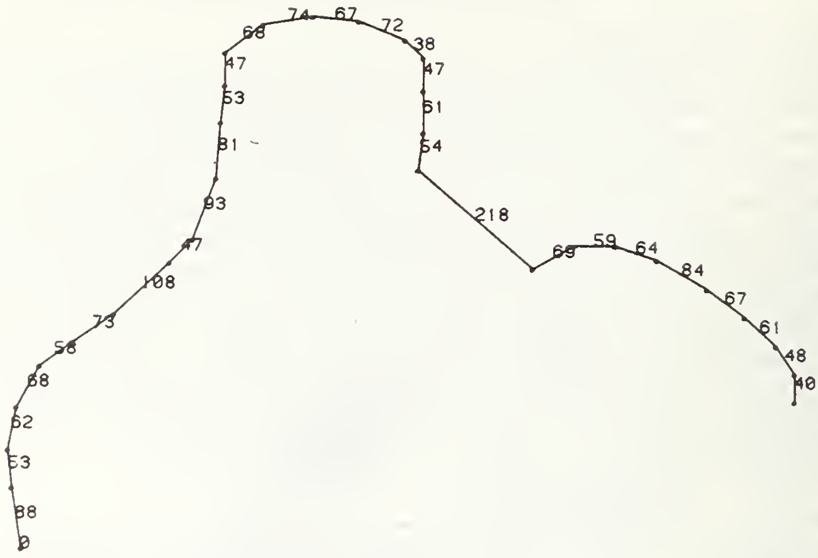

Figure 6: Distances between tracing points in single fiber

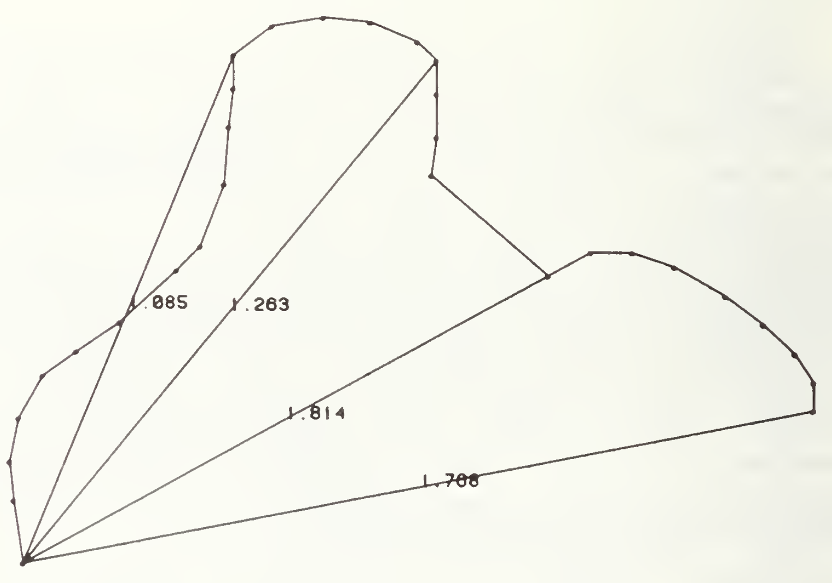

Figure 7: Ratios of length along fiber segments to distances between end points. Distances are calculated from lower left fiber end

measure is 1.085 whereas for the longer segments possessing more curvature, the measure is 1.263, 1.814, and 1.768 respectively.

We have displayed this curl measure for different segments in order to suggest an idea to be developed further, namely, that certain measures might be assigned not to whole fibers but rather to segments of fibers. These segments might possibly act like whole fibers in contributing structurally to the properties of the paper in which they are used. It can be seen that with little additional difficulty the algorithm for defining this curl measure can be extended to be a continuously varying function of position along a fiber rather than a simple function of the whole fiber.

We obtain another proposed measure for curl by calculating the angle of bend at each tracing point along a fiber. Since we are approximating the fiber by straight line segments, the angle between these segments is a well defined quantity. In Fig. 8 we see the same fiber plotted with small circles marking the plotting points and with the angle of bend at those points shown, in degrees. Thus, for the example shown, there are pairs of segments between which the angle varies in magnitude from 0 degrees to 69 degrees. The algebraic sign denotes the direction of bend along the fiber, a bend to the right

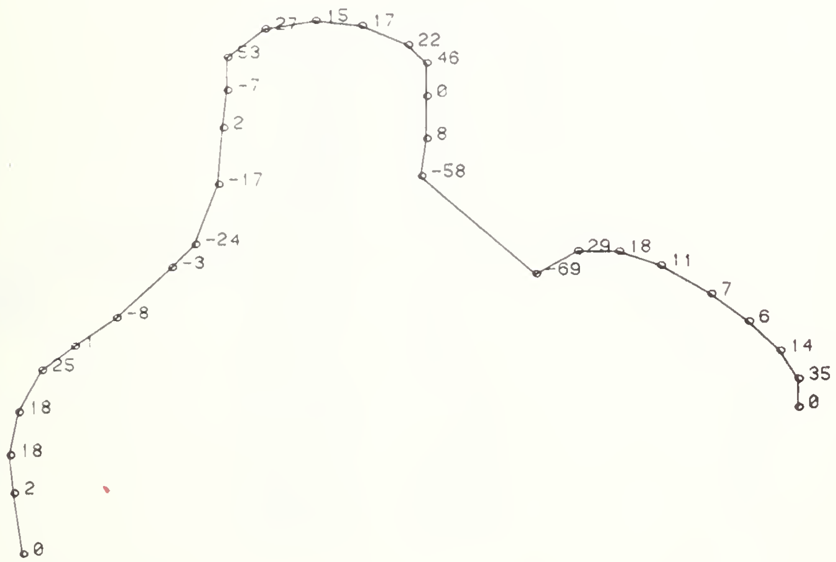

Figure 8: Angle of bend (degrees) at each tracing point along fiber

(going clockwise in the figure) being denoted positively, and one to the left negatively. At the end points where the angle is not defined, a zero is arbitrarily assigned.

One can imagine various ways of combining these angles. They can be averaged, their magnitudes can be averaged, the maximum value can be used to characterize the whole fiber, or they can be used in more complex ways, including the possibility of combining the angle measure with other measures. We have performed such experiments and will report one such combination involving angle and segment length below.

Yet another measure for curl can be obtained by generalizing the angle measure slightly. If we take sequences of three tracing points, they will, in general, uniquely define a circle passing through these points. The exception occurs in the case when the points are collinear, in which case the circle is degenerate, of infinite radius. In Fig. 9 we have plotted, at each tracing point, the radius in micrometers, of the circle passing through that point and the two points on either side of it. In the single case where three points are collinear, an 'S' for 'straight' is shown, as it is for the exceptional points at the ends of the fiber.

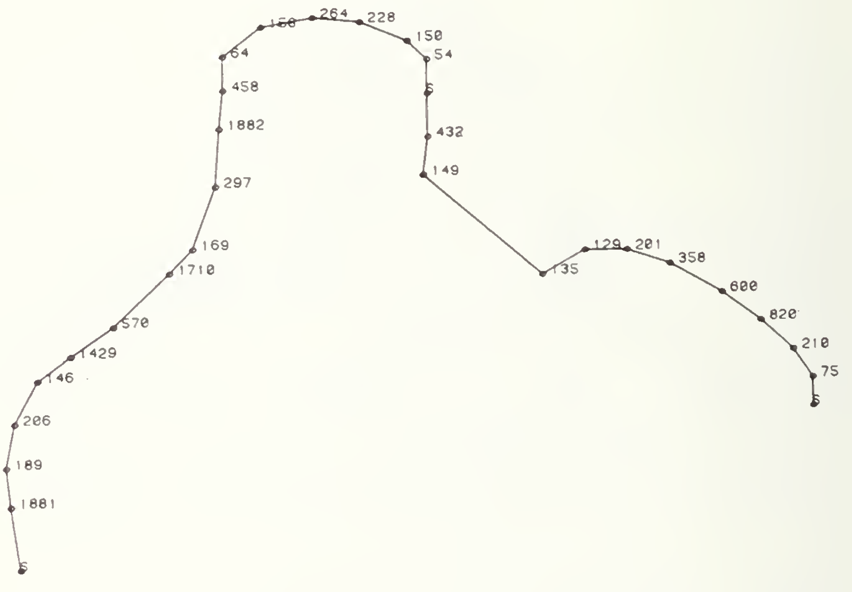

Figure 9: Radius (micrometers) of circle passing through each tracing point and adjacent point on each side

This radius of curvature measure is slightly more general than the simple angle measure because it allows information to be incorporated regarding the distance between tracing points. In the special case where tracing points are equally spaced, points of equal angle will have equal radius of curvature. For automatically scanned images this will generally be the case. A complication enters, however, when equally spaced tracing (or scanning) points straddle a sharp bend. It is desirable, in that case, to use an additional scanning point at the sharp bend, which introduces non-uniform spacing. In the manually traced data shown here, these critical points have been deliberately introduced. To see how the angle and radius of curvature measures are related, we have shown, in Table 1, the various angles of Fig. 8 in increasing magnitude order, and for each one have given the corresponding radius of curvature at that point. Note that the radius of curvature does not necessarily decrease with increasingly sharp angle of bend.

1.436	1428.6
2.234	1882.4
2.298	1880.9
2.771	1710.3
5.702	819.7
6.632	458.1
6.808	600.2
7.943	432.3
8.271	569.6
11.135	358.3
13.762	209.6
14.909	263.9
16.695	296.6
16.877	228.4
17.823	206.4
18.084	201.2
18.166	188.9
21.500	149.6
23.906	168.9
25.248	145.7
27.251	158.4
28.855	129.1
34.508	74.6
46.333	54.0
53.130	63.6
58.012	148.5
68.787	134.6

Table 1: Angle of bend (degrees) (left col.) as a function of radius of curvature (micrometers) (right col.) of the tracing points in Figures 8 & 9

We have suggested that the ease of defining algorithms on a general purpose research computer enables one to experiment with new, complex measurement procedures. One such algorithm enables us to incorporate information on fiber length and curl in a single measure. To see how this may be done, consider the diagram of Fig. 10. Here we see the same fiber, but with lengths indicated between various plotting points. The points chosen are just those at which the angle of bend exceeds, in magnitude, some prescribed amount, 10 degrees in this case. It is as though the fiber was considered to have been 'broken' at those points. If we change the 'breaking' criteria to, say, 20 degrees, we get the set of segment lengths shown in Fig. 11. Note that the segments become longer in some cases because some of the bends of Fig. 10 may exceed 10 degrees but not the 20 degrees necessary for segmentation in Fig. 11. If we segment at 30 degrees, we get the six comparatively long segments shown in Fig. 12. In the extreme case of segmentation at 180 degrees, there is no 'breaking' at all and we get the ordinary fiber length.

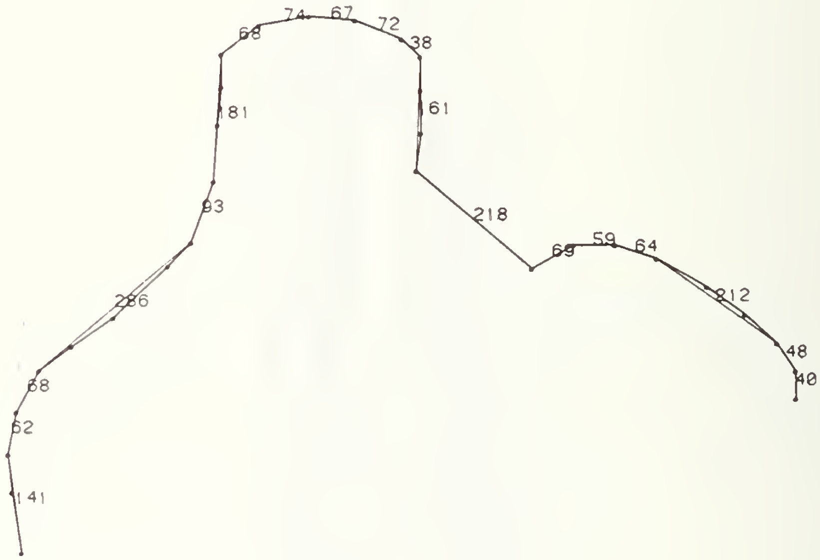

Figure 10: Segment lengths between bends sharper than 10 degrees

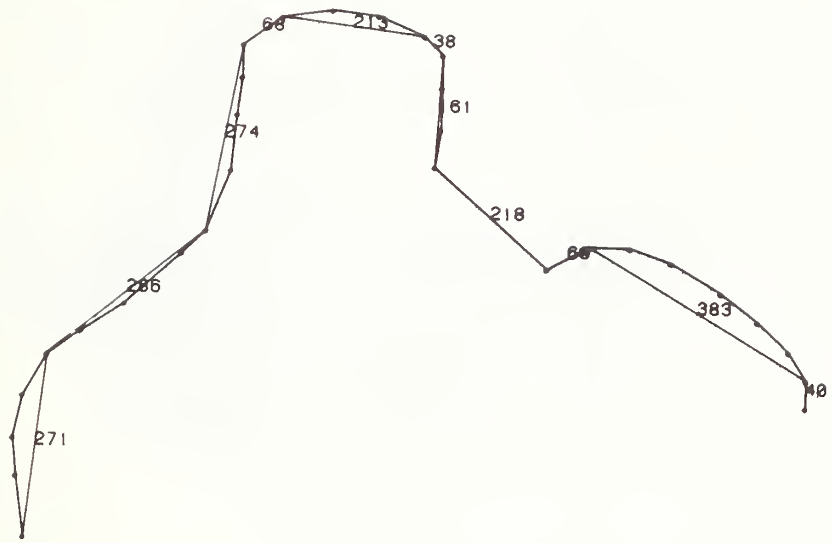

Figure 11: Segment lengths between bends sharper than 20 degrees

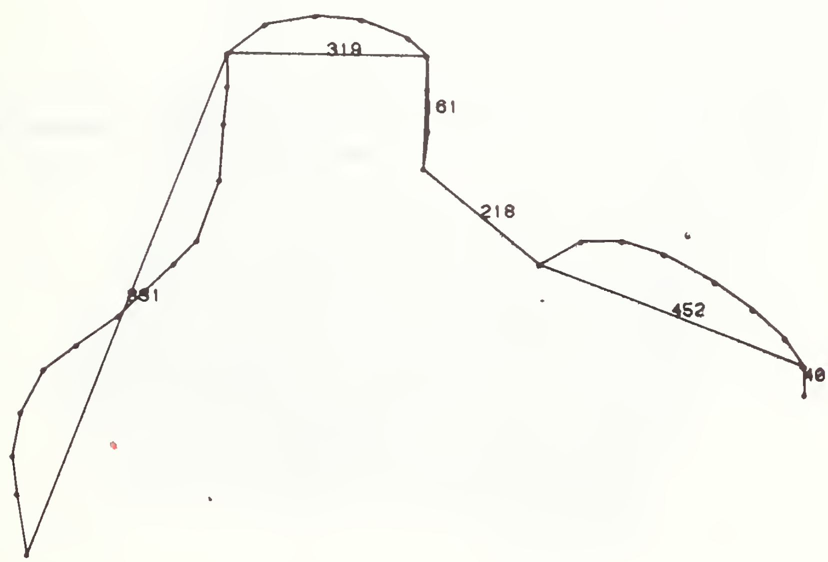

Figure 12: Segment lengths between bends sharper than 30 degrees

We have performed an extensive set of measurements, first reported in Ref. 3, using this last notion of curl. A set of three pulps were analyzed from semiautomatically traced photomicrographs. The images analyzed contained 2585 distinct fibers. They were drawn from three pulps, a northern and a southern softwood, and a hardwood. Fig. 1 is a typical image from the set of southern softwood images. It contains 28 of the 2585 fibers analyzed. For these 28 fibers, there were 317 tracing points.

In Fig. 13 we show the result of calculating a classical fiber length distribution for these three pulps. We find, as expected, that the hardwood fibers are primarily short, and the softwood fibers longer. The distribution of fiber lengths is also broader for the softwoods than for the hardwoods.

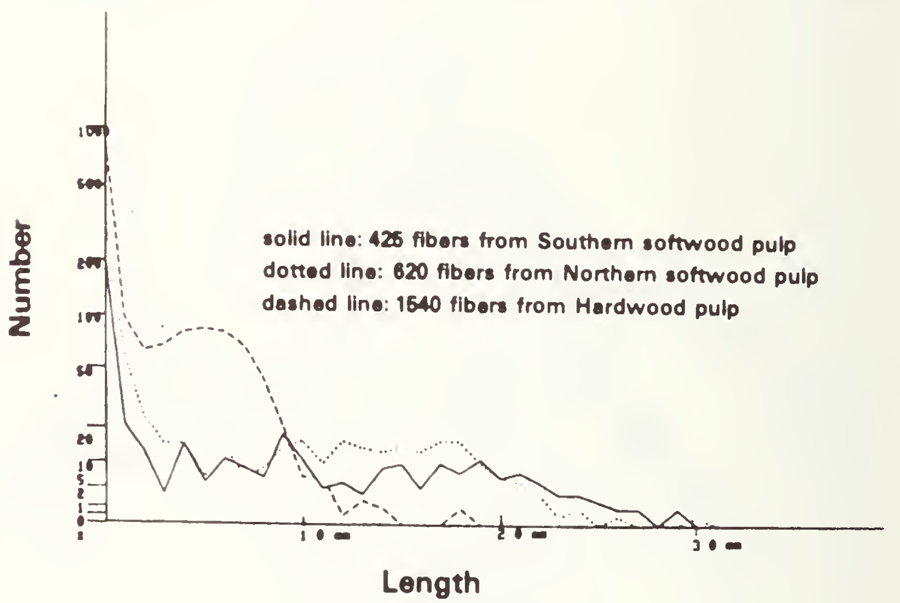

Figure 13: Histogram of number of fibers in length classes of 0.1 mm. for three different pulp sources

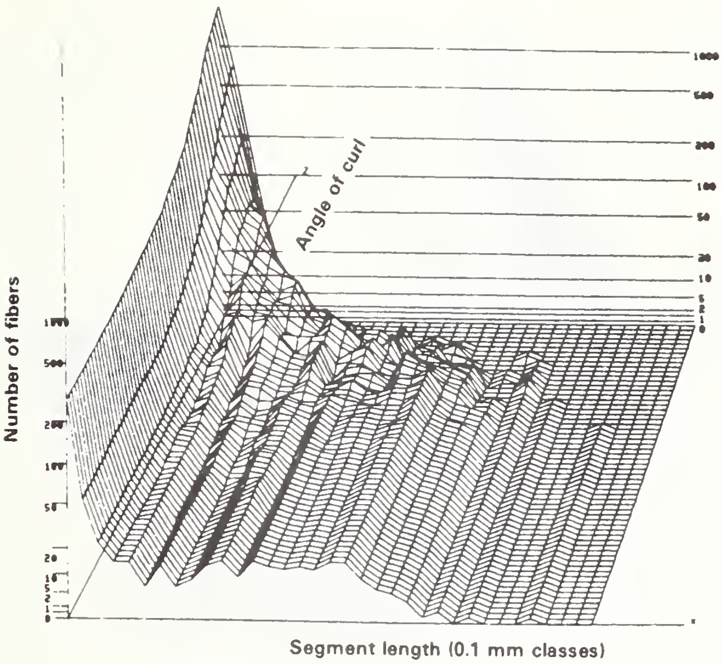

Figure 14: Histogram of number of fiber segments as a function of segment length for curl angles from 180 degrees (front) to 0 degrees (rear). Data from 620 Northern Softwood fibers

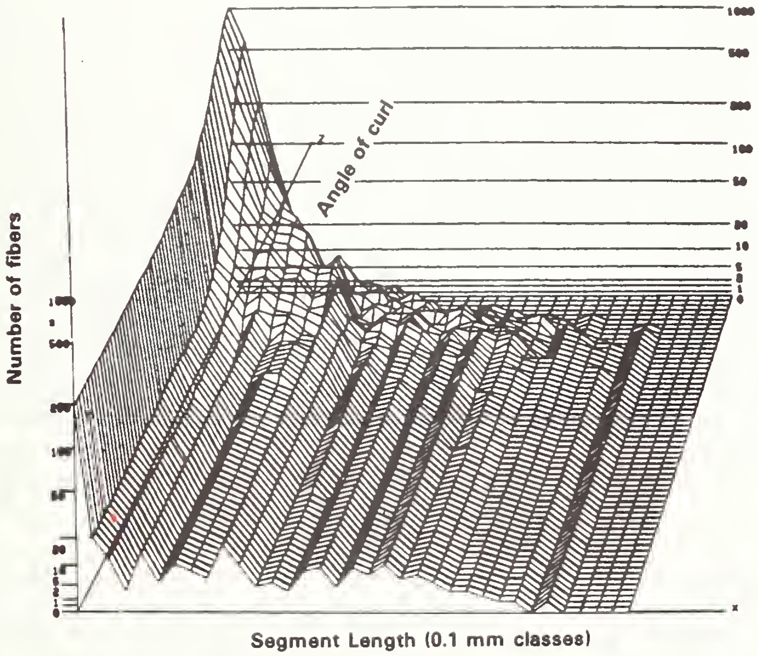

Figure 15: Histogram of number of fiber segments as a function of segment length for curl angles from 180 degrees (front) to 0 degrees (rear). Data from 425 Southern Softwood fibers

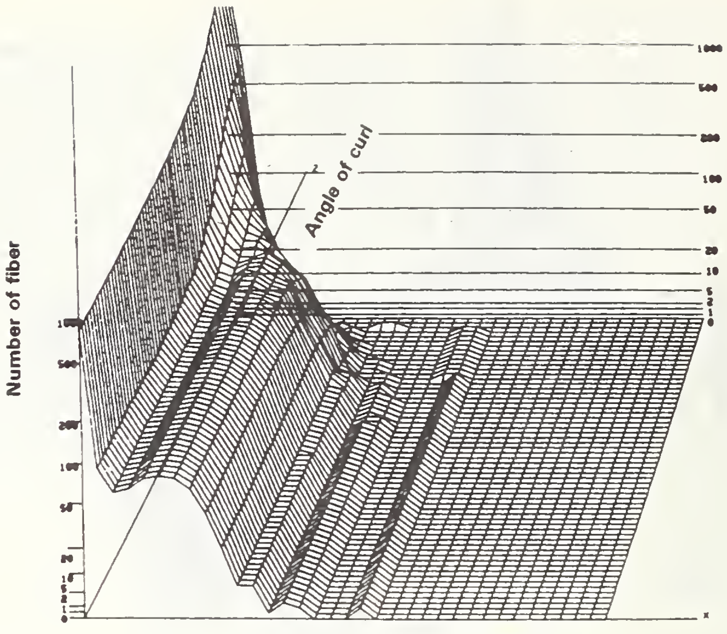

Figure 16: Histogram of number of fiber segments as a function of segment length for curl angles from 180 degrees (front) to 0 degrees (rear). Data from 1540 Hardwood fibers

A more interesting and informative way to view these data is shown in Figures 14, 15, and 16. In Fig. 14, data for the northern softwood is given. Here we see a three dimensional surface which depicts information both on the fiber length distribution and on the curl distribution. The front of the surface can be seen to be the curve of the northern softwood fiber length distribution as given in Fig. 13. However, the successive planes toward the rear of this figure show the distribution of segment lengths for various angles at which the fiber may be considered to have been 'broken' as discussed above. Thus the front plane has, for 180 degrees, the case of no breaking of the fibers. Each successive plane shows the result of a three degree change in the breaking criterion. In the rear plane, all bends of any angle greater than 0 degrees result in segments or breaks occurring, hence there are a large number of short segments obtained. Between the front and rear planes is found a great deal of information on both the length distribution and the curl distribution (in the last of the senses discussed above).

We can compare the three dimensional surface for the two softwoods and the hardwood. We see distinctions and similarities in the two softwoods, but a very significant difference between them and the hardwood surface of Fig. 16. One interpretation of Fig. 16 is that it shows a set of short, 'stiff' fibers. They are 'stiff' because for those planes in front of that plane corresponding to about 45 degrees there is virtually no change in the length distribution.

Other interpretations of these surfaces as descriptions of the morphology of these three pulps are possible, as are other definitions of the various morphological properties. Our purpose here is, however, to illustrate the range of definitions possible and to suggest by example, the ease with which these definitions can be modified and combined to yield insight into measurement methods for characterizing fiber morphology.

PATTERN RECOGNITION ALGORITHMS

The measurements described in the last section made critical (and we hope effective) use of the semiautomatic tracing data for fibers that we obtained with computer graphics tools operating on photomicrographs of pulps. Although these measurements illustrate the possibilities of algorithmic based measures, they do not yet demonstrate how these data may be obtained automatically and how the algorithm derived resulting automatic measurements may be obtained. For this, it is necessary to use automatic scanning methods and Pattern Recognition. We shall demonstrate here how such automatically scanned fiber images as that of Fig. 4 may be analyzed. It will also be necessary to show how the basic pattern recognition step performed by the human operator in selecting and separating distinct fibers in a fiber image may be accomplished by a computer.

We have shown, above, how a rectangular array of optical densities representing the fiber image in Fig. 4 can be obtained. A common method for extracting a pattern from such an automatic scan is to use density thresholding. Typically, a density value, suitably chosen from the range of densities that occur in the image, is used to distinguish all pixels with densities below that threshold from those at or above the threshold. This yields a binary image containing zeros for the points below threshold and ones for the remainder. From this binary image, it is possible, by means of several different algorithms, to find sets of connected binary ones. Two ones are connected if they are adjacent in the horizontal vertical, or diagonal direction. All the binary ones connected to some initial one or to those connected in turn to them, recursively, are treated as a single connected object, or blob. A blob is, then, a candidate for consideration as a fiber.

We see some of the problems that occur with this method in Fig. 17. Here we have chosen, for a threshold, the mean value of the densities that occur in the scanned image, namely 157. Then choosing the first pixel with density not less than the threshold we extract the blob containing that pixel. What we notice in Fig. 17 is that the blob thereby extracted contains three fibers. Two of these fibers are overlapping, and the remaining one is connected by a region in the image containing some dark parts that are over the threshold, thereby connecting two otherwise disjoint blobs. A more judicious choice of threshold can solve the problem of spuriously connected objects. In Fig. 18 we see the effect of merely raising the threshold by one density level to 158. Here the long fiber becomes separated from the two

Figure 17: Boundary points of single connected blob extracted from Figure 4 by thresholding at mean density, 157

Figure 18: Boundary points of single connected blob extracted from Fig. 4 by thresholding at slightly higher density, 158

crossed fibers. Thus selecting a point of the crossed fibers no longer causes the spurious fiber to be included. There still remains the problem of separating the crossed fibers. We will not treat this problem here, but defer it to a later study.

Another problem in automatic measurement and recognition of scanned images is seen in the three successive thresholdings of Figures 17, 18, and 19. Using thresholds of the mean, mean + 1, and mean + one standard deviation we get different extracted connected blobs. What we actually display in these figures is not the blobs themselves, but rather the results of a boundary tracing algorithm.

Once a blob has been extracted and exists as an image of 1's and 0's, it is possible to choose a single binary pixel on the boundary of the blob. This is a pixel which is a 1 and is adjacent to a 0 , unlike the interior pixels, all of which are surrounded by other ones. Starting with this pixel, it is possible to trace successively from one boundary pixel to another, going completely around the blob until the

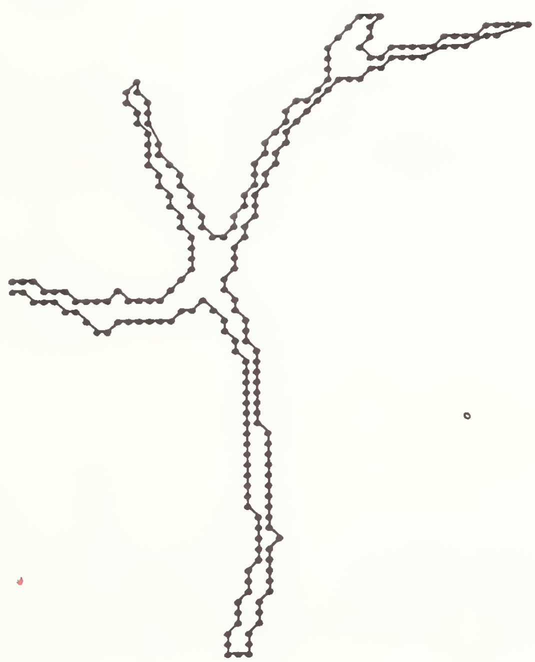

Figure 19: Boundary points of single connected blob extracted from Fig. 4 by thresholding at mean density plus 1.0 std. dev., 164

original pixel is encountered. The sequence of such automatically traced pixels is what we display in the three figures. Thus the boundary sequence of pixels is somewhat analogous to the set of tracing points obtained in the semi-automatic data acquisition case. Actually, it is necessary to obtain a set of points in the interior of the fibers to be strictly analogous to the semi-automatic case. We have connected this boundary sequence in the order that the boundary pixels are encountered to produce Figures 17, 18, and 19.

It is possible to make some useful measurements on these boundary sequences, and on the blobs from which they are obtained. First, by counting the pixels in the blobs, we can measure the area of the fibers as they are projected on the focal plane of the microscope. Since the images are not yet properly segmented, these areas are not meaningful as measures of individual fibers. The areas thus determined for the three thresholds of 157, 158, and 164 are, respectively, 65691, 26041, and 15268 pixels. Since these measurements are not calibrated in dimensional units, they should be viewed as only relative area measurements. We do note, as expected, that with increased threshold, the area of the objects measured decreases.

Another useful measure is the perimeter, or boundary length of the blobs. This is obtained by measuring the distance from one boundary pixel to the next and summing all these distances. These distances will all be proportional either to unity or the square root of two, corresponding to pixels that are adjacent vertically and horizontally or diagonally. In terms of the same relative units used in measuring the areas, we get boundary lengths of 5272, 2477, and 2272. The number of boundary points that contribute to these lengths are 500, 239, and 217 respectively. These results are summarized in the following table:

Threshold	157	158	164
Number of Boundary Points	500	239	217
Boundary Length, P	5272	2477	2272
Area, A	65691	26041	15268
P^2/A	423.1	235.6	338.1

The last measure shown in the table is a commonly used measure of degree of roundness of blobs, having a minimum value of 4 Pi for a circular blob and larger values corresponding to blobs like the present ones which are long and thin.

The simple pattern recognition algorithms demonstrated here can be greatly improved. As with the analysis of semi-automatically obtained data, the automatic scan data can be analyzed with algorithms that perform significantly different operations on the data to produce measurements that are superficially similar. An example, occurring above, is the measurement of boundary or perimeter length. Two different measures can be seen to give perimeter lengths. One is the count of the number of boundary pixels. Another is the length of the set of straight line segments joining boundary points. Similarly with area calculations, the count of interior pixels can give results different from the area included in the sequence of straight line segments constituting one of the above perimeter definitions. These definitions must be compared and tested in order to choose the most useful one for the kind of measurement desired. As with the fiber separation problem mentioned above, these pattern recognition algorithms will be the result of a future study.

CONCLUSION

We have attempted to demonstrate how images of pulp fibers can be measured in a variety of ways. No one of these ways is claimed to be superior to another, but we do claim that the different measurements are sufficiently different that a basis exists for choosing among them, given that criteria exist relating to how the measurements are to be used. Since all these measurements are based on algorithms written for large general purpose computers, they can not only be reproduced and the results duplicated, but they can also be improved in a special sense. In this sense, the improvement is in efficiency of performance of the measurement algorithms without changing the functional nature of their behavior.

When a new technology like image processing and pattern recognition is introduced to the pulp and paper field, it is important that exploratory investigations be permitted. It is also important that the results of this exploration be directly useable once assurance is available that the nature of the measurements to be made is well understood. The use of algorithms as the facilitator of this conversion from exploration to production has been demonstrated here. We hope that the subsequent stages in this process will be carried on by our colleagues in the pulp and paper industry and the electronics industry.

ACKNOWLEDGEMENTS

We would like to thank our colleagues E. L. Graminski for preparing the fiber image data used in these experiments and, in general, for suggesting more useful measurements to be made than even the powerful tools reported here can accommodate; D. J. Orser who demonstrated many times that algorithms that are elegant and carefully written also are economical and work; D. E. Rabin and D. A. Miller who helped get empirical data from one computer to another and from the computer into a form that can be understood by lesser beings like the author; and Mrs. S. Lightbody who, by preparing the manuscript has made it possible for the rest of the world to determine whether what we say makes any sense.

REFERENCES

1. Kirsch, R. A. et al, Experiments in Processing Pictorial Information with Digital Computers, Proc. 1957 Eastern Joint Computer Conf., Assn. for Computing Machinery, pp. 221-229 (1957).

2. Electronic Industries Assn., Computer Image Processing Equipment Buyer's Guide (1977), EIA Government Div., 2001 Eye St., N.W., Washington, D.C., 20006

3. Graminski, E. L. and R. A. Kirsch, Image Analysis in Paper Manufacturing, IEEE Computer Society Conf. on Pattern Recognition and Image Processing (June 6-8, 1977).

4. Kibblewhite, R. P., Structural Modifications to Pulp Fibers: Definitions and Role in Papermaking, TAPPI 60 (10):141 (1977).

NBS-114A (REV. 7-73)

U.S. DEPT. OF COMM. BIBLIOGRAPHIC DATA SHEET	1. PUBLICATION OR REPORT NO. NBSIR 78-1442	2. Gov't Accession No.	3. Recipient's Accession No.

4. TITLE AND SUBTITLE Algorithms For Image Analysis of Wood Pulp Fibers	5. Publication Date January 1978
	6. Performing Organization Code
7. AUTHOR(S) Russell A. Kirsch	8. Performing Organ. Report No.
9. PERFORMING ORGANIZATION NAME AND ADDRESS NATIONAL BUREAU OF STANDARDS DEPARTMENT OF COMMERCE WASHINGTON, D.C. 20234	10. Project/Task/Work Unit No.
	11. Contract/Grant No.
12. Sponsoring Organization Name and Complete Address *(Street, City, State, ZIP)* same as item 9	13. Type of Report & Period Covered final
	14. Sponsoring Agency Code

15. SUPPLEMENTARY NOTES

16. ABSTRACT *(A 200-word or less factual summary of most significant information. If document includes a significant bibliography or literature survey, mention it here.)*

Image analysis technology can be used to measure the visible morphology of pulp fibers. But before such measurements can be accepted, it is necessary to achieve precise definition of the necessary measurements in the form of suitable algorithms that have been experimentally tested on images of actual fiber data. We present such measurement results on both semiautomatically traced fiber data and on automatically scanned images. We explore the variety of definitions possible for some simple well known properties of wood fiber morphology by applying suitable algorithms to fiber image data. Finally, we suggest that this exploratory approach to the specification of the precise image analysis measurements needed in paper manufacturing can facilitate the introduction of a technology for process control that will result in savings in paper manufacturing cost and in reduction of energy requirements.

17. KEY WORDS *(six to twelve entries; alphabetical order; capitalize only the first letter of the first key word unless a proper name; separated by semicolons)*
Algorithms; artificial intelligence; image analysis; morphological analysis; paper fibers; pattern recognition; pulp characterization.

18. AVAILABILITY	19. SECURITY CLASS (THIS REPORT)	21. NO. OF PAGES
[XX] Unlimited [ ] For Official Distribution. Do Not Release to NTIS	UNCLASSIFIED	30
[ ] Order From Sup. of Doc., U.S. Government Printing Office Washington, D.C. 20402, SD Cat. No. C13 [XX] Order From National Technical Information Service (NTIS) Springfield, Virginia 22151	20. SECURITY CLASS (THIS PAGE) UNCLASSIFIED	22. Price

USCOMM-DC 29042-P74