DICTA2002: Digital Image Computing Techniques and Applications, 21--22 January 2002, Melbourne, Australia.

Linear Feature Shape Characterization Using Structured Sampling

Anthony Maeder and Karen Whelan

School of Electrical & Electronic Systems Engineering, Queensland University of Technology, GPO Box 2434 Brisbane Qld 4001.
a.maeder@qut.edu.au, k.whelan@qut.edu.au

Abstract

Shape characterization of major features in digital images of physical structures often requires considerable processing effort to cover the whole image. This paper describes an approach suitable for a class of simple features, which considers only a small fraction of the total image pixels in the computation. The sampling process uses randomly selected starting points as seeds for a local connectivity analysis, which establishes an estimate for the shape characteristics of linear features in the image. An example of a materials problem involving a metal alloy is used to compare the results of this approach with exhaustive analysis.

1. Introduction

Many problems in science and engineering require detailed inspection of images of samples of physical structures (such as biomedical or materials analysis) to locate, characterize and compare visible features of the often inhomogeneous medium. For example in petrology, the analysis of chemical processes leading to substance formation depends on assessing the dispositions, proportions and variabilities of different crystalline elements. A wide range of image processing methods (eg thresholding, segmentation, morphology) and shape analysis techniques (e.g. individual properties, boundary coding) are routinely applied to such images.

A particular problem in materials image analysis relates to the study of microstructures containing phases that exhibit “flaky” morphology. In polished samples the flaky phase appears as grouped lines often referred to as “wheat sheaf” particularly in the aluminium-silicon alloy system [2]. Aluminium-silicon alloys are widely used in a number of applications and the properties that they exhibit in service will depend on the nature of the phase structures formed during solidification from the liquid. It is possible to directly relate the service behaviour of the alloys to parameters that are extracted from two-dimensional images of the microstructure [1, 5].

It is usual to provide statistical distribution parameters alongside any quantitative results for these shape features, to indicate the extent of the variability caused by the sample inhomogeneity. This variability arises from the imperfect reaction aspects of the materials formation. The expense of processing these images by visiting every pixel in scan-order possibly many times, is most undesirable if many samples must be analysed or if real-time performance is required.

If an approach could be devised which only considers a small fraction of the pixels in the image, but nevertheless produces acceptable results by comparison with exhaustive processing, much computational effort could be saved. Previous work in random sampling of image pixels followed by structured sampling to perform neighbourhood search has previously been applied successfully to both texture [3] and shape [4] analysis. In these cases, good results were obtained when only 0.05% to 0.1% of the image pixels were randomly sampled, in characterizing simple image content features such as area, perimeter, separation or orientation.

2. Method

The work described in this paper applies a structured sampling approach to the characterization of linear features (eg phases exhibiting the flake morphology as occurring in some metal alloys). Some of the parameters of interest for these features are length (measured as the Euclidean distance between the line endpoints) and the angle (measured anti-clockwise from the horizontal). It will be assumed that the images are converted to binary pixel values (background and foreground) and are sufficiently noise-free that most linear features of interest are not accidentally broken. The linear features are assumed to be thin (from one to a few pixels wide) and straight, and un-linear features (such as small blobs) are assumed to be sufficiently infrequent that they are ignored.

The conventional method for processing such images to extract linear features is by exhaustive scan region labelling. In this approach, the pixels are visited in the usual line-sequential scan ordering, and when a new foreground pixel is encountered, it is assigned a unique region number. All pixels adjacent to any new foreground pixel belonging to that region are inspected to determine whether they are also foreground pixels connected to the region, and if so they are marked as part of the region. The inspection process can be undertaken recursively or sequentially across a scan line. This method of feature extraction requires each pixel to be visited at least once (if it is a background pixel far from a line) and at most 8 times (if it is a background pixel almost completely enclosed by a region). The minimum and maximum x and y values corresponding to extreme pixel positions of each region can be easily recorded. These values provide the means to calculate a good approximation for the length and angle of linear features directly. By traversing the boundaries of the extracted regions, various estimates of the linearness of the feature (such as average width and straightness) can also be computed, but this process is rather tedious.

The structured sampling method first randomly selects pixel locations for inspection. If the pixel has the background value it is ignored. If it has the foreground value, the adjacent pixels are then inspected to determine a first approximation for the line angle. Once the approximate angle has been determined, pixels adjacent to those of the initially adjacent pixels, in that direction, are selected for inspection. If they are background pixels, their adjacent neighbours are inspected until a new foreground value is found or the search terminates with all neighbours found as background. If a foreground pixel is found, the next adjacent pixels in the updated approximate angle direction are selected for inspection and the process is repeated. In a majority of cases, only one pixel needs to be inspected (along the direction of the approximate angle), sometimes two pixels (if the angle is not an integer multiple of 45 degrees) and occasionally three pixels. This method therefore visits approximately the number of pixels that lie on the centre of all lines in the image, plus a constant percentage overhead based on the average region angle, and on the proportion of background to foreground pixels over the whole image.

An example of how the structured sampling method might be applied is shown in Figure 1, for a synthetic image of size 8x8 pixels. Assuming the origin (0,0) is at the top left corner, this image can be seen to contain a linear feature located between (1,2) and (6,4), using (row,column) notation.

0	0	0	0	0	0	0	0
0	0	1	0	0	0	0	0
0	0	1	1	0	0	0	0
0	0	0	1	1	0	0	0
0	0	0	1	1	0	0	0
0	0	0	1	1	0	0	0
0	0	0	0	1	0	0	0
0	0	0	0	0	0	0	0

Figure 1: Sample 8x8 image

The sequence of pixels inspected by the structured sampling method after the detection of a new region at pixel (1,2) is as follows (hits on region pixels are shown as bold font):

(1,2), (1,3), **(2,3)**, (3,4), **(3,3)**, **(4,3)**, **(5,3)**, (6,3), **(6,4)**, (7,4), (7,5)

In this case, 11 inspections were undertaken, of which 6 yielded foreground pixels in the linear region. In comparison, the conventional exhaustive method would inspect pixels in the following sequence:

(1,2), (1,3), (2,0), (2,1), **(2,2)**, **(2,3)**, (2,4), (3,1), (3,2), (3,2), **(3,3)**, **(3,4)**, (3,5), (4,2), **(4,3)**, **(4,4)**, (4,5), (5,2), **(5,3)**, **(5,4)**, (5,5), (6,2), (6,3), **(6,4)**, (6,5), (7,3), (7,4), (7,5).

In this case, 28 inspections were necessary, of which 10 yielded foreground pixels in the linear region. In addition, the exhaustive method would require all remaining pixels in the image to be visited, to confirm whether any other foreground pixels existed. This would result in a further 28 pixels being inspected (in addition to the 10 previously inspected before the region was first detected). All these would be background pixels in this case, as this example deals with only one region. In practice, the start of the search would be based on a random selection of a pixel position, which if of foreground value would on average occur in the centre of a linear region. As the search for endpoints must be conducted in both directions, this is equivalent to starting at an endpoint. However, if only some of the total image pixels are selected via random sampling as possible region pixels, this will reduce the total number of pixel inspections correspondingly.

Cost-benefit is the most immediate question requiring examination, in assessing the usefulness of the structured sampling method over the conventional exhaustive method. If an image with N pixels in total contains fraction P of pixels as foreground, the exhaustive method will visit on average P*N*3 (or worst case P*N*4) potential foreground pixels, with on average P*N*3/2 of these being found to be background pixels. Checking for start of this and any further regions in the image requires inspecting (1-P)*N-P*N*3/2 background pixels. The structured sampling method which samples fraction F of the total number of pixels will visit about F*P*N/2 foreground and (F*P*N)/2+(F*(1-P)*N) background pixels. This method accurately finds the endpoints of lines, but is limited in accuracy by how broad a range of significant line length and angle values there are in the image, the broader ranges requiring a much larger sample (ie larger F) to achieve a good estimate.

3. Results

In order to test the method, a real world example from materials analysis was used. Figure 2 shows a high resolution image (approximately 2M pixels) of a section of aluminium silicon alloy with extensive linear feature formation, used to generate the results discussed here. It can be seen that this is a challenging problem, with a range of different line lengths and angles evident. The data in Tables 2 and 3 show the structured sampling results obtained for line lengths of L units and angles of A degrees for two different sample sizes (i.e. F values). These results should be compared with the distribution of actual values obtained by the exhaustive method, shown in Table 1.

An influential factor in collecting and interpreting the results is the degree of quantization or "binning" of lengths and angles into single representative values. This is necessary to compact the individual measurements obtained for each case of a linear feature detected, into gross classes, which are more representative of the whole distribution of actual feature characteristics. In this case, bins of 5 units of length and of 22.5 degrees of angle were adopted. This leads to some arbitrary clustering of those values close to these thresholds, which gives a more spread result in the sampled case due to the effects of using less data than the entire population.

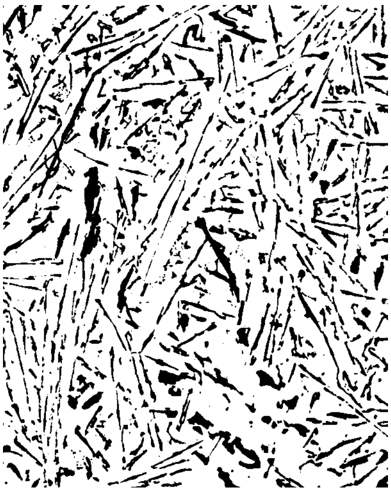


Figure 2: A materials sample containing linear features.

Table 1: Actual distribution of lengths and angles for all lines of Figure 2.

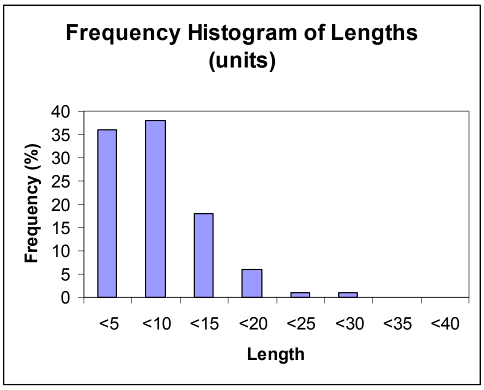
Length (units)	%	Angle (degrees)	%
<5	36	<22.5	4
<10	38	<45	6
<15	18	<67.5	30
<20	6	<90	14
<25	1	<112.5	3
<30	1	<135	23
<35	0	<157.5	15
<40	0	<180	4

Table 2: Results for Figure 2 using structured sampling (F=0.01).

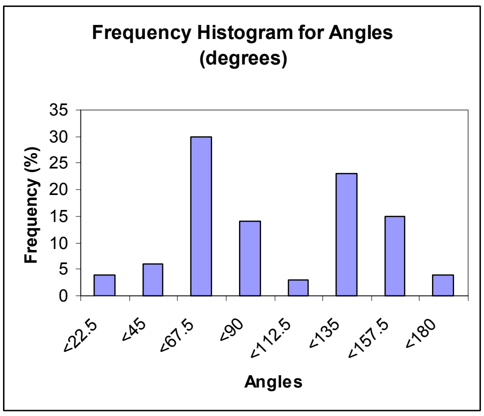
Length (units)	%	Angle (degrees)	%
<5	21	<22.5	10
<10	28	<45	21
<15	15	<67.5	19
<20	8	<90	8
<25	4	<112.5	8
<30	2	<135	19
<35	2	<157.5	4
<40	2	<180	6

Table 3: Results for Figure 2 using structured sampling (F=0.02).

Length (units)	%	Angle (degrees)	%
<5	23	<22.5	9
<10	38	<45	19
<15	12	<67.5	22
<20	14	<90	14
<25	5	<112.5	9
<30	1	<135	15
<35	1	<157.5	4
<40	1	<180	4

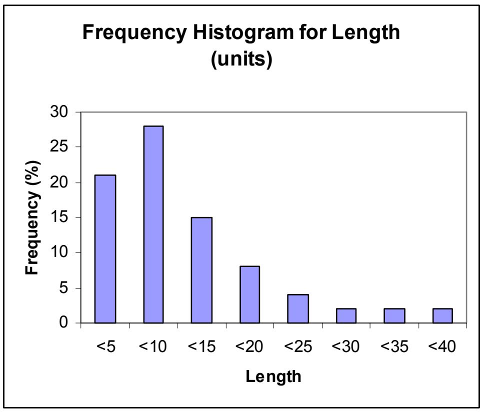


(a)

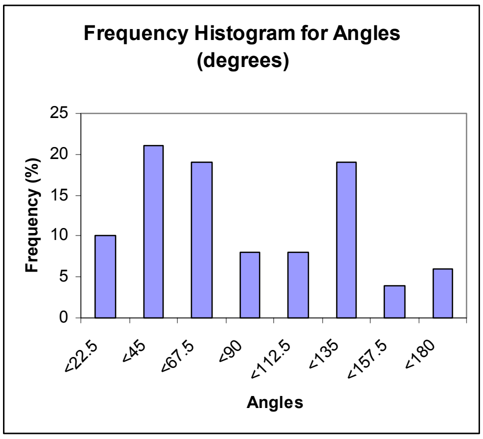


(b)

Figure 3: Frequency Histograms for actual distribution of all lines in Figure 2.

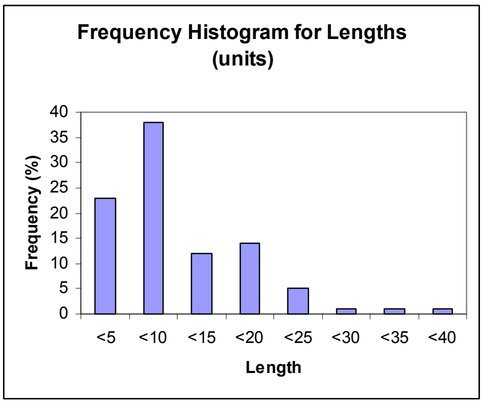


(a)

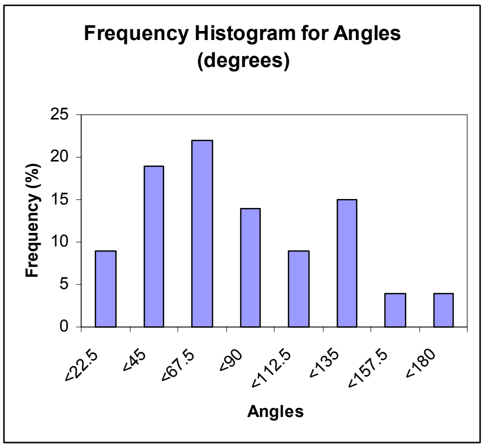


(b)

Figure 4: Frequency Histograms for Table 2 (structured sampling F=0.01).



(a)



(b)

Figure 5: Frequency Histograms for Table 3 (structured sampling F=0.02).

It can be seen from the results shown that at both the F values (0.01% and 0.02%) the structured sampling results are fairly reliable for both the lengths and for the angles. A bimodal distribution exists for the angle data, and a unimodal length distribution exists for length data in all three Tables of results. Artifacts of the low sampling rate are present: a broadening and shift in peaks for angle distributions, and a stretched tail plateau in the length distribution. There is little evidence that these types of artifacts will diminish rapidly as sample size increases, until a very substantial fraction of the image is being sampled (which in turn leads to much less computational cost saving). It is therefore necessary to make allowance for inherent spread when using results obtained from structured sampling.

4. Conclusion

The results indicate that it is feasible to use structured sampling for problems of this type at a very economical sampling rate. The parameters L and A were treated independently, but obviously the correlation between them could be determined as well by this method. Some improvements could be made to reduce the structured sampling cost further, if interval halving or other forms of prediction of line length could allow greater leaps across unvisited pixels. In the future, it would be interesting to compare these results with other common methods for line extraction (eg Hough transform).

5. References

[1] Caceres, CH, Davidson, CJ, Griffiths, JR, The Deformation and Fracture Behaviour of an Al-Si-Mg Casting Alloy, Materials Science and Engineering A – Structural Materials Properties, 197(2), 1995, pp171-179.

[2] Elliott, R, "Eutectic Solidification Processing: Crystalline and Glassy Alloys", Butterworths, :London, 1983.

[3] Maeder AJ, Texture characterization using random sampling, in Earnshaw RA & Wyvill B, "New Advances in Computer Graphics", Springer Tokyo 1989 p 603-612.

[4] Maeder AJ, Jenkins BM, Skottner LV & Rui Z, Shape characterization of particles in digital images using structured sampling, Proceedings of DICTA-95, Brisbane 6-8 Dec 1995, p607-612.

[5] Prasad, BK, Venkateswarlu, K, Modi, OP, Yegneswaran, AH, Influence of the Size and Morphology of Silicon Particles on the Physical, Mechanical and Tribological Properties of Some Aluminium-Silicon Alloys, Journal of Materials Science Letters, 15(20), 1996, pp1773-1776.