PROCESSING INFORMATION FROM SCANNING INSTRUMENTS
P. Hawkes

To cite this version:
P. Hawkes. PROCESSING INFORMATION FROM SCANNING INSTRUMENTS. Journal de Physique Colloques, 1984, 45 (C2), pp.C2-195-C2-200. <10.1051/jphyscol:1984244>. <jpa-00223957>

HAL Id: jpa-00223957
https://hal.archives-ouvertes.fr/jpa-00223957
Submitted on 1 Jan 1984

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
PROCESSING INFORMATION FROM SCANNING INSTRUMENTS

P.W. Hawkes

Laboratoire d'Optique Electronique du C.N.R.S., 28, rue Jeannes Marvig,
BP 4347, 31065 Toulouse Cedex, France

Résumé - Les tendances actuelles et quelques développements possibles dans le traitement des signaux provenant du MEB et du STEM sont analysés succinctement.

Abstract - Current trends and future possibilities in the processing of SEM and STEM signals are briefly surveyed.

I - EARLIER WORK

Digital processing of the image created by fixed-beam instruments suffers from the handicap that a time-consuming and expensive step separates images from computer: photographic recording and processing followed by digitization (and of course quantization). Strong though the arguments for digital processing may be, the proponents of optical methods can always retort that their channel is capacious and accepts a two-dimensional brightness distribution as input, two very attractive features. With the advent of scanning electron microscopes in the mid 1960's, however, the intermediate step evaporated and by 1968 papers describing image analysis of SEM pictures had begun to appear — indeed the first volume of the Scanning Electron Microscopy series contains a paper entitled "Computer Processing of SEM Images" (51). The reason is obvious: all scanning microscopes generate sequential signals, which are then used to write an image, usually on a CRT screen. It is thus comparatively easy to intercept this signal and modify it before allowing it to be displayed or, more ambitiously, to store it in computer memory and perform complex measurements or other operations on it before releasing it to a visual display device.

The first decade of SEM image processing was largely devoted to image analysis and simple types of image enhancement, essentially the matching of the image contrast range to that of the eye by expanding or contracting it as necessary and adjusting the mean brightness level. Image analysis reached a high level of sophistication, encouraged by the proliferation of SEM signals, and was routinely applied to many practical problems from fields as far apart as neurological disease and the mineralogy of coals. For the latter purpose, a specialized unit was designed at United States Steel (23, 29-31) which manipulates the signals generated by backscattered electrons, secondaries, transmitted electrons when possible and an energy-dispersive X-ray detector. Image features can be recognized by thresholding or topology after which the chemical and geometrical characteristics of the various regions pinpointed in the pattern recognition step can be tabulated or displayed. SEM image analysis needs a much more detailed review than can be given here, for so many specialized techniques have been developed in different fields that we are obliged to be unfairly selective. A generous, though still invidious, list was given in (17); before turning to the STEM, we just mention that the work on directionality in SEM images, a good example of the development of a technique for the needs of a very specialized field of study, has been surveyed not only in Scanning Electron Microscopy (50) but also in a recent book (42). Another elegant application is topographical surveying using...
an automatic focusing unit (22). This first decade also saw the arrival of commercial STEMs, both dedicated (VG Instruments and the short-lived Siemens ST 100 F) and as modifications to conventional TEMs. All the techniques already devised for the SEM are of course applicable here, provided that the signals in question are available, but some fundamentally new methods were added to these. In particular, the peculiar mode of image formation in the STEM, whereby a diffraction pattern of the pixel under the beam at any instant is created in the detector plane and subsequently sampled and/or integrated, has been a rich source of ideas. From the very beginning, the STEM detector was divided into a central disc and a surrounding ring, which collect to a good approximation an elastic dark-field image (ring) and an inelastic-plus-un-scattered bright-field image (disc), with the possibility of further electron energy subdivision by means of an energy analyser. In 1974, however, Dekkers and de Lang (7) made a highly original and ingenious suggestion, namely, that the detector should be divided into two semi-circular discs; by simple arithmetic operations on the resulting signals, some features of the specimen phase distribution would be mapped into signal variations that can be visualized directly. In the same year, Rose (35) also proposed subdivision of the detector to combat the aberrations of the probe-forming lens and a host of suggestions followed, reviewed in (15). All these proposals arise from a feature of the STEM that has no convenient analogue in the CTEM, namely the possibility of forming any desired weighted superposition of the current at each point of the far-field diffraction pattern of each individual pixel. When the earlier suggestions were made it was assumed that the weighted superposition would in practice be achieved by altering the detector geometry but not its response, so that the weights would be zero or (conventionally) unity. In recent years, however, the arrival of framestore memory and commercial systems for exploiting it has made it feasible to store the intensity distribution of the diffraction pattern from each pixel as it is generated and to perform one (or even several) weighted superpositions with arbitrary weights and of course geometries (46). This degree of flexibility makes it reasonable to contemplate using much more elaborate detector designs, such as the optimum detector proposed by Huiser and van Toorn (24), and the multiple segment geometries that hitherto seemed rather extravagant (since the segment pattern would vary with the operating conditions). The difference between the two situations is not unlike that between optical and digital processing: by using a detector of specific geometry, however complicated, we gain in speed and ease of operation but lose in flexibility; by measuring the entire diffraction pattern from each pixel, we have almost total flexibility but must be wealthy enough to buy the necessary storage units and peripherals. For further information on these points, see (2, 8-11, 20, 21, 25-28, 33, 38, 40, 43-45, 47, 49).

II - RECENT DEVELOPMENTS

2.1. Architecture

The last paragraph has taken us beyond the "earlier studies" for the use of frame-stores and the performance of complex operations on-line are distinctly more recent developments. There are a number of good accounts of the benefits of this newer hardware (See list at end of §I) and at least one system specifically intended for use with scanning electron microscopes is available commercially (from Toltec Computer Ltd). Many processing operations require large matrix transforms, and special purpose array processors are clearly attractive when many such transforms have to be performed, in iterative processing schemes for example. It is not quite so obvious what the large vector machines, the CRAYS and the Cybers (not to mention their imminent Japanese rivals), have to offer. The only work in electron microscopy even remotely relevant is that of Arnot et al (1), who used not a vector machine but a fast parallel processor (the ICL "DAP") to speed up tasks involving frequent, large, two-dimensional Fourier transforms. Nevertheless, it is clear that image processing is a task well-adapted to the architecture of the vector machines. It is easy to adapt the size of the digitized image to the preferred length of the vectors (the efficiency of such machines is acutely dependent on the match between the number of operations that can be performed in parallel on a particular computer and the sizes of the blocks of input data that are to be processed in parallel). There is no lack of calculations in which there is relatively little input-output (very wasteful compared with the actual computing) from among which we may cite:
multi-slice simulations (41), Gerchberg-Saxton or other algorithms to solve the phase problem, for example. At present, programs written in the familiar high-level languages, FORTRAN especially, can be run with relatively little modification on vector machines, sadly retrograde though this seems to specialists in comparative (computer) linguistics: "The history of the development of parallel programming languages would appear to be repeating many of the mistakes which occurred in the development of sequential languages. Many of the existing parallel languages have not benefited from advances which have been made in programming language design and implementation techniques" writes Perrott gloomily (32); he goes on to observe that "most programmers and researchers using these machines [vector and array processors] are expected to tackle a task on a machine of the latest hardware technology using a comparatively inferior software tool" and concludes "it is now possible to design a language which can exploit parallelism in the algorithm for the solution of a problem. In this way the user will be freed from the peculiarities of a parallelising compiler or a hardware-dominated syntax". We have quoted this paper at some length for it seems reasonable to anticipate that at least some image processing tasks will be confided to vector machines in the next few years and the question of transferring languages such as SEMPER (39), SPIDER (12), IMAGIC (18), or EM (19) (not to mention the eighty-odd other image-processing languages listed in (34)) will arise. It would be a great benefit to the user and, surely, to the general efficiency of the combination of language and machine if the versions transferred were written in a language such as that evoked by Perrott, intimately adapted to the nature of the computer architecture.

2.2. Coding

We have already mentioned briefly the use of digital framestores for short-term image storage but as Burge has pointed out (3), the problem of multiple-access long-term storage has by no means been solved and is likely to become a major impediment to the marriage between microscope and computer (already "joined together"). Hitherto, comparison between micrographs obtained in different laboratories has been made on the basis of published pictures or on individual collaboration. It would clearly be advantageous if institutions with common interests could deposit their images in a database, which would be consulted in the same way as literature searches are made in bibliographic bases, for example. Nevertheless, it is (at present) unthinkable that images, or even just the most interesting regions of images, should be stored in any quantity as raw arrays of typically eight-bit numbers. The problem of efficient coding has long been studied for one-dimensional signals and in recent years this has been extended to two-dimensional (and multi-dimensional) arrays. Of the various approaches to such coding, we mention two types: orthogonal transform coding and vector quantization. The principle of orthogonal transform coding is easy to state, less easy to put into practice: given a two-dimensional array (or set of arrays), can we find a reversible matrix transformation such that the elements of the resulting array are all uncorrelated? In other words, given that the original array contains much redundant information, how can we retain only significant data and jettison the rest? The formal solution to this question is known: for an $N \times X$ picture array $f$, the expansion coefficients $F(u, v)$, $F(u, v) = f \phi_{uv} \ , \ f = \sum_{uv} F(u, v) \phi_{uv}$ will be uncorrelated if the set of $N^2$ matrices $\phi_{uv}$ are the eigenmatrices of the autocorrelation function (matrix) of $f$, which we denote by $R$:

\[ N-1 \sum_{n=0}^{N-1} R(m,n,p,q) \phi_{uv}(p,q) = \gamma_{uv} \phi_{uv}(m,n) \]

(A proof of this is to be found in (36, §5.1).) Efficient this may be but convenient it certainly is not, for unlike the Fourier and related transform matrices, the expansion here is performed in terms of matrices the elements of which are a function of those of $f$: $R$ is determined by $f$ and the set $\phi_{uv}$ by $R$. Before coding any image using this "Karhunen-Loève" transform, the $\phi_{uv}$ would have to be found and worse, anyone consulting a database in which the images were so coded would have to know
the $N^2 \phi_{(u,v)}$ for each before being able to examine them.

Fortunately, there are other orthogonal transforms that are only marginally less efficient at image compression than the Karhunen-Loève transform and have fixed transform matrices and fast algorithms. The one that is closest to the Karhunen-Loève in performance is the discrete cosine transform, which has various attractive properties. Both Karhunen-Loève and discrete cosine image coding have been tested on electron microscope images. We refer to (4-6, 37, 52) for further details. A wasteful aspect of these coding methods can be eliminated and improve them still further: there is no need to use the same number of bits for every pixel. This idea is also exploited in "error-free" compression, where (unlike the transform methods) no information, however unimportant is discarded but shorter binary code words are allocated to the grey levels that occur more often in the picture. Once again we refer to (36) for details of this and the associated use of Huffman codes, or shift codes. The other procedure for storing images compactly to which we draw attention is vector quantization (13, 14). Here, instead of quantizing the grey levels of individual pixels, an ordered set of $i$ samples (an $i$-component vector) is mapped into one of a finite set of vectors. Thus instead of handling individual pixels, we now treat whole families of grey-level values as a single (vector) quantity. Each of these output vectors has a name, in the form of as short a binary word as possible, and it is these that are stored. At first sight, this is not particularly attractive, since the codebook required to decipher the picture might easily be huge. Recently, the notion of lattice quantizers to impose a pattern on the coding step has been introduced (13) and this appears to offer a means of "circumventing the need for a code book". Gersho concludes that "the subject of vector quantization is becoming increasingly important and a deeper understanding of the structural properties should play an important role in future studies of complexity, algorithm design and performance capabilities" (14).

2.3. Recursive methods

We have had occasion to mention the redundancy of much of the information in a typical image: in a landscape, for example, areas of blue sky or road, or walls or fields (at low resolution) might be essentially uniform, or such that a small region is typical of a larger one - a flock of sheep say. Mutatis mutandis, the same may well be true of an electron micrograph whence the success of transform coding methods. This notion also underlies one class of methods of restoring images degraded by noise and by the effect of a non-uniform transfer function. These methods are recursive, in the sense that information from adjoining pixels is assumed to be correlated to some extent and the grey-level values of pixels already acquired are used to correct each new value as it arrives. These techniques were originally developed for temporal signals, in which the notions of past (acquired), present (arriving now) and future have their obvious everyday meanings. When we move from the temporal to the spatial domain, however, the situation becomes more confused. If we consider a signal from a scanning microscope, then the temporal sequence is of course respected but the nearest neighbours of any given point are not only the points just acquired - there will be three points in the line above at least as near as the two points just acquired in the same line. It is important to incorporate into the recursion the fact that the picture statistics are truly two-dimensional. For this reason, recursive schemes that operate line by line rather than point by point have been proposed, which implies that adequate buffer storage must be available. Once the notion of point recursion has been abandoned, the division into past and future likewise becomes artificial and there is no reason why the correction should not be performed in a homogeneous fashion, using all the four (or even eight) nearest neighbours of each point, as in relaxation methods of solving partial differential equations. These methods have an extensive literature. They have some features that make them attractive for electron image processing, among which the fact that regions of any shape can be used is not the least interesting: jagged edges, even holes, are harmless though it is of course essential to operate in real space if such irregularities are present. Generally speaking, the methods are more efficient in reciprocal space but real space formulae are just as easy to derive. There is however an extra degree of complexity in electron microscopy, namely, the fact that at high resolution, the image intensity may well be the sum of two terms, each containing a transfer function and a specimen transparency function. It is not difficult to extend
the calculations formally to this situation but the statistical quantities that occur in the recursion formulae are now very difficult to measure or estimate. It is not yet clear to what extent, and more important, with what accuracy, these parameters can be estimated, by methods such as those used by Suresh and Shenoi (48) to obtain transfer function parameters. Nevertheless, it is probably possible to avoid the problem by methods such as those used in Wiener-Schiske filtering of focal series of TEM micrograph (16).

III - CONCLUDING REMARKS

The hardware and software of scanning signal manipulation have now reached a high degree of sophistication, as the following papers in this volume demonstrate convincingly. Let us hope that it will not be long before the speculations in the foregoing section are overtaken by reality.

REFERENCES


42. P. Smart and N.K. Tovey: *Electron Microscopy of Soils and Sediments*, Clarendon, Oxford 1982


