HISTORY OF THE CLASSIFYNDER

John Flenley
Massey University

In 1966 when I was a PhD student at the Australian National University, I often talked with my supervisor, Donald Walker, about the possibility of automating the identification of pollen grains. Palynologists spent so much time peering down microscopes, that it seemed to us difficult to justify. Computers at that time were developing rapidly, and had recently turned to the interpretation of visual images. So in October 1966 the CSIRO Computing Research Section sponsored a 'Workshop on Picture Interpretation', and Donald Walker urged me to contribute a paper. The proceedings were published as a volume of 24 papers, which the editors arranged in what they regarded as increasing order of difficulty. Palynology (Flenley, 1968) came next to last. The editors advised me to re-visit the problem in five years time (1973).

By then I was a lecturer at the University of Hull in the UK. I made friends with the electronic engineers (especially Gaynor Taylor) and put the palynology problem to them. They advised me to come back in another five years (1978). By this time the Scanning Electron Microscope was delivering beautiful pictures of pollen grains, showing their surface texture very clearly. Haralick *et al.* (1973) had published their ideas on texture analysis, leading to Haralick's (1979) book on the subject.

It now seemed to be the time to act, and Gaynor Taylor appointed a PhD student, Mitchell Langford, to apply Haralick's techniques to SEM images of pollen. This led to a most successful PhD thesis (Langford, 1988), preceded by a conference paper (Langford *et al.*, 1986), and followed by our first paper in the palynology literature (Langford *et al.*, 1990).

The SEM was expensive to buy and to operate. Possibilities of a 'desk-top' version never materialized, so the next problem was to start applying the ideas to optical images, which compared with the SEM meant lower magnification, lower resolution and less depth of focus. These problems were tackled by another PhD student at Hull, Walter Treloar. His PhD thesis (Treloar, 1992) showed a considerable advance, by applying Haralick's measures and Law's masks (Laws, 1990) to optical images, and also exploring software for finding pollen grains on a slide.

By then I had migrated to Massey University, and Walter Treloar was able to follow me with a post-doctoral scholarship from Massey. The results were presented at the INQUA Congress in Houston, Texas, and led to the introduction of neural network analysis, and achieved over 90% accuracy with the identification of 12 reference pollen types (Treloar, 1994, a, b).

Walter Treloar returned to a post in UK, and I then undertook a review of the research position with my daughter, Dr Eleanor Stillman, a lecturer in statistics at the University of Sheffield, UK. This culminated in a review paper (Stillman and Flenley, 1996), which set out the needs of palaeopalynology: more sites, fine resolution sampling, larger counts, rapid results, greater objectivity and finer determinations. It also set out the problems to be overcome: more taxa in datasets, classifier software, the need for operator-checking of results, the choice of variables to be measured, and statistical verification. Collaboration with

Sheffield continued and led to a presentation to the International Environmetrics Society (Fieller *et al.* (1968).

Meanwhile, I had obtained another postdoctoral fellowship from Massey, and had appointed Ping Li. He was a brilliant Chinese software specialist who had been forced to spend 10 years working in a coalmine during Chairman Mao's cultural revolution.

Ping Li improved our approach to neural networking and to the use of Law's masks. With technical help from Leighanne Empson and myself, he collected a larger database of images (both reference and fossil). We started to present the results at computer conferences (Li and Flenley, 1997; Li *et al.*, 1998; Flenley *et al.*, 1999). Results were now very promising, with up to 100% accuracy in identification being achieved, and so we put another paper in the palynology literature (Li and Flenley, 1999). I mentioned all this in a review paper (Flenley, 2003).

By now we were working in parallel with Professor Bob Hodgson who was Director of the School of Engineering and Technology in the Institute of Information Sciences and Technology at Massey. Bob was expert in optical technology and in computing, and had previously worked with Gaynor Taylor at Hull. Bob also obtained funding for a postdoctoral fellow and also appointed a Chinese postdoctoral fellow, Dr Zhang.

Zhang developed the neural network analysis still further, until no less than 43 parameters were being measured for each suspected pollen grain, and presented a paper on this at the International Conference on Artificial Intelligence in Las Vegas (Zhang *et al.*, 2003). Professor David Fountain, with an interest in modern pollen rain, also joined in this research.

It was again time to review the situation, which we did in three papers 'Towards automation of palynology 1, 2 and 3' published in the same journal (Treloar *et al.*, 2004; Li *et al.*, 2004; and Zhang *et al.,* 2004). Technical aspects were also published in the Proceedings of the International Symposium on Image and Signal Processing and Analysis, held in Zagreb, Croatia, in 2005 (Hodgson *et al.,* 2005).

We were now within sight of our goal, and applied for a Marsden Grant to complete the project. Two referees gave the project a Grade 1, and the third said it was impossible but gave it Grade 3 (out of 5). We were allowed to respond and quoted President Kennedy: "We want to go to the moon not because it is easy but because it is difficult." We did not get the grant.

Fortunately, Bob Hodgson was now able to supervise a Masters student, Gary Allen. He concentrated on the building of an automated XY stage, and preliminary results were reported by the team (Allen *et al.,* 2006). It was intended that Gary and myself should present the first test of the Autostage, as it was then called, at the International Palynological Congress in Bonn in 2008. Unfortunately Gary's funding failed, so I had to do the presentation alone (Flenley *et al.*, 2008). The organizers insisted that, as presenter, my name must go first, which was embarrassing in view of my tiny contribution. Gary completed a most successful M.Sc.

By now I was well past retirement age and had passed on my palynology teaching to Dr Katherine Holt. Kat embraced the automation project with enthusiasm and the Classifynder, as it is now called, is proceeding towards international acceptance. Bob Hodgson and his

team have put in tremendous effort and the machine is now in its third stage of development. As at June 2011, nine have already been sold and the present position is reviewed by Holt *et al.* (submitted 2011). Massey University is funding commercialization, and success seems assured.

REFERENCES

Allen, G.P., Hodgson, R.M., Marsland, S.R., Arnold, G., Flemmer, R.C., Flenley, J. & Fountain, D.W. (2006) Automatic recognition of light-microscope pollen images. In: *Proceedings of Image and Vision Computing New Zealand* (eds. P. Delmas, J. James & J. Morris), pp. 355-360

Fieller, N., Stillman, E. & Karaman, E. (1998) Automatic pollen analysis. Presentation to the International Environmetrics Society. Unpublished.

Flenley, J.R. (1968) The problem of pollen recognition. *Problems in Picture Interpretation* (ed. by M.B. Clowes & J.P. Penny), pp. 141-5. C S I R O, Canberra.

Flenley, J.R. (2003) Some prospects for lake sediment analysis in the 21st century. *Quaternary International*, **105**, 77-80.

Flenley, J.R., Hodgson, R., Allen, G., Marsland, S.R. & Arnold, G. (2008) Autopalynology: Autostage. Lecture presented at International Palynological Congress, Bonn. Unpublished.

Flenley, J.R., Li, P. & Empson, L.K. (1999) Identification of 13 pollen types by neural network analysis of texture data only. In: *Proceedings of Image and Vision Computing New Zealand 1999, University of Canterbury, Christchurch* (eds. D. Pairman & H. North), pp. 295-298. Landcare Research Ltd, Lincoln.

Haralick, R.M. (1979) Statistical and structural approaches to texture. *Proceedings of the Institute of Electrical and Electronic Engineers*, **67**, 786-804.

Haralick, R.M., Shanmagan, K. & Dinstein, I. (1973) Textural features for image classification. *IEEE Transactions Systems, Man & Cybernetics*, **SMC-3**, 610-621.

Hodgson, R.M., Holdaway, C.A., Zhang, Y., Fountain, D. & Flenley, J.R. (2005) Progress towards a system for automatic recognition of pollen using light microscope images. In: *Proceedings 4th International Symposium on Image and Signal Processing and Analysis, Zagreb, September 2005*, pp. 76-81

Holt, K.A., Allen, G.A., Hodgson, R.M., Marsland, S. & Flenley, J.R. (in review) Progress towards a lost-cost, automated trainable pollen location and classifier system for use in the palynology lab. Submitted to: *Review of Palaeobotany and Palynology*, May 2011.

Langford, M. (1988) Some applications of digital image processing for automation in palynology. Unpublished PhD thesis, University of Hull.

Langford, M., Taylor, G. & Flenley, J.R. (1986) The application of texture analysis for automated pollen identification. In: *Proc. Conference on Identification and Pattern Recognition*, pp. 729-739. Univ. Paul Sabatier, Toulouse, France, June 1986, Vol. 2.

Langford, M., Taylor, G.E. & Flenley, J.R. (1990) Computerized identification of pollen grains by texture analysis. *Review of Palaeobotany and Palynology*, **64**, 197-203.

Laws, K.I. (1990) Textured image segmentation. In: *Image Processing Institute, University of Southern California, Report No. 940*

Li, P. & Flenley, J.R. (1997) Classification and visualisation of pollen texture data using MLP neural networks. In: *Proceedings of the First Joint Australia and New Zealand Biennial Conference on Digital Image and Vision Computing - Techniques in Application, DICTA/IVCNZ '97* (ed. R.I. Chaplin), pp. 497-502, Massey University Albany Campus, Auckland.

Li, P. & Flenley, J.R. (1999) Pollen texture identification using neural networks. *Grana*, **38**, 59-64.

Li, P., Flenley, J.R. & Empson, L.K. (1998) Classification of 13 types of New Zealand pollen patterns using neural networks. In: *Proceedings International Conference on Image and Vision Computing* (eds. R. Klette, G. Gimel'farb & R. Kakarala), pp. 120-123, Auckland.

Li, P., Treloar, W.J., Flenley, J.R. & Empson, L. (2004) Towards automation of palynology. 2. The use of texture measures and neural network analysis for automated identification of optical images of pollen grains. *Journal of Quaternary Science*, **19**, 755-762.

Stillman, E.C. & Flenley, J.R. (1996) The needs and prospects for automation in palynology. *Quaternary Science Reviews*, **15**, 1-5.

Treloar, W.J. (1992) Digital image processing techniques and their application to the automation of palynology. Unpublished PhD Thesis. University of Hull, Hull.

Treloar, W.J. (1994a) Automation of pollen identification and counting using digital image processing techniques. *INQUA Commission for the Study of the Holocene, Working Group on Data-Handling Methods Newsletter*

Treloar, W.J. (1994b) Automation of palynology using image processing and pattern recognition techniques. Final Report to Massey University.

Treloar, W.J., Taylor, G.E. & Flenley, J.R. (2004) Towards automation of palynology. 1. Analysis of pollen shape and ornamentation using simple geometric measures, derived from scanning electron microscope images. *Journal of Quaternary Science*, **19**, 745-754.

Zhang, Y., Fountain, D.W., Hodgson, R.M., Flenley, J.R. & Gunetileke, S. (2003) Pollen classification via wavelet-based texture analysis. In: *Proceedings of the International Conference on Artificial Intelligence, June 23-26, Las Vegas, Nevada, USA, 2003*

Zhang, Y., Fountain, D.W., Hodgson, R.M., Flenley, J.R. & Gunetileke, S. (2004) Towards automation of palynology. 3: pollen pattern recognition using Gabor transforms and digital moments. *Journal of Quaternary Science*, **19**, 763-768.