Cytometry 5:244–249 (1984)

Image Processing for Mitoses in Sections of Breast Cancer: A Feasibility Study[1]

E.J. Kaman, A.W.M. Smeulders, P.W. Verbeek, I.T. Young, and J.P.A. Baak

Pattern Recognition Group, Department of Applied Physics, Delft University of Technology, Delft (E.J.K., A.W.M.S., P.W.V., I.T.Y.) and Department of Quantitative Pathology, S.S.D.Z., Delft (J.P.A.B.), The Netherlands

Received for publication April 6, 1983; accepted December 1, 1983

This paper describes an image analysis technique for the counting of nuclei in mitosis in tissue sections. Five experienced pathologists scored mitoses in photographs of preselected areas of tissue sections of the breast. Objects consistently labelled as mitotic cells by all five pathologists were considered "mitoses" in the analysis. In total, there were 45 mitotic nuclei, 68 possible mitotic nuclei and 1,172 nonmitotic nuclei.

The image analysis procedure was designed to give priority to a low false negative rate, i.e., misclassification of mitoses. The procedure consists of three steps:

1. Segmentation of the image.

2. Reduction of the number of nonmitotic nuclei by using feature values based on the brightness histogram of the objects.

3. Fully automatic classification of the remaining objects using contour features.

The objects remaining after the first two steps were visualized in a composite display for interactive evaluation: 10% of the mitotic nuclei were missed, and 85% of the nonmitotic nuclei were eliminated. The result of the fully automatic procedure described in this paper is rather disappointing and gave a loss of 37% of the mitoses while 5% of the nonmitotic nuclei remained.

Key terms: Image analysis, mitotic activity, tissue sections, automated analysis

In the sixties, when digital image processing entered the biomedical area, single cells were among the first topics to be studied. Due to the increased data acquisition and processing capabilities, entire cytological specimens are now ready to be analyzed. At the present time systems are operational for the screening of cervical smears (8,23,27) selection of well-spread metaphases in the field of cytogenetics (6), and automated analysis of blood smears (13,14).

In these screening systems, the use of single cell specimens in a monolayer suspension still is essential. Monolayers are obviously not present in a major part of diagnostic practice: tissue section analysis. As a result of this complication, the literature on the analysis of tissue sections is quite limited. The main issues have been the nuclear architecture of the epithelium of urinary bladder (4,25), intercorrelation of nuclear chromatin features in breast cancer (22), and the texture of the chromatin pattern of nuclei in endometrial hyperplasia and carcinoma (7).

Apart from the presence of multilayers of cells, several factors complicate the quantitative analysis of tissue sections.

1. Tangential cutting artefacts: The tangential cutting implies that image properties of single cells are not directly available, but can only be determined from a population of cells in a statistical sense using stereological methods (2,21). These artefacts may prohibit segmentation of single cells in the digital image.

2. Variability in image intensity: The uncontrolled variability of the tissue section thickness implies that attention should be paid to grey value normalisation (15).

This paper describes an image analysis technique for assessing mitotic activity in tissue sections. The number of mitotic nuclei (hereafter referred to as "mitoses") per unit area in tissue sections is an important feature in various grading systems of breast cancer (3) and also is important in the prognosis of breast cancer, ovarian cancer, and other tumors (1,16,19). The practical application in diagnostic pathology, however, is not very popular for two reasons. First, the accurate counting of mitoses is tedious and time-consuming work. Second, the reproducibility between different observers in a routine setting is not perfect (3). For example, the correlation between the first and second assessment by the

[1]Supported by grant 28-736 of the "Praeventiefonds."

Address reprint requests to Dr. J.P.A. Baak, Dept. of Pathology, S.S.D.Z., Reynier de Graefweg 7, 2625 AD DELFT, The Netherlands.

same observer of the total number of mitoses is approximately .90 (3).

MATERIALS AND METHODS

Tissue Processing and Image Selection

Tissue sections of breast cancer were used for this pilot study. Standard tissue fixation (4% buffered formalin, for approximately 24 h), dehydration in alcohol of increasing strength, and embedding in paraplast were performed. With this procedure, a majority of quantitative image features remain stable (24). Sections 4μm thick, were prepared and stained with hematoxylin-eosin, because of the necessity of visual inspection. From the image processing point of view a stochiometric stain is to be preferred. Photographs were made at $\times$1,000 film magnification $\times$100 objective, n.a. = 1.25 with green light, oil immersion, field diameter 200 μm using Kodak Panatomic-X (32 ASA) film. Standard grey wedges (kindly provided by Mr. M. Bisschoff, Carl Zeiss, Oberkochen, Germany) were included to confirm standard illumination conditions.

Description of the Material Used

An implicit complicating factor in the study of mitotic figures is the lack of absolute agreement among pathologists as to which objects constitute a mitosis. The inconsistency may be due to the inherent difficulty in defining the begin-and-end phase of a mitosis or be due to contradictory judgement on the same object. To reduce the inconsistency due to the definition of a mitosis, we consider as the endstage the moment when the nuclei of the two daughter cells are seen separately. As we could not rely on a data base selected by a single pathologist, images of cells were shown (Fig. 1) to a panel of five experienced, independent pathologists, who were asked to indicate mitoses, and objects which might, to their judgement, possibly be a mitosis (hereafter referred to as a "possible"). For ease of display these images were shown on photographs of preselected areas in the tissue. They were asked to be rather conservative in the labelling of an object as a mitosis. Thus, if they had any doubt, the object was labelled as a "possible." The results of this panel opinion are summarized in Table 1.

Objects consistently labelled by all five pathologists as either mitosis or cell not in division were considered, in the further analysis to be "mitosis" or "nonmitosis," respectively. All other objects were labelled as "possibles." In total, there were 45 mitoses, 68 possibles, and 1,172 nonmitoses including epithelial, inflammatory, and fibrocytic cell nuclei.

Due to the preselection, the images used were enriched with mitoses and tissue elements resembling mitoses. The percentage of mitoses in the photographs is about 3.5%, whereas in tissue sections of breast cancer this percentage is in the range of 0.1%–1.5% (J.P.A. Baak, unpublished results).

Instrumentation and Image Acquisition

The images were processed and evaluated in a large image processing and measurement system including the DIP (Delft Image Processor) pipelined image processor (10). The system is programmable through a command string interpreter encompassing some 250 operations on binary and grey-scale images and a variety of image measurements. Numerical data resulting from feature value measurements can be stored in a data base. A statistical package, ISPAHAN (12), is used for interactive evaluation and classification.

The black-and-white photographic negatives were scanned by a flying spot scanner and reduced to 5 bits (32 levels), with a 256$\times$256 pixel field of scan. The pixel-to-pixel distance ccorresponded to 0.25μm at the specimen level. A data base of 107 images of 64$\times$64μm^{2}was compiled. Half of the images were put aside to be used as an independent test set for the evaluation of the developed image processing procedure.

Image Processing

The image processing procedure was designed to give priority to a low false negative rate—that is, very few mitoses should be missed. At the same time the false positive rate should be reduced to a low percentage as well. The reduction of nonmitoses and final classification of the remaining objects was implemented by means of the following steps:

1. Normalisation: In spite of the standard conditions during the treatment of the film, one can expect small differences in the exposure level of the photographic negatives. Therefore, prior to the image processing, the "minimum" grey level in the image, defined by skipping the first 10 low grey values, was subtracted from the image in order to normalize all the images.

2. Global thresholding: In the first image processing step the image was segmented with a global grey level threshold (heuristially set at 8 grey levels over the minimum) to localize objects. Pixels with a grey level greater (brighter) than the threshold were background pixels; the others were defined as object pixels. The selected part of the image consisted of mitoses and other elements such as artefacts, pycnotic nuclei, inflammatory cells, and nuclear fragments.

3. Elimination of small objects: Half of these fragments were removed by application of a threefold erosion operation (21) which eliminated objects having an enscribed circle with a diameter smaller than 1.5μm.

4. Local thresholding: Subsequently, the objects were thresholded a second time, now locally, to approximate their contours more accurately. This was of importance for the calculation of the feature values derived from the contour later on. The image was again thresholded globally, now at a grey level slightly less than the grey level of the background. To be sure that an entire object is included in this local area, the area was dilated by three steps. The histogram of brightness within this new local area was then calculated and smoothed with a one-dimensional linear filter with coefficients 1,2,1.

If there were not two clearly distinct peaks present in this histogram, it was assumed that the local area did not contain a mitosis. By application of this "peak-to-peak criterion" this object was removed from further

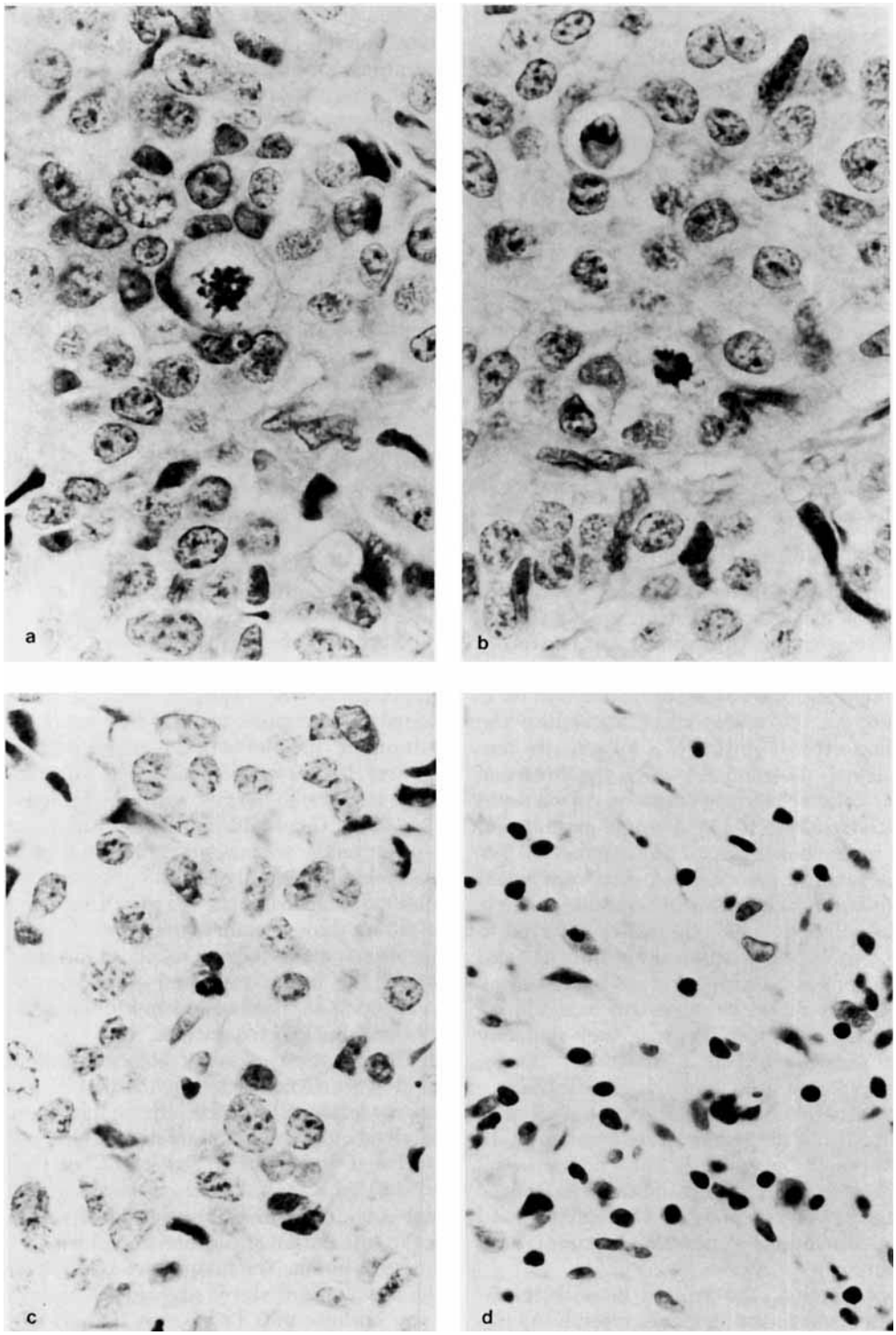

FIG. 1. Four hematoxylin-eosin-stained images of breast cancer with, in the centre, objects labelled by a panel of five pathologists. Image a) contains a mitosis consistently labelled by the five panel members. b) Four out of five label mitosis, one pathologist was in doubt. c) Two pathologists label mitosis, two were in doubt, one concluded no mitosis. d) Inflammatory cells. Note the large variety in overall staining intensity level. Magnification ×640.

Table 1
The a Priori Assignment by a Panel of Five Pathologists

	Absolute agreement	4 out of 5 agreement
Mitoses	45	55
Possibles	68	33
Nonmitoses	1,172	1,197

processing. The local histogram of dark objects, like mitoses, in a bright local environment yields a local histogram in which the two peaks are present. The local threshold is set at the minimum between these two peaks.

5. Local histogram features: The following three features were then calculated from the local histogram and used to further reduce the number of objects: the area (AREA), the average grey value (AGV), and a measure for the homogeneity of the object (HOM). The value of the homogeneity was calculated by dividing the number of pixels under the local threshold by the number of pixels under the second global threshold.

6. Reduction using local histogram features: Reduction of the number of objects was achieved by setting the following limits to the feature values: if an object was either too small (AREA $< 15\mu m^2$) or too large (AREA $> 90\mu m^2$) or was too bright (AGV > 8) or had too much internal contrast (HOM < 0.4) then the object was removed. In that case the object presumably was not a mitotic nucleus. The remaining objects were all gathered in a composite image and displayed for visual control and interactive evaluation (Fig. 2).

7. Contour features: For these objects two other features were calculated as well: average Roberts gradient (20) along the contour of the objects in a 5×5 window and the bending energy (5,26) which is a measure for the roughness of the contour.

8. Fully automatic classification: In the learning phase of the statistical classification only two classes of objects were considered: The class of mitoses and the class of nonmitoses. For fully automatic classification, the complete data-set, including the half previously put aside, was used. This complete data-set was subjected to the processing steps 1–7. The objects remaining after these steps were divided into two equal groups; first one group was used as a learning set and the other as a test set and then the two roles were reversed (11). A linear discriminant function (12) was used for classification. In the test phase the objects labelled by the panel as "possibles" were classified as well. The classification matrix thus contains three a priori classes and two a posteriori classes.

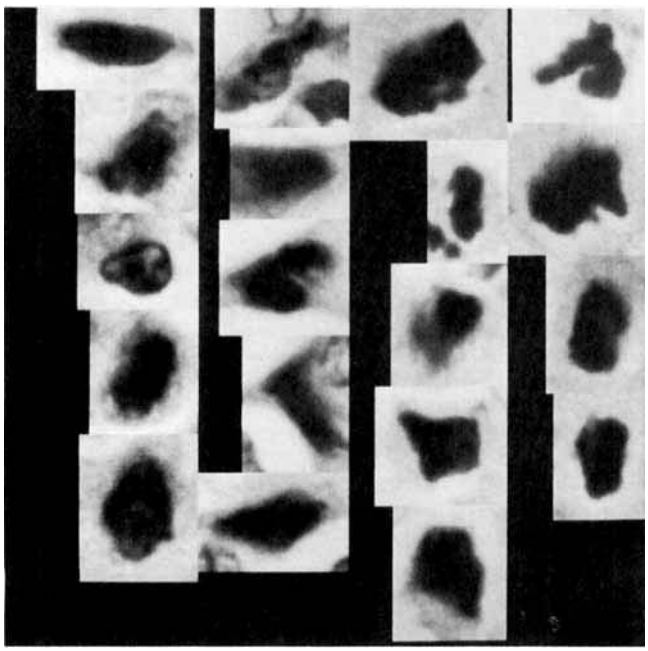

FIG. 2. Monitor display of composite image of automatically selected objects, suitable for interactive evaluation.

RESULTS

The result of the analysis after each image proessing step is shown in Table 2.

After segmentation with the global grey-value threshold the number of non-mitoses was reduced by a half, while the number of mitoses did not decrease. The reduction by three erosion steps again removed half of the remaining non-mitoses while again none of the mitoses was removed. Thus after these two steps 25% of the initial nonmitoses remained. Subsequently, an additional 2.5% of the nonmitoses were removed by application of the "peak-to-peak criterion."

At this point the objects were eliminated by subsequent application of a window on the feature values area, homogeneity, and the average grey value. This step eliminated 40% of the remaining nonmitoses and 10% of the mitoses, giving a total elimination of 85% of the nonmitoses and 10% of the mitoses.

With a Fisher linear discriminant function (9) the remaining objects were classified into two classes. This classification can be represented in a confusion matrix (Table 3).

When the results of the image processing and the classification were combined we obtained the results given in Table 4.

A Receiver Operating Characteristic (ROC) curve was generated with a kernel-based classifier (12); see Figure 3. This curve is given for the entire classification procedure. Presently, the best result of a fully automatic procedure has been a loss (false negative) of 37% of the mitoses while 5% of the nonmitoses remain (false positive).

DISCUSSION

The objective of this pilot study is to investigate the possibility of scoring semiautomatically the mitotic activity in tissue sections. The basic approach is to detect and count mitoses in a diagnostically interesting area of the section.

The uncertainty in the a priori assignment of mitoses was overcome, in this study, by referring preselected

Table 2
The Percentage of Reduction of Mitoses and Nonmitoses After Each Processing and Classification Steps (possibles are not taken into consideration for the present table)

	Elimination (%)		Cumulative remaining (%)	
	Mitoses	Nonmitoses	Mitoses	Nonmitoses
Image Analysis				
Preselection				
Global threshold	0	50	100	50
Three erosion steps	0	50	100	25
Minimum peak distance	0	2.5	100	24
Further analysis				
Classification with features (area, homogeneity, grey level)	10	40	90	15
Further classification (see Table 3)	30	60	63	5

Table 3
Confusion Matrix for the Objects Remaining After Processing and Classification With the Features Area, Average Grey Level, and Homogeneity

	Mitoses	Nonmitoses	Total
Absolute agreement			
A priori			
Mitoses	28	12	40
Possibles	31	22	53
Nonmitoses	43	77	120
Total	102	111	213
4 out of 5 agree			
A priori			
Mitoses	36	14	50
Possibles	15	9	24
Nonmitoses	51	88	139
Total	102	111	213

Table 4
Confusion Matrix for all the Objects With Absolute a Priori Agreement

	Automatic classification		
	Mitoses	Nonmitoses	Total
Mitoses	28	17	45
Possibles	31	37	68
Nonmitoses	43	1,129	1,172
Total	102	1,183	1,285

areas of tissue sections to a panel of five pathologists. In this way, a reliable data set of mitoses and nonmitoses was composed.

There are several causes for the large number of "possibles" resulting from the panel assignments. First, the pathologists were asked to be conservative in their labelling of a mitosis. Second, because of the use of black-and-white photo prints, neither depth information nor color information was available. This is, of course, not a normal diagnostic situation. The total number of possibles (68) decreases to 33 if objects for which four out of five pathologists agree are removed from the set of possibles and either added to the mitoses or nonmitoses. The images used in this pilot study were selected for the presence of cells in division and objects which show strong resemblance to mitoses to investigate the true selective power of the developed procedure. It is likely that the use of such a data-base degraded the classification result in this study.

Although the data-set may seem relatively small the following preliminary conclusion may be drawn. In the first stage of processing, three quarters of the nonmitoses were removed and all mitoses were retained. This is an encouraging result for a preselection system. The results change to a removal of 85% of the nonmitoses losing 10% of the mitoses by the application of simple-to-compute object features.

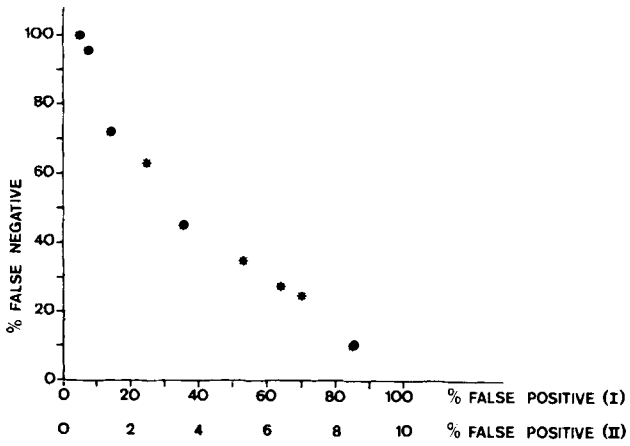

FIG. 3. ROC-curve. On the vertical axis the perentage of the false negative classification (loss of mitoses) and on the horizontal axis the total percentage of false positive classifications (the number of nonmitoses classified in the whole recognition procedure as mitoses) are presented for the number of nonmitoses after the reduction steps (I) and for the total number of nonmitoses (II).

The display of the remaining objects in a composite image makes the quantitative assessment of the mitotic activity feasible. Now, as an alternative for automatic classification the remaining objects could be eliminated interactively giving a considerable work reduction.

Analysis of the false negative and false positive classifications indicates the following two main causes for misclassification. First, the local contrast of some of the mitoses was low with respect to the surrounding background value. This seems to occur particularly in late stages of division. Second, the segmentation procedure is inadequate in some cases. In these cases a local threshold is not sufficient to retain a contour as perceived visually. The contour is of prime importance for the classification based upon shape features. So far, the results obtained with the shape analysis are disappointing. Application of an adaptive contour algorithm e.g., (18) or (17), might improve the result. In addition, gain might be expected from the use of a higher sampling resolution as the characteristic little folds in the shape of mitoses are obscured by the size of the present digitization grid. The presence of these folds may then be established by the bending energy feature (5,26).

In summary, the study presented here indicates the feasibility of an interactive image processing system for the scoring of mitotic activity. In a subsequent study more sophisticated processing techniques, such as determination of the integrated optical density and texture measures, and specimen decision strategies are planned. These techniques will be implemented in a real-time system working directly from microscopic slides and thus making the photographic step unnecessary.

ACKNOWLEDGMENTS

The authors wish to thank Mrs. D.I.M. de Jong for her help in the preparation of this report.

LITERATURE CITED

1. Agrafojo Blanco A, Gibbs ACC, Langley FA: Histological discrimination of malignancy in mucinous ovarian tumours. Histopathology I:431–443, 1977.
2. Aherne BA, Dunnill MS: Morphometry. E. Arnold Ltd., London, 1982.
3. Baak JPA, Kurver PHJ, Snoo-Nieuwlaat AJE de, Graaf S de, Makkink B, Boon ME: Prognostic indicators in breast cancer—morphometric methods. Histopathology 6:327–339, 1982.
4. Bartels PH, Bibbo M, Wied GL: Modeling of histologic images by computer. Acta Cytol 20:62–67, 1976.
5. Bowie JE, Young IT: An analysis technique for biological shape. II. Acta Cytol 21:739–746, 1977.
6. Castleman KR, Melnyk JH: An automated system for chromosome analysis. Final report, Internal document no. 5040-30. Jet Propulsion Laboratory, 1979.
7. Diegenbach PC, Baak JPA: Quantitative nuclear image analysis: Differentiation between normal, hyperplastic and malignant appearing uterine glands in a paraffin section. IV. The use of Markov chain texture features in discriminant analysis. Eur J Obstet Gynaeol Reprod Biol 8:157, 1978.
8. Driel AMJ van, Pleom JS : The use of LEYTAS in analytical and quantitative cytology IEEE Trans Biomed Eng 29–92-100, 1982.
9. Duda RO, Hart PE: Pattern Classification and Scene Analysis. John Wiley and Sons, New York 1973.
10. Duin RPW, Gerritse FA, Groen FCA, Verbeek PW, Verhagen CJDM: The Delft Image Processing System, design and use. Proceedings 5th International Conference on Pattern Recognition, Miami Beach, Florida, 1980, pp 768–773.
11. Fukunaga K: Introduction to Statistical Pattern Recognition. Academic Press Inc., New York, 1972.
12. Gelsema ES: ISPAHAN; an interactive system for pattern analysis: structure and capabilities. In: Pattern Recognition in Practice, Gelsema ES, Kanal LN, (eds). North-Holland Publishing Company, Amsterdam, 1980, pp 481–491.
13. Graham MD, Norgren PE: The Diff3 analyzer. A parallel/serial Golay image processer. In: Real-Time Medical Image Processing. Onoe M, Preston K, Rosefeld A (eds). Plenum Press, New York, 1981, pp 163–182.
14. Green JE: The Abbot laboratories, ADC-500. In: Real-Time Medical Image Processing. Onoe M, Preston K, Rosefeld A (eds) Plenum Press, New York, 1981, pp 119–148.
15. Haralick RM, Shammugan K, Dinstein I: Textural features for image classification, IEEE Trans SMC 3:610–621, 1973.
16. Heul RO van der: Het periostale ossificerende fibrosarcoom en de gradering van osteosarcomen. 1962.
17. Lester JM, Williams HA, Weintraut BA, Brenner JF: Two graph searching techniques for boundary finding in white cell blood cell images. Comp Biol Med 8:293–308, 1978.
18. Meyer F: Empiricism or idealism . . . In: Pattern Recognition in Practice. Gelsema ES, Kanal LN (eds) North Holland Publishing Company, Amsterdam, 1980, pp 21–33.
19. Norris HJ: Mitosis counting (editorial). Human Pathol 7:483–484, 1976.
20. Rosenfeld A, Kak AC: Digital Picture Proccessing. Academic Press, New York, 1982, Vol 1 and 2, 2nd ed.
21. Serra J: One, two, three, ... Infinity. In: Quantitative Analysis of Structures, Chermant JL (ed). Riederer, Stuttgart, 1978, pp 9–24.
22. Stenkvist B. Bengtsson E, Eriksson O, Jarkrans T, Nordin B, Westman-Naeser S: Correlation between cytometric features and mitotic frequence in human breast carcinoma. Cytometry 1:287–291, 1981.
23. Tanaka N, Ideda H, Ueno T, Mukawa A, Kamitsuma K: Field test and experimental use of CYBEST model 2 for practical gynecologic mass screening. Anal Quant Cytol 1:122–126, 1979.
24. Thunissen EBJM, Baak JPA, Diegenbach PC. Fixation induced variations in quantitative nuclear image features in sections. Acta Histochem 68:218–226, 1981.
25. Young IT, Vanderlaan M, Kromhout L, Jensen R, Grover A, King E: Morphologic changes in rat urothelial cells during carcinogenesis II. Image cytometry. Cytometry 5:5, 1984 (in press).
26. Young IT, Walker JE, Bowie JE: An analysis technique for biological shape. Information Control 25:357–370, 1974.
27. Zahniser D, Oud PS, Raaijmakers MCT, Vooys GP, Walle RT van der: Field test results using the bioPREP cervical smear prescreening system. Cytometry 1:200–203, 1980.