ACTA STEREOL 1990; 9/1: 53-60
PROC ECS 5 FREIBURG I. BR. 1989
ORIGINAL SCIENTIFIC PAPER

COMPARISON AND DISCUSSION OF SOME TEXTURAL PARAMETERS

Catherine Souchier, Anne Marie Manel, Pascale Felman,
Odile Gentilhomme and Paul Andre Bryon

Laboratoire de cytologie analytique, 8 avenue Rockefeller,
69373 LYON Cedex 08. FRANCE, and laboratoires d'hématologie,
Hopital Debrousse (AM) 69322 LYON cedex 02, Hopital E.Herriot (OG)
69437 Lyon cedex 08, Hopital Lyon Sud (PF,PAB) 69310 Piere Bénite.
France

ABSTRACT

Texture being one of the main nuclear characteristics, there is a constant increase of methods for its quantification. As a better understanding of some textural parameters is needed, this study examined three methodological groups : analysis of grey level distributions (coefficients of variation, of skewness and of kurtosis), of the darkest or lightest points and of neighbored points like the cooccurrence matrix method. The discussion was done after an experimental study on three groups of real images (non-Hodgkin's lymphoma nuclei), representing three different textures (clumped, homogenous, fine chromatin). Of all the parameters analyzed, the coefficient of variation of nuclear brightness proved to be the best model of a textural variable.

Keywords : Chromatin pattern, cooccurrence matrix, image analysis, quantitative microscopy, texture measurement.

INTRODUCTION

In biology, image analysis is often used to study nuclei and describe their size, their shape, their DNA content and their chromatin texture.

Numerous methods to analyze texture have been proposed in literature, and in some methodological groups, the possible options are many. Global distribution of nuclear brightness is studied and summarized by the coefficients of variation, of skewness, and of kurtosis (Bacus et al, 1987, Jagoe et al 1984). The darkest/whitest areas are defined and analyzed by the percentage of points labeled black or white, by the granularity estimation through a sieving technique such as opening, and by the margination measurement (Lockart et al, 1984, Young et al, 1986). The variations of grey level values in the neighborhood of a picture point are studied and cooccurrence matrix (Bacus and Grace, 1987, Haralick et al, 1973, Pressman, 1976), as well as run lengths matrix (Galloway, 1985) defined. The relative frequencies p_{ij} of occurrence of one pixel in the grey level class i, while its neighbor is in the grey level class j, are computed in cooccurrence matrix, and the frequencies of occurrence of collinear connected set of pixels in a same grey level class i and of length j are computed in run lengths. Several texture parameters are determined from the matrix. Filtering approaches such as top hat transformations (Meyer, 1980), gradient, Laplace and median filters (Jagoe et al, 1984, Kriete et al, 1987) are also proposed.

The main questions asked by the biologist are : Which is the best choice for nuclear study ? What is the biological meaning of all the parameters ? What can be done with the image analysis system I am using ? This study aims at a better understanding of some textural parameters and concentrates on three methodogical groups; the discussion followed an experimental study on three groups of real images (non-Hodgkin's lymphoma nuclei) representing three different textures (clumped, homogenous, fine chromatin).

MATERIAL AND METHODS

Biological material

Nuclei of non-Hodgkin's malignant lymphoma (NHL) coming from specimens classified by two pathologists according to the 'Working Formulation for Clinical Usage' (The NHL classification project, 1982) were analyzed. Two hundred nuclei were selected from lymphocytic NHL specimens, two hundred from lymphoblastic NHL, and two hundred from immunoblastic NHL. These nuclei were chosen as being representative of three different kinds of texture. In lymphocytic NHL, the chromatin is condensed in clumps; in lymphoblastic NHL, it is fine and the nuclei appear to be more homogeneous; and in immunoblastic NHL, the chromatin is fine and evenly distributed with small aggregates (Fig. 1). Analysis was performed on lymph-node imprints, colored with a panoptic stain (Wright), as is usual for clinical classifications.

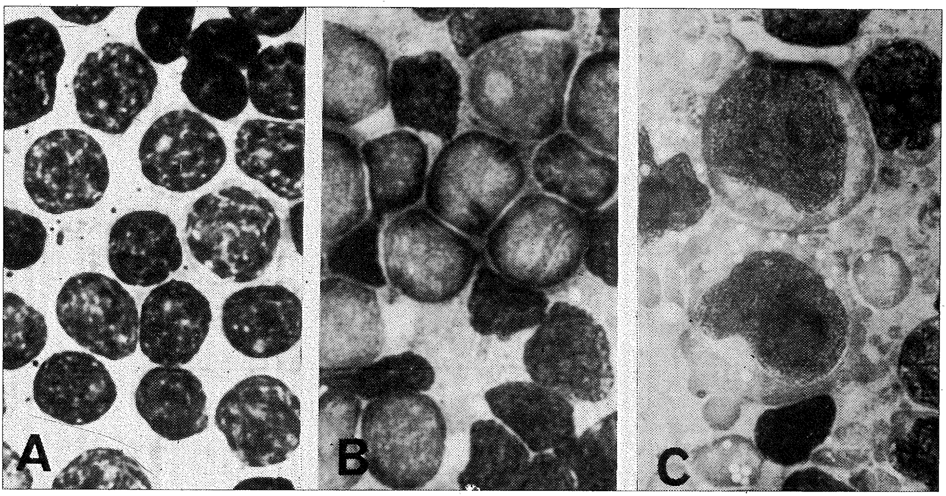

Figure 1. Nuclear texures of non-Hodgkin's lymphomas
A. Lymphocytic NHL, B. Lymphoblastic NHL, C. Immunoblastic NHL.

Image analysis method

The study was performed on an image analysis system (Quantimet 900, Cambridge Instruments) (Jenkinson, 1985), connected with a microscope (Zeiss Universal). The observation of nuclei was done with a 100X planachromat objective (NA = 1.25).

The images of nuclei were interactively selected, stored on the computer storage device, and then measured. The routines were developped with the image analysis ready-to-use software (QUIPS) and in extended

Pascal, which allows complete interfacing with Quips.

The study concentrated on three different methods.

- 1. The nuclear distribution of grey levels was described using the coefficient of variation (COVAR= standard deviation/mean grey value), and the Fisher coefficients of skewness and kurtosis.

- 2. Two methods were used for each nucleus to label the blackest and whitest points. In the first one, the cutpoints were defined around the mean grey value. The points above the grey value (1.1 * mean value) were considered as white and the points under the grey value (0.9 * mean value) as black. In the second method, the 10% (or 20%) blackest points were labeled black, and similarly the 10% (or 20%) whitest points were labeled white. They were analyzed separately or together in a same union image. The quantities of black/white clumps were evaluated on the basis of the percentage of points that resisted opening, and the percentages of black/white points were also evaluated in the first method.

- 3. The cooccurrence matrices (Haralick, 1973) were computed according to different procedures. In this method, two points at a distance d are considered as neighbors, and the matrix defines the relative numbers (p_{ij}) of bi-points being, one inside a grey level class i and the other inside a grey level class j. Eight or sixteen grey level classes were considered. The grey level classes were either defined for each nucleus or globally for all the nuclei. The classes were computed on the basis of equal numbers of grey levels, or of equal numbers of pixels in each class (equalization procedure). The variables (angular second moment, contrast, entropy, inverse different moment and correlation) were computed from the relative frequencies, as defined by Haralick (1973) (Table 1).

Table 1. Textural features of the cooccurrence matrix

Angular Second Moment	$\mathrm{ASM} = \sum_i \sum_j p(i,j)^2$
Contrast	$\mathrm{CON} = \sum_i \sum_j (i-j)^2 \,.\, p(i,j)$
Entropy	$\mathrm{ENT} = -\sum_i \sum_j p(i,j) \,.\, \log(p(i,j))$
Inverse Difference Moment	$\mathrm{IDM} = \sum_i \sum_j \frac{1}{1+(i-j)^2} \; p(i,j)$
Correlation	$\mathrm{COR} = \frac{\sum_i \sum_j p(i,j).(i-\mu).(j-\mu)}{s^2}$

Statistical analysis

Data were transferred to a 286 PC compatible computer and analyzed using BMDP statistical software (Dixon et al, 1988). The efficiency of a group of variables was measured in terms of its discriminating power, and the relations between variables by the coefficients of correlation.

RESULTS

Analysis of brightness distribution

An overall discrimination of 66% was obtained with the coefficient of variation of nuclear brightness alone. The results were improved by only 3% when the coefficients of skewness and of kurtosis were added. Textures of lymphoblastic NHL were especially easy to identify, 87 percent of nuclei were correctly classified.

Analysis of black/white points

The results are given in Table 2.

Table 2. Percent of correct classification obtained in Black/white analysis

BLACK / WHITE POINTS DEFINED :	BLACK / WHITE POINTS ANALYSED				
	TOGETHER	SEPARATELY			
		Black only	White only	Black / White	
					+COVAR
Around the mean value	63 %	58 %	69 %	67 %	66 %
As 10% Blackest / Whitest points	41 %	30 %	45 %	44 %	69 %
As 20% Blackest / Whitest points	38 %	38 %	43 %	47 %	69 %

Better results (+22%) were obtained when the black/white points were defined around the mean nuclear grey level and not as the 10% or 20% blackest/whitest points.

The percentage of correct classification was increased by 5% when the black and black points were analyzed separately, and not in a same union image. The correlation between black/white clumps (r=.5) is lower than the correlation between black/white points (r=.9).

Moreover, better results were obtained when dealing with white points only than with black points only. In this optimum option, the classification of the lymphoblastic NHL was higher (88%) than the classification of the two other subtypes (67%, 53%), as was the case in the analysis of brightness distribution.

Analysis of cooccurrence matrix

Very high correlations (>0.9) were either obtained between the angular second moment variable and the entropy, or between the contrast variable and the inverse different moment. When, instead of only two variables (angular second moment and contrast), four variables were added to the correlation variable, the percentage of correct classification was only improved slightly (3%).

The main results are given in Table 3.

Table 3. Percentage of correct classification obtained with cooccurrence matrix.

The results were obtained with the angular second moment, the contrast and the correlation variables.

COOCURRENCE MATRIX EQUAL NUMBER OF GREYS / CLASS	16 GREY CLASSES Distances between neighbors				8 CLASSES Distances	
	1	4	8	16	1	16
Min=0 , Max=255 for all nuclei	63 %	70 %	61 %	71 %	51 %	62 %
Min=63, Max=238 for all nuclei	62 %	71 %	61 %	73 %	61 %	66 %
Min , Max of each nucleus	56 %	56 %	62 %	65 %	53 %	65 %
EQUAL NUMBER OF POINTS / CLASS	55 %	62 %	58 %	71 %	62 %	73 %
Min=0 , Max=255 for all nuclei	55 %	62 %	58 %	71 %	62 %	73 %
Min , Max of each nucleus	54 %	59 %	56 %	70 %	46 %	64 %

Better results (+ 4%) were obtained when the grey level classes were defined with an equal number of grey levels per class and for all the nuclei than for each nucleus or with an equal number of points in each class. In the best cases, the texture of the lymphoblastic NHL was recognized better than the two other subtypes. On the contrary, the texture of the lymphocytic NHL was easier to recognize than the two other NHL, when the grey level classes were defined for each nucleus, mainly at small distances (d=1,4) and with an equal number of grey levels per class.

No significant differences were observed between either the results obtained with eight and sixteen classes of grey levels, or between the results defined inside the real extreme values (63-248) and those by default (0-255). The interest of a precise definition of extreme grey values was nevertheless shown for a reduced number of grey classes (=8).

Mean absolute and relative values were dependent on the distance defined between the two points considered as neighbors. The increase in contrast value is more important for a clumped chromatin (X10) than for a fine or homogenous chromatin (X6).

Analysis with the coefficient of variation of nuclear brightness

The coefficient of variation of nuclear brigthness (COVAR) was added to the variables obtained from the analysis of black/white points or cooccurrence matrix. There were significant improvements in almost all cases (Tables 2,4), and the differences between results obtained previously with the different types of cooccurrence matrix lose significance. The smallest increases were effectively observed in the cases which had previously shown the best results: cooccurrence matrix defined for all the nuclei with the same number of grey levels in each class, analysis of black/white points defined in relation to each mean nuclear grey value. In these cases, the variables angular second moment, the contrast and the total quantity of black/white points are highly correlated with the coefficient of variation of nuclear brightness (r > .8), and cover the same information. In all the variants except one (case with 82% total correct classification), the percentage of correct classification is about 10% better for the lymphoblastic NHL than for the two others.

Table 4. Percentage of correct classification obtained with coefficient of variation of nuclear brightness and cooccurrence matrix.
The variables of cooccurrence matrix are contrast, correlation and angular second moment.

COEFFICIENT OF VARIATION AND COOCCURRENCE MATRIX WITH :	DISTANCES BETWEEN NEIGHBORS			
EQUAL NUMBER OF GREY LEVELS / CLASS	1	4	8	16
Min=0 , Max=255 for all nuclei	70 %	73 %	71 %	74 %
Min , Max of each nucleus	72 %	82 %	75 %	78 %
EQUAL NUMBER OF POINTS / CLASS				
Min=0 , Max=255 for all nuclei	69 %	76 %	74 %	79 %
Min , Max of each nucleus	65 %	70 %	70 %	80 %

Optimisation of the percentage of correct classification

Several methods and, for one methodological group, several options, were brought together for a better classification, as was done for the coefficient of variation of nuclear brightness, and the most significant results were noted. The combination of different distances in cooccurrence matrix gave the best results and an overall discrimination of 85% was obtained with the following variables : coefficient of variation of nuclear brigthness, contrast and correlation obtained at smaller and greater distances (4 and 16 points).

DISCUSSION

Three different methodological groups were compared. All of them lead to nearly 65% of correct classification for at least one condition, and in nearly all the cases, the most homogeneous texture -the texture of lymphoblastic Non-Hodkin's lymphomas- is the texture the most easily recognized. However, the analysis of nuclear brightness distribution can easily be considered as the best, at least when textures with different levels of chromatin heterogeneity have to be discriminated. The results were obtained with only one variable : the coefficient of variation, which is moreover well defined without any need for grey level compression and nuclear partition, and easy to interpret because clearly related to nuclear homogeneity. The other methods do not have these advantages.

In the analysis of black/white points, the correct cutpoints or percentages for labelling the points as black/grey/white must be defined, as well as the correct opening for clump selection. In this sense, the optimal solution would depend on the types of nuclei to be analyzed and on the staining used, and would rarely be obtained. Moreover, with this method only, it is better to define the points in relation to each mean nuclear grey value, the percentage of black/white points being related, like the coefficient of variation, to the nuclear homogeneity, although this solution seems more difficult to standardize than when the percentage of black/white points has been fixed. Nevertheless, it is the easiest and quickest option. It should also be noticed that, on our data set, better results were obtained with different data for the black and white points, the data for white points being better than for black. These observations should lead to a slight modification of the method proposed by Young et al (1986) which makes a global evaluation of the black/white points. Similarly, our observations contradict the propositions of other authors who analyze only the black phase (Komitowski,1985, Lockart,1984). Other data sets should at least be separately analyzed. In our opinion, the main advantage of this methodological group is that it makes it possible to obtain textural parameters with only mean brightness and area measurements associated with image morphological transformations, and this can be done with most of the commercial image analysis systems.

The cooccurrence matrices do not lead to well-defined and easy to interpret variables with the chromatin pattern, as was the case with the analysis of the black/white method. For instance, when the cooccurrence matrix is defined for all nuclei between the extreme default values, and with an equal number of grey levels in each class, the angular second moment variable -index of matrix heterogeneity- covers both the nuclear heterogeneity and the dependence intensity between neighbors points. Part of the information is thus close to the coefficient of variation of nuclear brightness, and this explains why, when this last parameter is added, the total percent of correct classification is not really improved. Grey level equalization of each nucleus is of interest because it eliminates the heterogeneity factor and leads to a clearer variable. The meaning of a variable and the results of the discriminant analysis thus depends on the matrix definition. Another example of the difficulty of a precise interpretation can be illustrated by the results obtained with the cooccurrence matrix defined between the extreme values of each nucleus. In this case, the grey levels are transposed on the same grey scale, and the zones of transition of homogeneous nuclei such as the lymphoblastic nuclei are amplified more than those of heterogeneous nuclei such as 'lymphocytic. The visual perception is thus transformed, and the highest discrimination of the lymphocytic NHL was more a result of observation than could have been predicted by the method.

Our results give some pointers as to the choice of the best cooccurrence matrix. The decision would depend on whether or not it was possible to measure the coefficient of variation of nuclear brightness at a low cost. Without this last variable, the best solution would be to work with a small distance and sixteen classes defined between the extreme default values. As concerns the coefficient of variation of nuclear brightness, a grey level histogram equalization of each nucleus as recommended by Haralick (1973) would be preferable for a better interpretation of results. The very high correlation between some variables also suggests that the number of variables coming from a same cooccurrence matrix should be as low as possible and that only the most significant should be kept. The combination of two distances is of more interest. The greatest distance should be, as it was in our situation, similar to the clump size, and as pointed out by Pressman (1976), not too great.

The coefficient of variation of nuclear brightness alone is by no means sufficient for recognition of all the nuclear textures. It should be associated to other variables, and those coming from the cooccurrence matrix represent one of the solutions. Nevertheless, we only obtained an 85% correct classification on our three relatively well-defined groups of nuclear textures, while a more significant rate (>90%) was expected, in spite of the nuclear variability. The grey levels compression computed in the method of cooccurrence matrix as well as in the black/white points analysis are one of the possible explanations. This transformation is sensitive to slight grey levels variations inside nuclei and boundary effects which are not representative of nuclear texture and are observed in some cases of nuclei, more particularly on imprints. A solution without any grey level compression would probably give more significant results, and whould also be better defined. The development of grey processors could lead to new developments of textural variables, without any nuclear segmentation or grey levels compression, such as the coefficient of variation of nuclear brightness. Moreover these new variables should each, in our opinion, reflect a characteristic of the chromation texture, such as homogeneity, condensation, margination, if they are to be useful for nuclear texture.

ACKNOWLEDGMENTS

The author wishes to thank the "Association pour la Recherche sur le Cancer" for a research grant.

REFERENCES

Bacus JW, Grace LJ. Optical microscope system for standardized cell measurements and analyses. Applied Optics 1987; 26:3280-3293.

Dixon WJ, Brown MB, Engelman L, Hill MA, Jenrich RI. BMDP statistical software manual. University of California Press, 1988.

Galloway MM. Texture analysis using gray level run lengths. Computer Graphics and Image Processing 1985; 4:172-179.

Haralick RM, Shanmugam K, Its'hak Dinstein. Textural features for image classification. IEEE Transactions on Systems, Man, and Cybernetics 1973; smc-3:6:610-621.

Jagoe R, Sowter C, Slavin G. Shape and texture analysis of liver cell nuclei in hepatomas by computer aided microscopy. J Clin Pathol 1984; 37:755-762.

Jenkinson G. Image analysis takes off. Nature 1985; 316:87-88.

Komitowski D, Zinser G. Quantitative description of chromatin structure during neoplasia by the method of image processing. Analyt Quant Cytol Histol 1985; 7:178-182.

Kriete A, Schäffer R, Harms H, Aus HM. On-line transmission electron microscopic image analysis of chromatin texture for differentiation of thyroid gland tumours. Analyt Quant Cytol 1987; 9:268-272.

Lockart RZ, Pezella KM, Kelley MM, Toy ST. Features independent of stain intensity for evaluating Feulgen-stained cells. Analyt Quant Cytol 1984; 6:105-111.

Meyer F. Quantitative analysis of the chromatin of lymphocytes; an essay on comparative structuralism. Blood Cells 1980; 6:159-172.

Pressman N.J. Markovian analysis of cervical cell images. J Histochem Cytochem 1976; 24:138-144.

The non-Hogkin's lymphoma pathologic classification project. National Cancer Institute sponsored study of classifications of non-Hodgkin's lymphomas, summary and description of a working formulation for clinical usage. Cancer 1982; 49:2112-2135.

Young IT, Verbeek PW, Mayall BH. Characterization of chromatin distribution in cell nuclei. Cytometry 1986; 7:467-474.