

Classification

Physics Abstracts

07.05.Pj — 42.30.Sy

Global and Non Parametric Classification Methods Using Mathematical Morphology: Application to DNA Ploidy Measurement of Archival Solid Tumors

Christophe Boudry ^(1,2,3), Michel Coster ⁽¹⁾, Paulette Herlin ⁽²⁾, Brigitte Sola ⁽³⁾ and Jean-Louis Chermant ⁽¹⁾

⁽¹⁾ Laboratoire d'Études et de Recherches sur les Matériaux, URA CNRS 1317, ISMRA, 14050 Caen Cedex, France

⁽²⁾ Service d'Anatomie Pathologique, Centre F. Baclesse, 14021 Caen Cedex, France

⁽³⁾ CYCERON, Université de Caen, URA CNRS 1829, 14074 Caen Cedex, France

Résumé. — La majorité des tris d'objets biologiques effectués en analyse d'images fait appel à la mesure de paramètres (objet par objet) suivie d'une analyse des données. La morphologie mathématique offre des outils susceptibles de traiter les images globalement. Jusqu'à présent, cette opportunité a été sous employée pour les tris d'objets. Le but du présent article est de proposer des méthodes globales de tri, appliquées à l'élimination des agrégats et des petits débris, pour la mesure de l'ADN ploïdie des tumeurs solides. Les résultats illustrent la puissance et l'efficacité des traitements globaux par morphologie mathématique utilisés comme méthode de tri.

Abstract. — Most of sorting methods applied to images in biology uses parametric measurements (object per object) followed by data analysis. Mathematical morphology offers tools able to treat images globally. Up to now, this opportunity has been under employed. The goal of the present paper is to propose global sorting methods, applied to the automatic elimination of aggregates and small debris, for DNA ploidy quantification. Results obtained illustrate the powerfulness and efficiency of global mathematical morphology treatments as sorting method.

1. Introduction

Most often, image analysis leads to a quantitative evaluation by way of parameters which take into account physical phenomena. When belonging to an homogeneous class, objects are simply segmented and then measured. In biology, the complexity of images often requires sorting of objects before measurement. Sorting can be done thanks to different criteria: texture, grey level, color, size, shape... Depending on the criterion studied, sorting will be obtained using individual or global methods.

Individual methods treat object per object. One or several parameters are measured for each object, then a data analysis allows object sorting [1, 2]. On the contrary, global methods are based on the study of the whole image and can be separated in two categories. The first category does not need image transformations: this can be easily illustrated by an example of grey level image threshold which allows the global sorting of bright objects *versus* dark objects. The second category is based on global morphological transformations: the simplest example refers to granulometry based on opening/reconstruction which allows thickness dependent sorting of objects. These global methods do not consider objects individually, but the whole image globally, whatever the number of objects per image.

Individual methods are the most popular methods [3, 4] and global methods are scarcely used as classification tool [5-7]. Here a global sorting method which uses complex mathematical morphology (MM) operators based on shape, size and grey level analysis is proposed.

If flow cytometry allows the DNA (DesoxyriboNucleic Acid) ploidy measurement of large populations of nuclei in a short delay, it is a well known blind technique. Image analysis, an alternative tool for DNA ploidy quantification, allows sorting of elements under visual control, which is reputed time consuming. To compete with flow cytometry, automatic sorting procedures dedicated to image DNA ploidy quantification must be developed. For this task, methods here proposed are devoted to the automatic classification of debris and aggregates *versus* undamaged nuclei, for image DNA ploidy measurement of solid tumors. The main problem encountered in DNA ploidy quantification is the purity of samples to be analyzed. The preparation of samples requires a dissociation of tissue, which generates undamaged nuclei to be measured, but also debris and aggregates. These unwanted elements represent up to 60% of the total events and can be classified in four categories: small debris, sliced nuclei, damaged nuclei and aggregates. Small debris and aggregates introduce a great bias in image DNA ploidy measurement and consequently must be automatically removed [8].

2. Materials and Methods

The study was done on an archival tumor corresponding to a case of astrocytoma grade 2 and prepared according to Van-Driel Kulker *et al.* [9] as previously described [10]. Image acquisition was done at a 512×512 pixels resolution in 8 bits (1pixel = $0.11 \mu\text{m}^2$) thanks to a BH2 Olympus microscope, a moving stage, a Matrox PIP 1024 frame grabber and a Sony CCD camera. Fifty images representing 1079 elements were studied. Segmented events were classified in six categories of normal and abnormal nuclei, and in 4 categories of debris as previously described [8]. For each category, sensivity and false positive rates of global automatic method were assessed by reference to manual sorting. Interactive sorting and DNA measurements were done using DRACCAR[®] software (Gestinfo Corp., Caen and ADCIS Corp., Caen, France) running under UNIX operating system. Visilog4[®] software (Noesis, Orsay, France) also running under UNIX provided the mathematical morphology routines. Access (Microsoft[®]) data base management software was used to compare labelling of events, in order to assess sensitivities and false positive rates for each studied category.

3. Elimination of Aggregates

Aggregates are due to overlapping of nuclei during the sedimentation procedure for sample preparation. Aggregates are always characterized and interactively identified by their concavities (Fig. 1a) whereas nuclei are convex (Fig. 1b) or slightly concave (Figs. 1c, d). The ultimate eroded set can be used to characterize concavity, keeping in mind that convex shapes have one

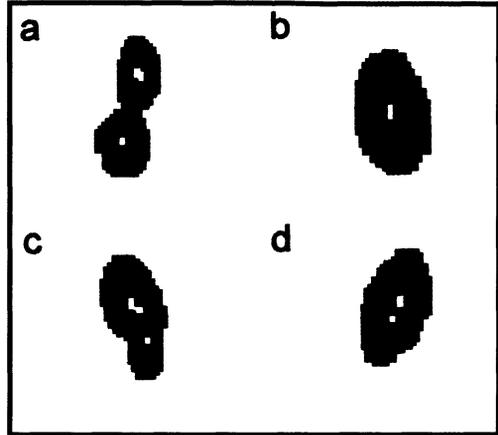
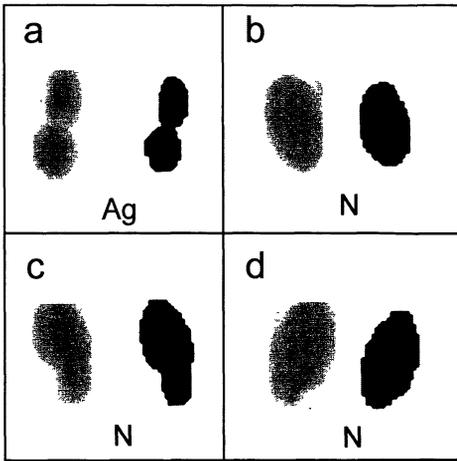


Fig. 1.

Fig. 2.

Fig. 1. — Examples of aggregates (Ag) and convex or slightly concave nuclei (N): grey level and binary images.

Fig. 2. — Ultimate eroded set (represented as white dots in objects).

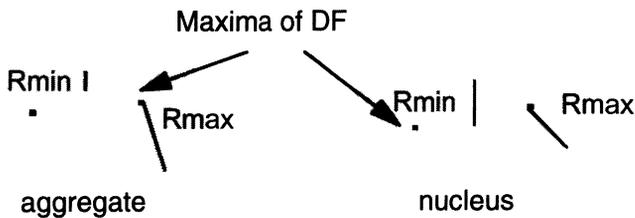


Fig. 3. — R_{min}/R_{max} evaluation: example of an aggregate and of a concave nucleus.

ultimate eroded set, whereas concave shapes have two or more [11, 12]. It must be noticed that each ultimate eroded set corresponds to the maximum of the distance function (DF). Concerning the present study, objects with one ultimate eroded set have more chances to belong to nuclei (Fig. 2b). Objects with 2 or more ultimate eroded set can correspond to aggregates (Fig. 2a) or convoluted nuclei (Figs. 2c, d). This problem cannot be solved simply by counting ultimate eroded sets. To further discriminate these nuclei from aggregates, the ratio between the minimum radius divided by the maximum radius (R_{min}/R_{max}) of each object was used (Fig. 3). As shown in Figure 3, theoretically nuclei must have high R_{min}/R_{max} whereas aggregates must have low ones. The watershed transformation is currently used to separate connected objects with the DF and its maxima as markers [5, 13, 14]. In the present study, the dodecagonal distance function (DDF) was used because of its greater sensitivity than the hexagonal one (data not shown). In order to compute R_{min}/R_{max} with global transformations, watershed is reconstructed conditionally to the objects (Fig. 4a). Resulting segments represent $2R_{min}$ (Fig. 4b). A geodesic propagation function (Fig. 4c) allows the computation of the geodesic radius, corresponding to the minimum of the geodesic propagation function. This geodesic radius is equal to R_{min} . The corresponding grey level is then associated to each segment. As shown in Figure 4d, the grey level of each segment

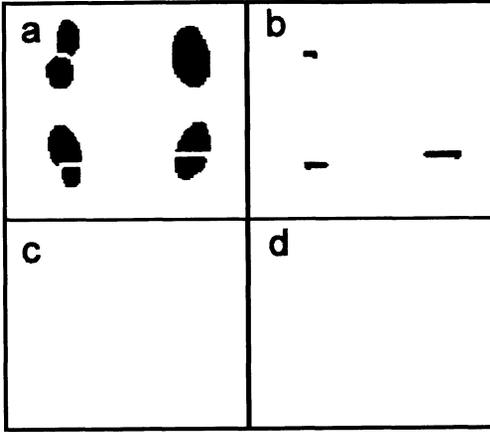


Fig. 4.

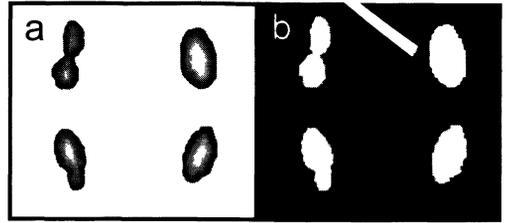


Fig. 5.

Fig. 4. — R_{\min} computation: a) conditional watershed (represented as a white segment in objects); b) $2R_{\min}$; c) geodesic propagation; d) grey level of each segment corresponding to the value of R_{\min} for each object.

Fig. 5. — R_{\max} computation: a) dodecagonal distance function; b) the grey level of each object corresponding to the value of R_{\max} for each object.

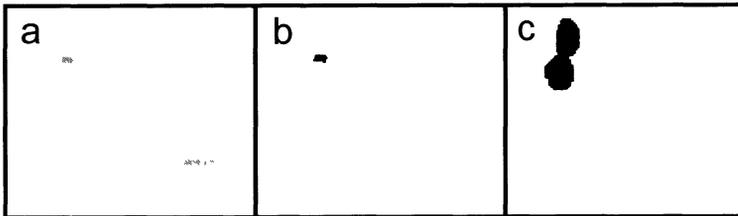


Fig. 6. — Computation of R_{\min}/R_{\max} : a) result of image 4d divided by image 5b, the grey level of each object corresponds to the value of R_{\min}/R_{\max} ; b) thresholding; c) reconstruction of the original image condition to image 6b.

then corresponds to the value of R_{\min} for each object. Computation of R_{\max} is done in the same way using the maximum of the DDF which represents the thickness of each object (Fig. 5a). The value of this maximum corresponds to R_{\max} . The corresponding grey level is associated to each object. The grey level of each object corresponds to the value of R_{\max} for each object (Fig. 5b). Dividing the two images (Figs. 4d and 5b) leads to an image where the grey level of each object corresponds to the value of its R_{\min}/R_{\max} (Fig. 6a). For each object, the value of the parameter to be estimated then corresponds to its grey level. A simple threshold giving the best contrast recovers markers for objects of interest (*i.e.* aggregates) (Fig. 6b). A conditionnal reconstruction referring to the original image allows to separate nuclei from aggregates (Fig. 6c).

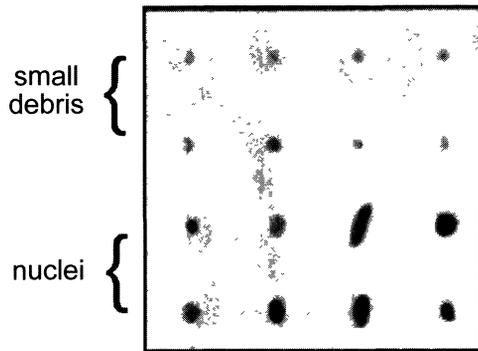


Fig. 7. — Small debris and undamaged nuclei.

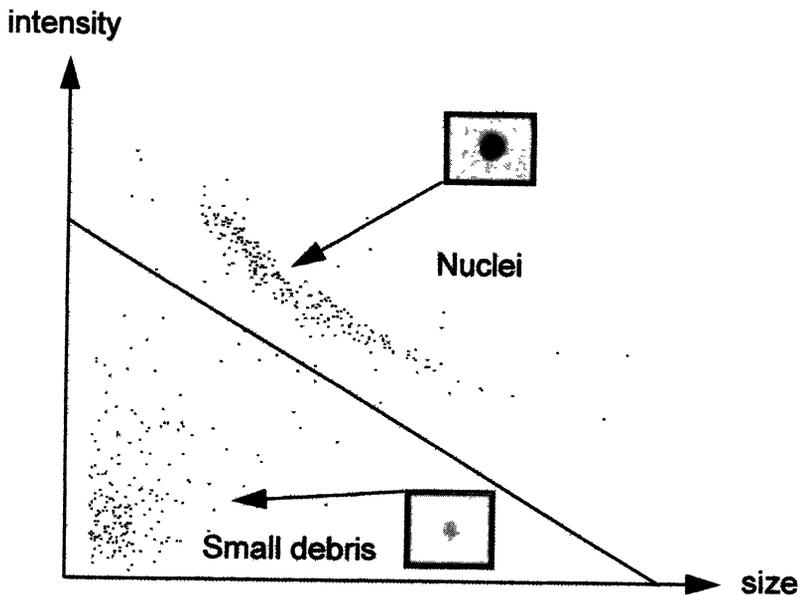


Fig. 8. — Graph intensity as a function of size, for nuclei and small debris populations.

4. Elimination of Small Debris

The small debris represent the main fraction of unwanted elements. They are characterized by a small size and a high grey level (Fig. 7). As shown on a graph of intensity as a function of size (Fig. 8), there is no intensity threshold value able to separate correctly and completely nuclei from small debris. The reason is that intensity value, which discriminates nuclei from small debris, is size dependent. To separate nuclei from small debris, the intensity threshold value must decrease when size increases. Top-hat transformation allows the detection of dark area of image $f(x)$ using closing ($C^{\lambda B}$) transformation of $f(x)$, followed by a threshold (t) of $C^{\lambda B}(f(x)) - f(x)$. The closing function acts as a sieve, the width of which is determined by the size of the closing (λB). The threshold (t) retains only objects which intensity is higher than the threshold value. By changing the size of the closing (λB) and the value of the threshold (t), it is possible to choose the width and intensity of selected objects. For each transformation and each size of top-hat, over-topping objects

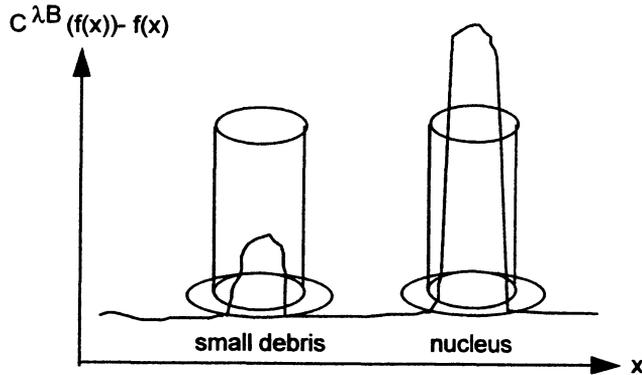


Fig. 9. — Example of top-hat transformation applied to the characterization of small debris. Over-topping objects are considered as nuclei; objects in the top-hat are labelled as small debris.

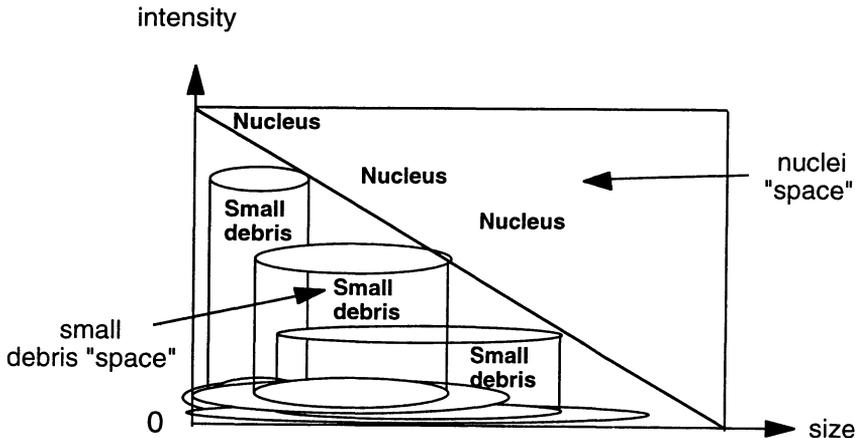


Fig. 10. — Graph intensity as a function of size. Top-hat of increasing size and decreasing width are used to fit the graph intensity as a function of size.

are labelled as nuclei, other objects are considered as small debris (Fig. 9). Top-hat of increasing width and decreasing height are applied sequentially (Fig. 10) allowing an optimal separation of small debris from nuclei.

5. Results

From a set of 50 images, 86.6% of aggregates are characterized with a false positive rate of 0.8% (nuclei falsely labelled as aggregates). 98.7% of small debris are characterized with a false positive rate equal to 3.2%. Sensitivities and false positive rates are not sufficient to evaluate the efficiency of the method, referring to the information of interest (*i.e.* the restitution of DNA ploidy abnormalities information). So, DNA ploidy curves obtained after automatic elimination of debris and aggregates must be compared to those obtained by interactive elimination (Fig. 11). Nuclei located in the tetraploid range after global automatic method represent 19.4% of the total events: a value very similar to the 20.5% obtained after interactive elimination (reference value).

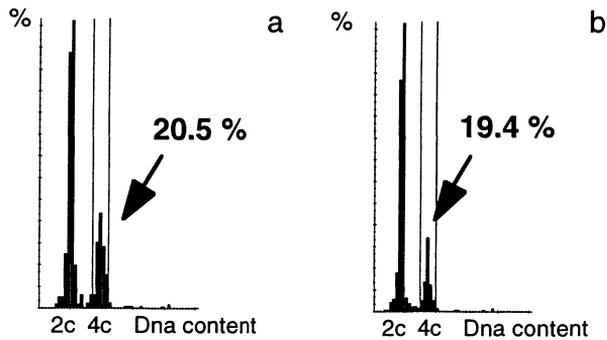


Fig. 11. — a) DNA ploidy curve of undamaged nuclei after interactive elimination of aggregates and small debris. b) DNA histogram of undamaged nuclei after automatic elimination of aggregates and small debris by global method.

6. Conclusion

Parametric methods are more often used in biology as sorting methods. Furthermore, a lot of parameters are used without true knowledge of their usefulness. Parametric methods refer to knowledge bases and used data analysis such as multiparametric analysis or logistic regression. Their computation requires powerful machines.

Mathematical morphology treatment is guided by an *a priori* knowledge of the morphological characteristics of objects to be analyzed. Depending of this knowledge, only the most suited morphological operators are used to solve a problem, reducing the number of operations to perform. Furthermore, mathematical morphology transformations are very fast, allowing time saving. However, the finding of strategies and the elaboration of such methods can be relatively long, global transformations requiring a perfect knowledge of the morphology of objects to classify. Global transformations by mathematical morphology are relevant for cell sorting because they allow high sensitivities with low false positive rates, while restoring pertinent information (DNA ploidy abnormalities) correctly.

As a conclusion, mathematical morphology global transformations are efficient alternative tools for object sorting.

Acknowledgements

This work was done under the auspices of “Pôle Traitement et Analyse d’Images (TAI) de Basse-Normandie”. Christophe Boudry has a fellowship of “Ministère de l’Enseignement Supérieur et de la Recherche”.

References

- [1] Wheelless L., Robinson R., Lapets O., Cox C., Rubio A., Weintraub M. and Benjamin L., *Cytometry* **17** (1994) 159-166.
- [2] Boudry C., Herlin P., Coster M., Chermant J.L., Sola B. and Henry-Amar M., *Acta Stereol.* **14** (1995) 121-128.

- [3] Masson E., Herlin P., Galle I., Duigou F., Belhomme P., Bloyet D. and Mandard A.M., *Acta Stereol.* **13** (1994) 75-81.
- [4] Palcic B., Mac Aulay C., Shlien S., Treurniet W., Tezcan H. and Anderson G., *Anal. Cell. Pathol.* **4** (1992) 429-441.
- [5] Coster M. and Chermant J.L., *Precis d'analyse d'images* (Les Presses du CNRS, 1989).
- [6] Meyer F., *J. Histochem. Cytochem.* **27** (1979) 128-135.
- [7] Meyer F. and Van Driel A., in *Quantitative Image Analysis in Cancer Cytology and Histology*, J.Y. Mary and J.P. Rigaut, Eds. (1986) pp. 33-49.
- [8] Boudry C., Herlin P., Coster M., Sola B. and Chermant J.L., *Analyt. Quant. Cytol. Histol.* (July 1996) in press.
- [9] Van-Driel Kulker A.M., Mesker W., Van der Burg M. and Ploem J.S. *Analyt. Quant. Cytol. Histol.* **3** (1987) 225-231.
- [10] Herlin P., Duigou F., Masson E., Bloyet D. and Mandard A.M., *Acta Stereol.* **11** (1992) 411-417.
- [11] Meyer F., *Cytologie quantitative et morphologie mathématique*, Thèse de Doctorat de l'École Nationale Supérieure des Mines de Paris, Fontainebleau (1979).
- [12] Talbot H., *Analyse morphologique de fibres minérales d'isolation*, Thèse de Doctorat de l'École Nationale Supérieure des Mines de Paris, Fontainebleau (1993).
- [13] Serra J., *Image Analysis and Mathematical Morphology* (Academic Press, 1982).
- [14] Schmitt M. and Mattioli J., *Morphologie mathématique* (Masson, 1993).