 Cytometry 28:289–297 (1997)

Applying Watershed Algorithms to the Segmentation of Clustered Nuclei

Norberto Malpica,[1]* Carlos Ortiz de Solórzano,[1] Juan José Vaquero,[1] Andrés Santos,[1] Isabel Vallcorba,[2] José Miguel García-Sagredo,[2] and Francisco del Pozo[1]

[1]Grupo de Bioingeniería y Telemedicina, Universidad Politécnica de Madrid, Madrid, Spain
[2]Servicio de Genética, Hospital Ramón y Cajal de Madrid, Madrid, Spain

Received 3 July 1996; Accepted 16 April 1997

Cluster division is a critical issue in fluorescence microscopy-based analytical cytology when preparation protocols do not provide appropriate separation of objects. Overlooking clustered nuclei and analyzing only isolated nuclei may dramatically increase analysis time or affect the statistical validation of the results. Automatic segmentation of clustered nuclei requires the implementation of specific image segmentation tools. Most algorithms are inspired by one of the two following strategies: 1) cluster division by the detection of internuclei gradients; or 2) division by definition of domains of influence (geometrical approach). Both strategies lead to completely different implementations, and usually algorithms based on a single view strategy fail to correctly segment most clustered nuclei, or perform well just for a specific type of sample. An algorithm based on morphological watersheds has been implemented and tested on the segmentation of microscopic nuclei clusters. This algorithm provides a tool that can be used for the implementation of both gradient- and domain-based algorithms, and, more importantly, for the implementation of mixed (gradient- and shape-based) algorithms. Using this algorithm, almost 90% of the test clusters were correctly segmented in peripheral blood and bone marrow preparations. The algorithm was valid for both types of samples, using the appropriate markers and transformations. Cytometry 28:289–297, 1997. © 1997 Wiley-Liss, Inc.

Key terms: FISH; interphase nuclei; fluorescence microscopy; cluster division; digital image analysis; mathematical morphology; automation

Segmentation of clustered nuclei is necessary in biological studies where knowledge of the morphology of cells or nuclei, the distribution of fluorescence signals in them, and/or the organization of cells in the tissue specimen is required. Nuclear morphology and staining quality strongly depend on the sample source, the preparation protocol, and the staining technique. This study concerns the automatic segmentation of fluorescent-stained nuclei which have suffered a complex combination of fixation, dehydration, embedding, and endogenous enzyme inactivation procedures in order to perform a non-isotopic (fluorescent) in situ hybridization of their DNA content.

In situ hybridization allows the detection of quantitative and structural aberrations in the genomic content of malignant or premalignant lesions, which are correlated with the diagnosis and prognosis of their related diseases. In fluorescence in situ hybridization techniques (FISH) (9,16,21), nucleic acid probes are labeled by fluorescent markers. These markers allow the localization of their attached DNA sequences. Automated fluorescence microscopes with charge coupled devices of high sensitivity (CCD cameras) and computer controlled scanning processes have been developed in the last years to assist in the analysis of FISH specimens (14,18). These systems achieve the detection and scoring of DNA targets using image analysis algorithms.

In fully automated applications, accurate measurements of the DNA targets assume a correct segmentation of the

A preliminary report of this work was presented at the XVIII Congress of the International Society for Analytical Cytology (Rimini, April 1996) and the abstract published in the special Cytometry Supplement 8, March 1996.

Contract grant sponsor: ARCADIM Project; Contract grant number: CICYT TIC92-0922-C02-01 (Comisión Interministerial de Ciencia y Tecnología); Contract grant sponsor: European Concerted Action CA-AMCA; Contract grant number: BMH1-CT92-1307; Contract grant sponsor: Comunidad Autónoma de Madrid (CAM); Contract grant sponsor: Universidad Politécnica de Madrid (UPM).

Carlos Ortiz de Solórzano can be reached at the Resource for Molecular Cytogenetics, Lawrence Berkeley National Laboratory, Berkeley, CA.

E-mail: Carlos@white.lbl.gov

Juan José Vaquero is now at the Department of Nuclear Medicine, National Institutes of Health, Bethesda, MD.

*Correspondence to: Norberto Malpica, Grupo de Bioingeniería y Telemedicina, E.T.S.I. Telecomunicación, Universidad Politécnica de Madrid, 28040 Madrid, Spain.

E-mail: norberto@teb.upm.es

counterstained nuclei. This is a relatively easy task, usually implemented with thresholding, region growing, or edge detection algorithms. Most of these algorithms take into account either the morphological information (size, shape, connectivity) or the pixel information (light intensity) of the objects present in each image (13). Problems arise when trying to segment clusters of nuclei in which it is difficult to make an accurate definition of each individual nuclear domain. Clusters may appear in a great percentage of the total number of nuclei, which justifies the development of specific image segmentation tools. That is the topic of a recent paper (12) where the authors review some previous strategies for cluster segmentation based on morphological information (20,24,26) or on the alternate use of morphological and intensity information (1,7).

To solve the related problems, the authors suggest that both morphological and pixel information are to be used to get a reliable segmentation. Among the algorithms that use all the information available, gray scale mathematical morphology algorithms are very attractive because they have a geometrical approach and deal easily with object-oriented criteria such as shape, size, contrast, connectivity, etc. Moreover, morphological transformations can be implemented very efficiently, both in software and hardware, and are appropriate for parallel computing implementations.

In the paper mentioned (12), the algorithm was tested on only seven images, and the number of samples from which they are taken is not specified. We have tested our algorithm using images from nine different samples, therefore taking into account a broader range of nuclei features and imaging conditions.

A method to segment nuclear clusters efficiently, based on watershed lines, will be demonstrated in this paper. This technique, which may appear to be close to region-growing methods, leads in fact to a more general segmentation methodology. The results, described below, demonstrated the adequacy of this method to segment nuclei clusters. The modularity of the proposed algorithm allows one to easily apply it to nuclei with features different from the ones presented here, just by changing specific steps of the implementation parameters.

MATERIALS AND METHODS

Sample Preparation

Nuclei were analyzed using standard cytogenetic preparations from bone marrow and peripheral blood specimens. All slide preparations underwent standard FISH procedures, except actual probe hybridization, in order to produce all possible variations in shape and size that these techniques may cause.

Bone marrow samples: Following standard laboratory cytogenetic procedures, microscope slide spreads were carried out directly, after 24 h of unstimulated culture in bone marrow aspirates. Before slide spreads, cells were harvested according to routine methods: KCl hypotonic treatment, 30 min, and 3x methanol-acetic acid (3:1) fixation; the suspension of cells was placed on microscope slides and air-dried at room temperature.

Peripheral blood samples: After 72 h of phytohemagglutinin-stimulated lymphocyte culture, cells were harvested, KCl hypotonic-treated for 10 min, methanol:glacial acetic acid-fixed, and air-dried after being placed on slides.

Both sets of slides were stored at -20°C until they were treated following procedures for normal hybridization routine (22). Probe hybridization was achieved by denaturing slides for 2 min in 70% formamide/2xSSC at 65°C, then quickly quenched in ice-cold 70% (v/v) ethanol and dehydrated in serial ethanol washes (80%, 90%, 100%). Finally, slides were stained with propidium iodide (PI).

System Description

Microscope. The microscope used is an Ergolux (Leica, Wetzlar, Germany) with a motorized scanning stage (Marzhäuser, Wetzlar, Germany), eight slides wide. The excitation/emission filter blocks, interference filters, and objectives can be placed and removed automatically. Motor control is performed by a stepping motor controller (SMOC) (Metasystem, Sandhausen, Federal Republic of Germany). This controller communicates with the CPU via a serial RS-232 connection.

Camera. We used a MicroImager 1400 CCD camera (Xillix Technologies, Richmond, BC, Canada) as an acquisition device. It contains a Kodak KAF 1400 CCD which has 1,344 x 1,038 pixels with a 6.8 x 6.8 µm pixel size. The camera is attached to the microscope using a standard C mount adapter. The CCD can be clocked in pixel additive mode (binning), in which four adjacent pixels are combined during clock out to increase sensitivity and to obtain a higher frame transfer rate. Hence a true 2 x 2 binning is achieved. The camera is controlled by a DC1 i/o board (Access Dynamics, Alamogordo, NM) which uses a VSB (VME subsystem bus) with a 16 Mb memory board to store the acquired images. The memory board and the controller are attached to the VME workstation backplane.

CPU. The core of the system is a SparcStation 4/370 (SUN Microsystems, Mountain View, CA), 32 Mb RAM, 1 Gb HD, VME bus, with a UNIX SUN OS 4.1.1 operating system. It controls the SMOC via RS232, and the Xillix camera through the DC1 board attached to its VME backplane. Images are retrieved and stored in RAM from the 16 Mb memory board through the VME bus.

Image Processing

Figure 1 summarizes the algorithm that has been used for both peripheral blood and bone marrow cluster detection and division.

Step numbers 1, 2, 3, 6, and 7 are common for both types of samples. Steps 4 and 5 are specific for each type and define the nature of the algorithm. The features used to divide the clustered nuclei are either the internuclear gradients, or shape and size measurements, or a combination of them. In order to provide a better understanding of the overall process, some insight on the watershed algorithm (steps 4, 5, 6, 7) is needed.

The watershed is a morphological algorithm which permits the detection of crest lines in images (3). Consider-

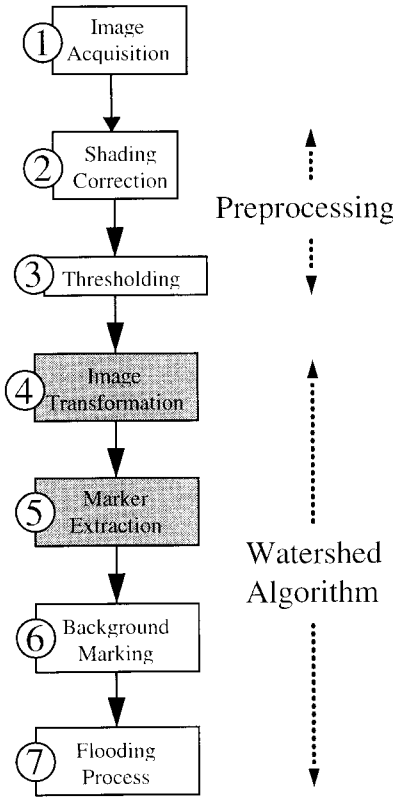


FIG. 1. Flow diagram of the overall cluster division process.

ing the gray-level image as a topographic surface, water falling on it will flow down the walls of the catchment basin corresponding to each minimum. The points where the water can flow down either one of two sides are the crest lines, which are to be detected. Efficient implementations, such as the one used in this work, may be obtained following an approach by immersion simulations. If we drilled a hole in each minimum and filled the surface progressively with water, the watersheds would correspond to the points where water coming from two different basins would meet. The concept of immersion can be closely simulated by making use of a hierarchical queue, with priorities defined by the gray level of image points. Points are recursively inserted and extracted from the queue, in an order defined by their gray levels. This way we ensure that the lowest gray levels are processed before the highest ones (4).

Then, in order to be able to apply this algorithm to detect lines dividing clustered objects, two main actions have to be carried out:

1) The original image has to be transformed into a different image, where—following the topographical simile—''crest lines'' correspond to the original image object boundaries and their inner and outer parts are ''valleys.''

2) Singular markers are to be defined on every valley of the transformed image as starting points of the flooding process.

An incorrect choice of the image transformation usually leads to an incorrect segmentation, and a failure assigning a single marker to each valley of the transformed image leads to oversegmentation.

The choice of both transforming and marking strategies has to be made regarding the image features. This is now accomplished manually but work is in progress to obtain a knowledge-based discrimination stage which will permit an automatic selection of the correct strategy, using some nuclei and cluster features (nuclei width and shape, texture, intra- and internuclei gradients, etc.). What follows is a description of each step of the overall process, including the criteria for strategy selection when necessary.

Two examples, one from a bone marrow (BM) sample and another from a peripheral blood (PB) sample will be used to illustrate the different algorithm steps. This samples where selected because they cover two very different cases.

Image acquisition. A filter block N2.1 from Leica (excitation wavelength BP 515-560, suppression filter LP 580) was used on the acquisition of the counterstained image, fitting the PI excitation/emission wavelengths (520/610). The objective used was a x63 Fluor, NA 1.30 (oil immersion).

Images have a resolution of 672 x 519 pixels and have been acquired using the binning facility of the Xillix camera for SNR improvement. That means a final pixel width at the sample of 0.21 μm. The exposure time is 0.3 sec and images are read from the camera at 8 Mbps.

Images were automatically focused using a three-phase maximum contrast search algorithm (5). Contrast was measured using an autocorrelation-based formula (25).

$$F_{\text{vollath\_1}} = \sum_{i=1}^{M-1} \sum_{j=1}^{N} g(i, j) \cdot g(i + 1, j) - \sum_{i=1}^{M-2} \sum_{j=1}^{N} g(i, j) \cdot g(i + 2, j)$$

where *M,N* are the image dimensions; *i,j* the pixel coordinates, and *g* the image gray level at the pixel. A comparative study among 15 different focusing functions was carried out. This function showed the best results. The images used to illustrate the algorithm are shown in Figure 2.

Shading correction and background substraction. Optical aberrations in objective lenses and optical filters, cell-to-cell variations in the CCD, as well as focusing misalignment of the light source, produce the same image nonuniformity on every acquired field. Correction of these inhomogeneities is achieved by multiplying each image by a background control matrix. This matrix was obtained

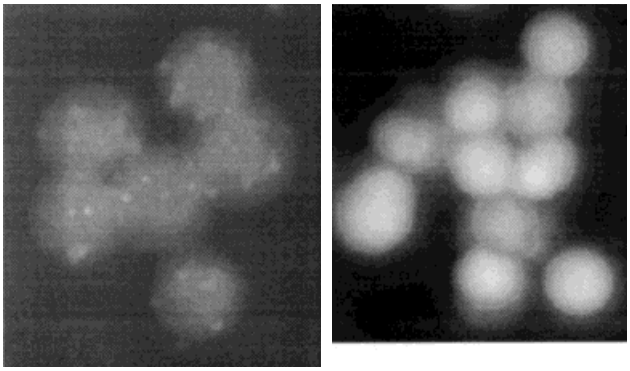


FIG. 2. Original images. (**a**) Sample from peripheral blood (PB). (**b**) Sample from a bone marrow (BM) specimen. These totally different images have been used in order to show the algorithm steps and the final algorithm performance.

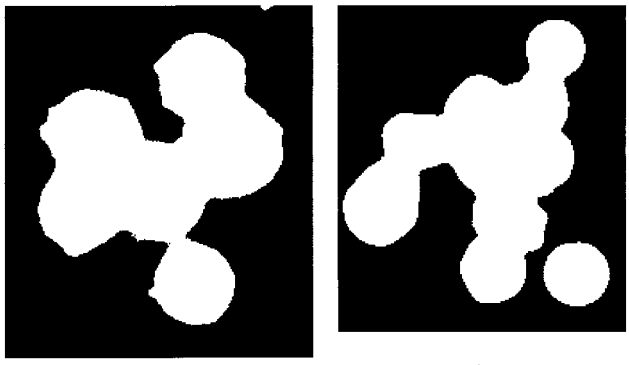


FIG. 3. ISODATA threshold of the original images. (**a**) PB sample. (**b**) BM sample.

from 25 manually selected empty images where deviation of each pixel from the image mean intensity was calculated. The shading control matrix was obtained as the mean of these 25 deviation matrices, excluding the four extreme values (highest absolute values) in each case. In this method, we assumed that the gray level given by the camera is proportional to the illumination, and not simply linear with it. This is a reasonable assumption, since for the integration time we used to acquire the images (0.3 sec) dark current can be neglected compared to the gray level of the fluorescent signal of the nuclei.

This method corrects the sample-independent background shading functions, provided that they remain unaltered. There are also background variations among images from different fields due to factors such as different fluorophore concentrations, autofluorescence, etc. This is corrected by subtracting the modal value of the background in the histogram to every pixel in the image.

Image thresholding. Even when a fixed exposure time is used in the acquisition, as in the present case, it is not correct to use a fixed threshold for the acquisition of images all along the sample, since the image mean gray level intensity usually varies among different zones in the sample due to several (uncontrolled) reasons (different counterstain-to-nuclei fixation, previous and unequal light exposure of different zones in the image, etc.).

Once corrected, images are roughly segmented by thresholding the image histogram. Several thresholding algorithms listed in (8) were compared using an objective-based evaluation as described in (13) on a set of 20 images with a mean content of five nuclei per image. The best results were obtained using the ISODATA or INTERMEANS thresholding algorithm (15).

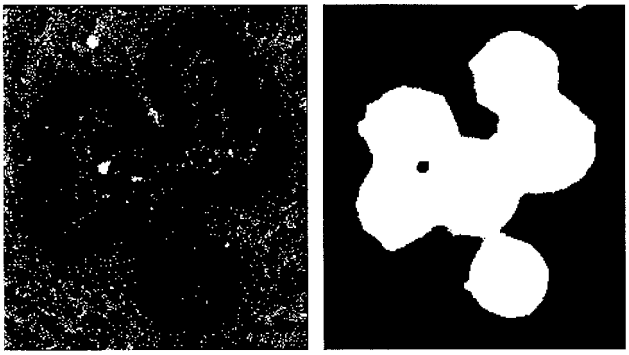
This algorithm correctly segments most isolated nuclei, but it is unable to segment nuclei in clusters (Fig. 3). The algorithm fails when trying to separate many clusters because it classifies pixels into two different groups (background and object), and there must be some pixels with background intensity between every two nuclei in order to separate them. This does not occur in most cases, where no or only little difference in intensity exists between nuclei.

After thresholding, the resulting objects are labeled and measured in order to reject non-nuclear objects. Features used to discriminate between nuclei and debris are object area, perimeter, and circularity (10). The decision step classifies the different objects into nuclei, debris, or clusters of nuclei. These clusters will be separated in a second step.

In order to evaluate the discrimination power of the three mentioned features, their Fisher Ratio (20) was computed from the values measured on the test images. For the nucleus/cluster discrimination, the three features were found discriminant enough, and therefore a Fisher Linear Discriminant function (FLD) was computed as a linear combination of these features. Coefficients were calculated using the algorithm on the training test set of images.

To determine the optimum threshold of the discriminator, its complete Receiver Operating Characteristic (ROC) curve (2) was drawn using the test set of objects. Since false negatives (clusters classified as nuclei) are much more harmful than false positives, the operating point chosen was the first point of the curve in which the False Negative Rate was zero (FNR = 0). This point is the first one where the True Positive Ratio equals one (TPR = 1), which corresponds to the one with less FPR while FNR = 0.

Threshold improvement. In clusters (or in parts of them) composed of three nuclei whose centers roughly define a triangle, if the thresholding algorithm fails to define the internuclei zone as a background area, it is useful to extract a background marker in order to be able to apply the distance transform to the image (see Image Transformation, below). This is achieved by extracting the *h*-domes of the image of the original image. *h*-domes (23)



(a) (b)

FIG. 4. Background marking applied to the image in Figure 2a. (**a**) Inverse top hat of the image. As may be seen, the internuclei background zone is extracted. Spurious minima are rejected, considering their position and size, using a morphological filter. (**b**) The background marker is superimposed on the binary image obtained after the ISODATA thresholding.

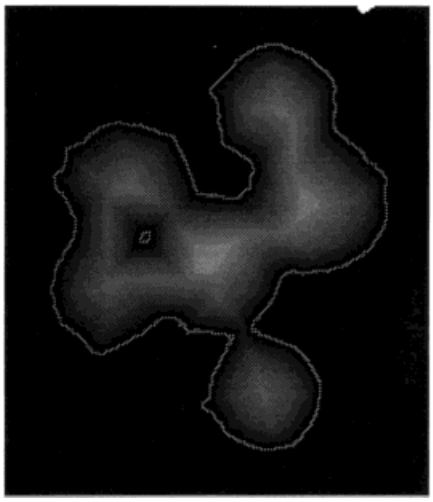


FIG. 5. Distance transform of the image in Figure 4b. Image has been processed to show more contrast and the object contour has been marked.

are the image zones fulfilling the following criteria:

- Every pixel in the dome has a gray value greater than any of the pixels surrounding it.
- The maximum gray level difference between two pixels in the dome is smaller than or equal to *h*.
- The *h*-domes are extracted using the following morphological formula:

$$M_h(I) = I - \rho_I(I - h)$$

where $M_h(I)$ represents the image of maxima, I is the initial image (original image after preprocessing), I–*h* represents the result of subtracting a constant value h to the original image, and $\rho_I(I–h)$ the morphological reconstruction of the original image from I–*h*. Morphological reconstruction (23) is obtained by conditional dilation of the decremented image under the original one, i.e., by recursively dilating I-*h*, keeping the result with value less than I, until convergence.

This transform was applied to all clusters, except those with no minima. The value of *h* used was 28, the value which performed best on the test set of images. Morphological reconstruction was achieved with a 3 x 3 square structuring element. The result of the transform, as well as the thresholding of the cluster after imposing the minima defined by it on the original images, are shown in Figure 4.

Image transformation. To be able to apply the watershed to segment nuclei, we should apply it on an image where the contours to be calculated correspond to watershed lines, and the nuclei to catchment basins surrounded by them. The gradient transformation of the original image often suits this condition. The morphological gradient of the initial gray-level image, defined as the difference between the dilation and the erosion of the gray value image, is to be used (19) to calculate the gradient image:

$$Gr(I) = \partial(I) - \epsilon(I)$$

The gradient was calculated using a 3 x 3 square structuring element. When high intranucleus gradients exist and they are higher than the internuclei ones, or when none or a small internuclei gradient exists, a different strategy should be applied to generate the surface for the watershed.

PB Clusters. In PB clusters, where nuclei are granular and no internuclei gradient exists, nuclei, even those heavily packed, preserve their almost rounded shape. We propose then the use of a geometrical approach.

The traditional way of applying geometry to segmentation problems begins with a region-growing process on a blank image, starting from markers defined for each nucleus. A watershed is applied over the binary image that contains the contours defined by the Isodata threshold. This watershed defines the Skeleton by Zones of Influence (SKIZ) (19) of the original image. The result of this method is purely geometrical and it yields incorrect results in clusters containing nuclei of different sizes. An improvement can be added by using an image with gradient information instead of the blank image in the first step.

We suggest using the inverse of the distance transform as the work image for the watershed operator. Image distance transformation is obtained by using the erosion transform. This morphological operator assigns to each point in an object the maximum number of iterative erosions of the object which still include the point.

Working with a 3 x 3 mask, the result is similar to the classical distance transform. This transform is computed in an efficient way with a recursive algorithm (6). In Figure 5, the result of the distance transformation applied on the PB sample is shown.

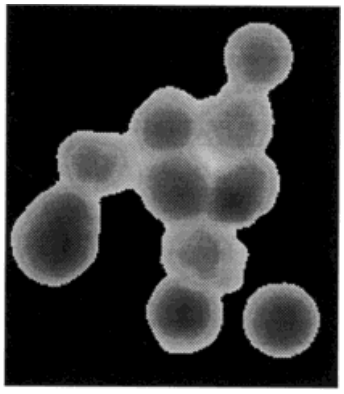


FIG. 6. Inverse of BM original image (Fig. 2b).

BM clusters. In the BM case, nuclei intensity decreases somewhat uniformly from the center to the boundaries, but gradients are low in the internuclei edges. Therefore, the inverse of the original cluster clearly defines crest lines on the nuclei boundaries, surrounding the nuclei as a crest line surrounds its catchment basin. Figure 6 illustrates this idea for the BM example.

Nuclei marking. Once the transformed image has been obtained, singular markers have to be defined and imposed as minima on the transformed image. From this minima, the watershed will find the crest lines in the transformed image by means of the simulation of the flooding process.

If a watershed transform is performed using all local minima of the gradient image, an oversegmented image will result, in which all the contours between two minima will be present. Several solutions to this problem are possible. One of them is to start the watershed only from selected points. In this case, only the contours dividing marked regions are detected. Therefore, a unique marker per nucleus has to be found. Several markers have been proposed in the literature (3), and we have tried to find the most appropriate for each type of image.

1) In PB samples, nuclei have an almost round shape. That means that the distance transform of the binary thresholded image of the nuclei will present one regional maximum per nucleus. These regional maxima are used as markers to begin the watershed. The regional maxima of the transform are extracted using the h-dome transform (see Threshold Improvement, above), with $h = 1$. This allows one to obtain a unique marker per nucleus in cases where more than one would appear using simple maxima detection (Fig. 7). Height of the *h*-domes was chosen empirically, using a test set of images. In the blood samples, maxima of the distance function are being sought, and the minimum value $h = 2$ always yields good results.

2) In BM slides, nuclei present a great contrast with the background, and a decreasing intranuclei gray level. In such images, markers can be obtained for nuclei in a more direct way, by maxima or *h*-dome extraction algorithms applied directly to the original image, using $h = 15$ as a parameter, chosen as above. (Fig. 8).

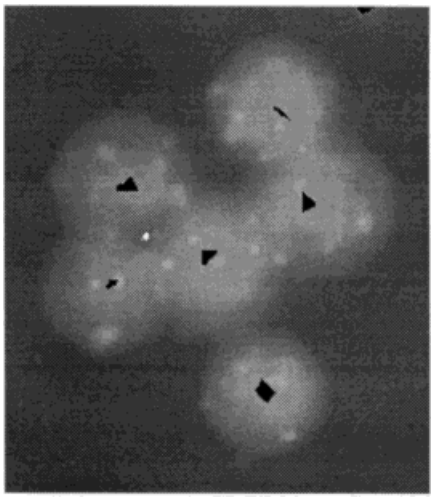
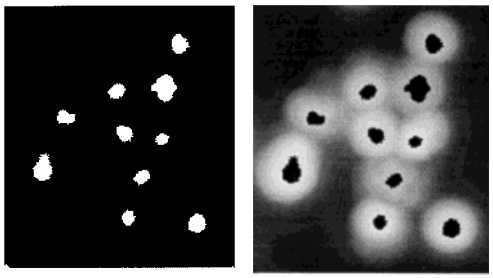


FIG. 7. Image from Figure 2a with nuclei markers (in black) and background marker (in white) superimposed.



(a) (b)

FIG. 8. (**a**) Nuclei markers for Figure 2b, obtained as the 5-domes of the original image. (**b**) Markers superimposed on the original image.

Background marking. Once an internal marker for each nucleus has been obtained, a marker for the background is needed in order to define the external contours of the clusters. Two possibilities have been considered for this purpose:

1) A background marker formed by lines which are equidistant to the internal markers. It is obtained by applying a watershed transform to a uniform gray-scale image, starting from the internal markers. This gives a result the SKIZ of the markers.

2) A background marker placed on a point which is external to the object. In this case, an external maker placed at the upper-left corner of the image was used.

Flooding process: final segmentation. The final separation of the clusters is obtained by applying the watershed to the transformed image, starting from every marker, both internal and external. Two different possibilities appear, depending on the external markers used:

Table 1
Results of the Application of the Algorithm to Bone Marrow Samples

	Nuclei analyzed	Correctly separated	Oversegmented	Fused
Training set	64	52 (81.25%)	2 (3.12%)	10 (15.62%)
Test set	106	99 (93.39%)	0 (0%)	7 (6.6%)

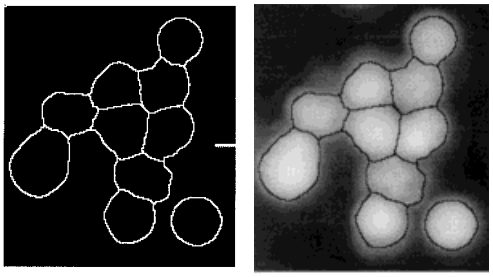


FIG. 9. Segmentation of a BM cluster into its components. (**a**) Crest lines as a result of the watershed. (**b**) Crest lines superimposed on the original image.

1) Using the SKIZ of the internal markers, nuclei are segmented as completely separated objects.

2) With the background marker artificially imposed, catchment basins originating from each of the markers touch each other, resulting in a common line of separation between both nuclei.

RESULTS

For each type of cells, a certain number of nuclei (64 in BM and 65 in PB) were used as a training set to adjust the parameters of the algorithms. A different set of nuclei, containing 106 in the bone marrow samples and 364 in the PB samples, was then used to test the algorithms. Isolated nuclei were not taken into account in the results, as the algorithm does not affect the correct segmentation of these nuclei. Five different samples (1 slide per sample) were used for the PB experiments, while four were used for the BM experiments. In both cases, training images were taken from the first two slides, and the test images were taken from remaining three PB and two BM slides.

BM Nuclei

The images were taken from four different samples (two hybridized at the same time and two more at different times). The results are shown in Table 1. The final segmentation of the BM cluster used to illustrate the algorithm is shown in Figure 9.

Some of the errors found are shown in Figures 10 a,b,c,d. Incorrect nuclear marking is the reason for the oversegmentation problems as can be seen in Figures 10 a,b. The error in the nuclei marking process is mainly due to the existence of several intensity maxima in one nucleus. In normal conditions of lighting and focusing, the existence of maxima big enough to affect the algorithm is very unusual (it only accounted for 5% of all of the nuclei).

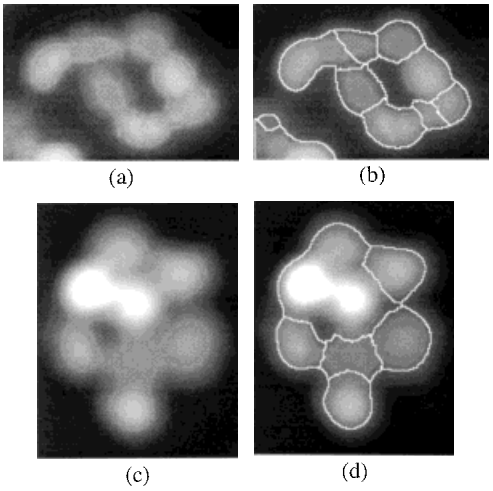


FIG. 10. Examples of incorrect BM cluster segmentation. (**a**) and (**c**) Original images. (**b**) Final segmentation showing two oversegmented nuclei. (**d**) Final segmentation showing three fused nuclei.

Table 2
Results of the Application of the Algorithm to Peripheral Blood Samples

	Nuclei analyzed	Correctly separated	Oversegmented	Fused
Training set	65	58 (89.21%)	1 (1.53%)	6 (9.23%)
Test set	364	320 (87.91%)	17 (4.67%)	27 (7.41%)

Nuclei may appear fused (Figures 10 c,d) due to two opposite causes: overexposure or underexposure of the image. In both cases, the algorithm fails when trying to find significant image maxima which can be used as markers of the clustered nuclei.

The final results are closely associated with a correct selection of the microscope and camera settings regarding to the image characteristics. Microscope lamp misalignment, incorrect exposure time for the image acquisition, and out-of-focus image acquisition influence the performance of the algorithm.

PB Samples

Images were obtained from five different samples, two of which were hybridized at the same time. The results obtained using the transformation and markers described in the previous sections are shown in Table 2. The results obtained on the image example are shown in Figure 11.

Oversegmentation and fusion uncertainties (Fig. 12) are due to morphological and artifactual problems, and are more difficult to remove than in the BM samples.

DISCUSSION

The performance of watershed-based algorithms for the segmentation of fluorescent-stained nuclei clusters has been demonstrated in this paper. This mathematical mor-

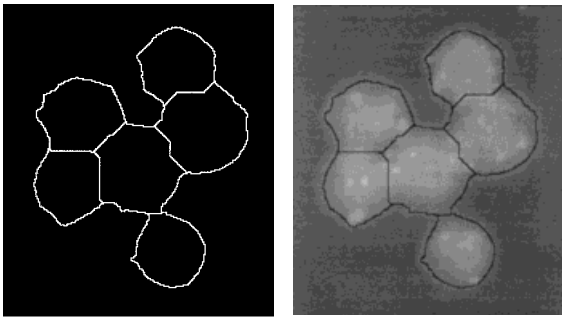


FIG. 11. Segmentation of a PB cluster into its components. (**a**) Crest lines as a result of the watershed. (**b**) Crest lines superimposed on the original image.

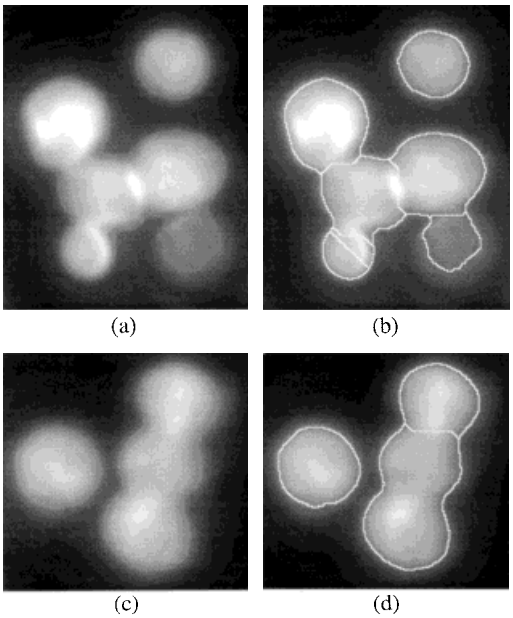


FIG. 12. Examples of incorrect PB cluster segmentation. (**a**) and (**c**) Original images. (**b**) Final segmentation showing an oversegmented nuclei. (**d**) Final segmentation showing two fused images.

phology-based algorithm uses both intensity and morphology information in order to solve the ambiguities due to overlapping nuclei on the samples. We have focused our study on the selection of the growing and marking strategies, regarding the brightness, morphology, and internuclei gradients of the clusters, because an incorrect selection of these methods dramatically affects algorithm performance.

The algorithm was tested on a broader range of samples (four BM samples and five PB samples) than previously presented algorithms (12) and the results obtained are at least as good as the ones obtained previously.

The proposed algorithm is very robust in the sense that it can easily be fitted to other types of clusters by finding appropriate markers for them. It is more flexible, as it is made up of several subroutines which can be adapted to specific problems. New types of nuclei can easily be segmented by changing specific modules or parameters in the global segmentation process.

Results are globally satisfactory, since almost 90% of the nuclei have been segmented correctly in both PB and BM samples. The remaining 10% of incorrectly segmented nuclei are due to an incorrect selection of the microscope and camera parameters for the image acquisition or to artifactual or morphological problems that would also produce errors in the visual analysis of the samples, and would therefore be rejected anyway.

In the case of PB samples, most segmentation problems (lack of markers in some of the nuclei) are due to nuclei which are too close together for the distance transform to separate both nuclei. This type of cluster would not be taken into account by a human counter for the counting of FISH spots anyway, because it would be impossible to assign a certain spot to a specific nucleus inside the cluster.

The algorithm parameters (background and nuclei marker size) used for the evaluation have been set manually after visual inspection of some nuclei on both samples.

Future work will focus on the automatic setting of watershed parameters, based on the automatic discrimination of the basic cluster and nuclei features, such as shape, size, texture, and brightness.

A further advantage of the watershed algorithm is that it is easy to implement in a parallel configuration. Work is in progress to implement and evaluate a parallel version of the whole cluster separation process.

LITERATURE CITED

1. Ahrens P, Schleicher A, Zilles K, Werner L: Image analysis of Nissl-stained neuronal perikarya in the primary visual cortex of the rat: Automatic detection and segmentation of neuronal profiles with nuclei and nucleoli. J Microscopy 157:349–365, 1990.
2. Beck JR, Chultz EK: The use of Relative Operating Characteristic (ROC) curves in test performance evaluation. Arch Pathol Lab Med 110:13–20, 1986.
3. Beucher S: The watershed transformation applied to image segmentation. Scanning Microsc Suppl 6:299–314, 1992.
4. Beucher S, Meyer F: The morphological approach to segmentation: The watershed transformation. In: Mathematical Morphology in Image Processing. E. Dougherty (ed). Marcel Dekker, New York, 1992.
5. Boddeke FR, van Vliet LJ, Netten H, Young IT: Autofocusing in microscopy based on the OTF and sampling. Bioimaging 2:193–203, 1994.
6. Chen S, Haralick M: Recursive erosion, dilation opening and closing transforms. IEEE Trans Image Proc 4:335–345, 1995.
7. Garbay C, Chassery J-M, Brugal G: An interactive region-growing process for cell image segmentation based on local color similarity and global shape criteria. Anal Quant Cytol Histol 8:25–34, 1986.
8. Glasbey, CA: An analysis of histogram-based thresholding algorithms. CVGIP: Graphical Models and Image Processing 55:532–537, 1993.
9. Griffin DK: Fluorescent in situ hybridization for the diagnosis of genetic disease at postnatal, prenatal, and preimplantation stages. Int Rev Cytol 153:1–40, 1994.
10. Haralick RM, Shapiro LG: Computer and Robot Vision, Vol. I. Addison-Wesley, Reading, 1992.
11. Joensuu H: DNA flow cytometry in the prediction of survival and response to radio therapy in head and neck cancer. A review. Acta Oncol 29:513–516, 1990.
12. Lockett SJ, Herman B: Automatic detection of clustered, fluorescent-stained nuclei by digital image-based cytometry. Cytometry 17:1–12, 1994.
13. MacAulay C, Palcic B: A comparison of some quick and simple threshold selection methods for stained cells. Anal Quant Cytol Histol 10:134–138, 1988.
14. Mascio LN, Verbeek PW, Sudar D, Kuo W, Gray JW: Semiautomated DNA probe mapping using digital imaging microscopy. I. System development. Cytometry 19:51–59, 1995.
15. Otsu N: A threshold selection method from gray level histogram. IEEE Trans System Man Cybernetics SMC 8:62–66, 1978

16. Pinkel D, Straume T, Gray J: Cytogenetic analysis using quantitative high sensitivity fluorescence hybridisation. Proc Natl Acad Sci USA 83:2934–2938, 1986.
17. Poddighe PJ, Ramaekers FCS, Hopman AH: Interphase cytogenetics of tumours. J Pathol 166:215:224, 1992
18. Sakamoto M, Pinkel D, Mascio L, Sudar D, Peters D, Kuo W, Yamakawa K, Nakamura Y, Drabkin H, Jericevic Z, Smith L, Gray JW: Semiautomated DNA probe mapping using digital imaging microscopy. II. System performance. Cytometry 19:60–69, 1995.
19. Serra J: Image analysis and mathematical morphology. London: Academic Press, 1982.
20. Schalkoff R: Pattern recognition: Statistical, structural and neural approaches. New York: Wiley & Sons, 1992.
21. Trask B: Fluorescence in situ hybridization. A review of the role in genome analysis, supporting background information and detailed protocols. In: Genome Analysis: A Laboratory Manual. B. Birren, E. Green, P. Hieter, and R. Miers (eds). Cold Spring Harbor Laboratory Press, New York, 1995.
22. Vázquez-Mazariego Y, Cabello P, García-Sagredo JM, Lopez-Yarto A, Vallcorba I, Resino M, Muñoz R, Pérez I, Mayayo M, Ferro MT, SanRomán C: Burkitt lymphoma with a duplication of der(8) t(2;8) derivative: Interpretation of a complex caryotype by chromosome painting. Cancer Genet Cytogenet 76:136–139, 1994.
23. Vincent L: Morphological grayscale reconstruction in image analysis: Applications and efficient algorithms. IEEE Trans Image Proc 2:176–201, 1993.
24. Visscher DW, Zarbo RJ, Greenawald KA, Crissman JD: Prognostic significance of morphological parameters and flow cytometric DNA analysis in carcinoma of the breast. Pathol Ann 25:171–210, 1990.
25. Vollath D: The influence of the scene parameters and of noise on the behaviour of automatic focusing algorithms. J Microsc 151:133–146, 1988.
26. Wolf G: Use of global information and a priori knowledge for segmentation of objects: Algorithms and applications. Proc SPIE 1660:397–408, 1992.