Segmentation of Color Images from Serous Cytology for Automated Cell Classification

Olivier Lezoray, Abderrahim Elmoataz, A.P., Hubert Cardot, A.P., Gilles Gougeon, A.P., Michel Lecluse, C.T., Hubert Elie, M.D., and Marinette Revenu, Pr.

OBJECTIVE: To design an automated system for the classification of cells based on analysis of serous cytology, with the aim of segmenting both cytoplasm and nucleus using color information from the images as the main characteristic of the cells.

STUDY DESIGN: The segmentation strategy uses color information coupled with mathematical morphology tools, such as watersheds. Cytoplasm and nuclei of all diagnostic cells are retained; erythrocytes and debris are eliminated. Special techniques are used for the separation of clustered cells.

RESULTS: A large set of cells was assessed by experts to score the segmentation success rate. All cells were segmented whatever their spatial configurations. The average success rate was 92.5% for nuclei and 91.1% for cytoplasm.

CONCLUSION: This color information–based segmentation of images of serous cells is accurate and provides a useful tool. This segmentation strategy will improve the automated classification of cells. (Analyt Quant Cytol Histol 2000;22:311–322)

Keywords: image analysis, computer-assisted; computer-assisted diagnosis; mass screening; segmentation of color images.

Human screening is a manual activity that has its limits. Its aim is the detection of abnormal or suspicious cells to establish an accurate diagnosis. This manual screening of cytologic slides is described as "intense, particularly complex and whose result relies on the interpretation of the human being."[6] The small amount of abnormal cells observed in the screening process requires high concentration from the cytotechnologist. Because of screening subjectivity, some errors can occur and cause false negative diagnoses. The risk of false negative slides is currently the main criticism of conventional cyto-

From the University Laboratory of Applied Sciences and Department of Pathology, Louis Pasteur Hospital Center, Cherbourg, and the Research Group in Computer Science, Image and Instrumentation, Caen, France.

Mr. Lezoray is Doctoral Candidate, University Laboratory of Applied Sciences of Cherbourg.

Mr. Elmoataz is Assistant Professor, University Laboratory of Applied Sciences of Cherbourg and Research Group in Computer Science, Image and Instrumentation.

Messrs. Cardot and Gougeon are Assistant Professors, University Laboratory of Applied Sciences of Cherbourg.

Mr. Lecluse is Cytotechnologist, Department of Pathology, Louis Pasteur Hospital Center.

Dr. Elie is Head, Department of Pathology, Louis Pasteur Hospital Center.

Ms. Revenu is Professor, Research Group in Computer Science, Image and Instrumentation.

This work was funded by the Manche Department Commitee, National Cancer League, France.

Address reprint requests to: Olivier Lezoray, Service d'Anatomie et de Cytologie Pathologiques, CHLP, 46 rue du val de saire, 50100 Cherbourg, France.

Financial Disclosure: The authors have no connection to any companies or products mentioned in this article.

Received for publication January 13, 1998.

Accepted for publication November 18, 1999.

0884-6812/00/2204-0311/$15.00/0

logic diagnoses. None of the corrective methods are completely accurate (rescreening of a certain number of slides, screener training, and so forth).[11] A promising approach is to assist the cytotechnologist in finding abnormal cells in the smear.

Research has been undertaken to use semiautomated systems for the rescreening of slides.[13,14] Such systems help in the detection of abnormal or suspicious cells in slides: they allow the recognition of possibly abnormal cells and select the most significant for cytotechnologist review. These semiautomated systems therefore contribute to the reduction of screening errors. Some automated rescreening systems have been developed. They are designed to operate in conjunction with a human screening process by rescreening all standard smears that are determined to be normal by cytotechnologists.[5-9,12] This is also the objective of a system that we are developing; its role is to assist screening by using image analysis and cellular classification: it is called Aide à la Recherche en Cytologie par le Tri Informatique Cellulaire (ARCTIC), which stands for help with research in cytology by the computerized sorting of cells. The system uses color images in various color spaces. The intended use is the cytologic assessment of serous effusions. In this paper the color-based segmentation method of the system is presented, and the percentage of correctly segmented cells is scored. The segmentation step is the most difficult and the most critical in an automated system. It is intuitive for the human observer to segment an image based on the color of the cells. A machine, however, relies on digital processing techniques to determine the set of pixels that defines each region of the image. Our work concerns the development of a segmentation strategy for color images of cells from serous cytology. This strategy relies on the extraction of regions by mathematical morphology using color information in many color spaces. It is used to obtain the cytoplasmic and nuclear boundaries of cells from serous effusions.

Materials and Methods

Preparation of the Specimens

Cell preparations are obtained by centrifugation (1,800 rpm, 20 minutes) followed by fixation in acetone-alcohol. Slides are stained by the Papanicolaou international standard of coloration.[8] The study is based on a database of digitized cell images, collected from pleural and peritoneal effusions with different pathologies. Images are either systematically taken at random from a slide or from different slides in which cells were interesting from an image analysis point of view.

Imaging System

The slides are examined using an Olympus BX-50 microscope (Tokyo, Japan), with a 20× objective, equipped with a Sony 3CCD XC-003P color video camera (Tokyo, Japan). The magnified microscopic image is captured by the three-chip camera and digitized to a 512×512-pixel color image at a photometric resolution of 3×8 bits. The pixel separation corresponds to 0.349 μm. The digitized images are stored on a Windows NT™–based computer hard drive (Microsoft France, Les Ulis Cedex, France). They are referenced in an ACCESS™ database (Microsoft France) according to the number of the slide, the code corresponding to the pathology of the cells in the image and comments entered by a pathologist on the organ and sample sources. This database enables us to test the segmentation process for various cell types, such as normal or abnormal cells; for particular spatial configurations, such as clusters, overlapping or barely touching cells; and for different types of image background (homogeneous, hemorrhagic or mucoproteinic).

Image Acquisition

Multispectral images are used for further interpretation. The camera is calibrated with a blank-field image to correctly adjust the white balance. The acquisition procedure always remains the same: 20× objective, 0.3 NA of the condenser and a stabilized light source for constant and uniform illumination of the slide field. The voltage is maintained at a value of 9 V, corresponding to the color temperature of 5,500°K. This voltage is chosen in compliance with the manuals of the microscope and camera and that of Commission International de l'éclairage standard illumination.

Before each acquisition, both the camera and microscope are left turned on for a minimum of an hour; this is necessary to reach the thermal equilibrium of the system and avoids a drift in the system response during acquisition. This fact has been established and verified by comparing the difference between two images in the same blank field taken at different times after switching on the camera and microscope. Then the remaining noise, which corresponds to a gaussian random variable, is studied. This noise is corrected later, during the image processing steps, by a smoothing operation. Along

with the noise, some nonuniformity remains in every acquired field. To correct those inhomogeneities (coming from optical aberrations of the camera), division of each recorded image by a background reference image is performed. This image is obtained from an empty image field.

User Interface

The images are stored in a Compaq Deskpro 6000 (Compaq Computer S.A.R.L., Issy-les-Moulineaux Cedex, France) on Windows NT 4.0 (Microsoft, Redmond, Washington, U.S.A.). The ARCTIC program is used to display and analyze the images in BMP, TIFF or PAN[2] format. Different morphologic operators can be used that can also perform the whole segmentation on the images. The segmentation can be manually assessed by pathologists or image analysis experts.

Color Image Analysis Procedure

Formulas

The color of a point is given by a vector, each component giving the amount of red, green and blue (RGB) of the point. There are other three-dimensional color spaces based on variables obtained from the RGB color space. In the segmentation, hue, saturation luminance (HSL) and L*u*v* color spaces are used. They can be computed with the formulas given below[7,10,15]:

$$H = ar\cos\left[\frac{((R-G)+(R-B))}{2\sqrt{(R-G)^2+(R-B)(G-B)}}\right], \text{ and if } G<B: H = 2\pi - H,$$

$$S = 1 - \frac{3\min(R,G,B)}{L}.$$

$$L = R+G+B.$$

$$L^* = 116\left(\frac{Y}{Y0}\right)^{\frac{1}{3}} - 16 \text{ if } \left(\frac{Y}{Y0}\right) > 0.008856.$$

$$L^* = 903.3\left(\frac{Y}{Y0}\right) \text{ elsewhere.}$$

$$u^* = 13L^*(u' - u'_0).$$

$$v^* = 13L^*(v' - v'_0).$$

$$u' = \frac{4X}{X+15Y+3Z} \qquad v' = \frac{9X}{X+15Y+3Z}$$

$$u'_0 = \frac{4X0}{X0+15Y0+3Z0} \qquad v'_0 = \frac{9X0}{X0+15Y0+3Z0}.$$

$$\text{with } \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 2.76 & 1.7518 & 1.13 \\ 1 & 4.5907 & 0.0601 \\ 0 & 0.565 & 5.5943 \end{pmatrix} * \begin{pmatrix} R \\ V \\ B \end{pmatrix} \text{ and } \begin{pmatrix} X0 \\ Y0 \\ Z0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

The HSL color space is more intuitive and allows an interpretation much closer to human vision. The L*u*v* color space is perceptually uniform. The use of color in image analysis is recent. In cytology some studies using color can be found; they deal with the segmentation[3] of cells and characterization of a texture.[4]

Aim of the Segmentation of Images

Both cytoplasm and nuclei of cells have to be segmented, the cytoplasm to obtain the context information (to characterize isolated or clustered cells) and the nucleus for grading malignancy. Once segmented, the cell can be classified among cellular types (from normal to abnormal). To segment the images correctly, the contexts of the images should be known precisely. The images are color images that can be divided into three groups according to their background: homogeneous, hemorrhagic or mucoproteinic (Figures 1–3).

Cells have green cytoplasm and a blue nucleus except for the erythrocytes, which are red. Both the spatial configuration and the color of the cells have extreme variability. Cells with a pale or dark nucleus (in a blue tone) or clustered cells having very dark nuclei (primarily owing to their overlapping) or chromatin irregularities can be found in the images. The segmentation strategy takes place in a sequential manner. The fact that there are some nuclei in the processed image is determined first. If not, no further examination is performed. When nuclei are present, all the erythrocytes are eliminated because they interfere with the determination of cell regions. At this stage the cytoplasm and nuclei are separately subjected to a color-based segmentation process. Finally, touching nuclei are separated and their cytoplasm dissociated in order to have only one nucleus or cluster of nuclei per cytoplasm. The whole strategy is summarized in Figure 4.

Recognition of Nucleated Cells

The first step in the analysis of an image is to determine whether there are nucleated cells in the image. In this application, cells are of interest if they have a blue nucleus, so the fact that there are some blue objects in the images has to be established. To achieve this, the HSL color space is used; based on hue intensity, a nuclear presence or nonpresence is determined. A double thresholding of hue selects the objects having a given blue hue—i.e., nucleated cells. The two thresholds are determined by measuring various hues of nuclei. This step does not allow

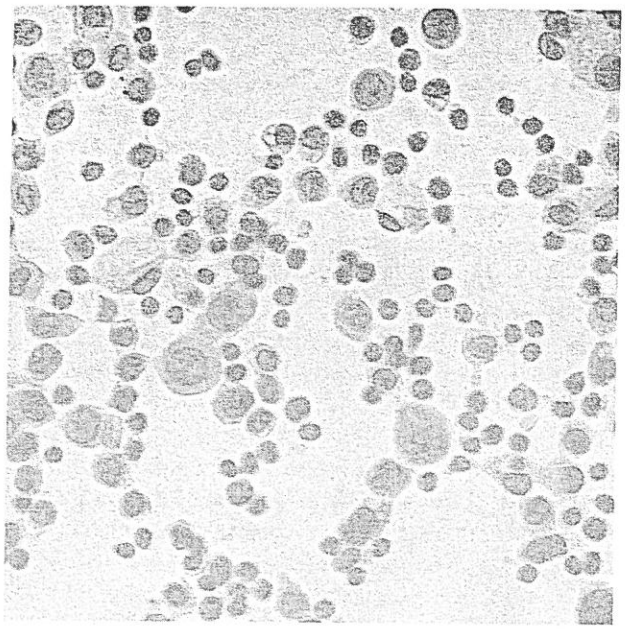


Figure 1 Homogeneous background.

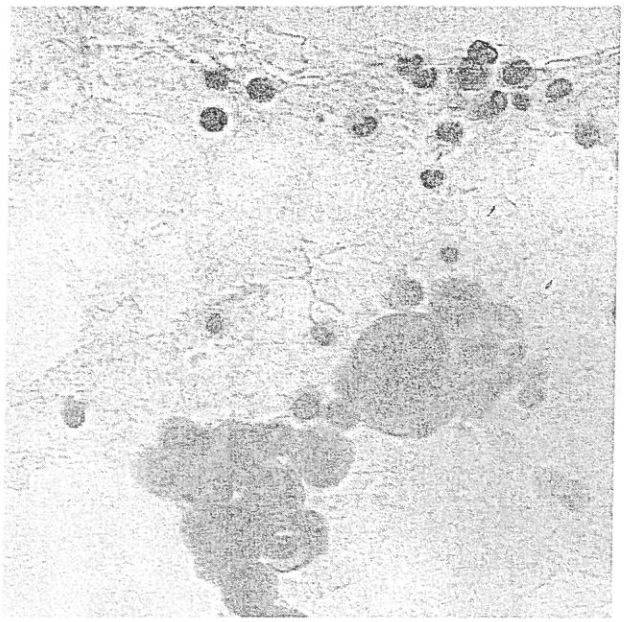


Figure 3 Hemorrhagic background.

recognition of the nucleated cells as such but focuses on blue objects and determines the presence of nucleated cells of interest even with erythrocytes in the background.

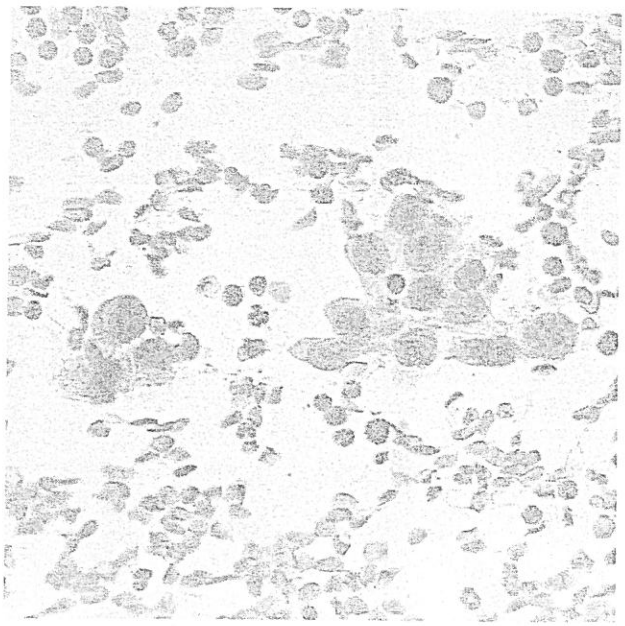


Figure 2 Mucoproteinic background.

Erythrocyte Elimination

Once the presence of nucleated cells in the image

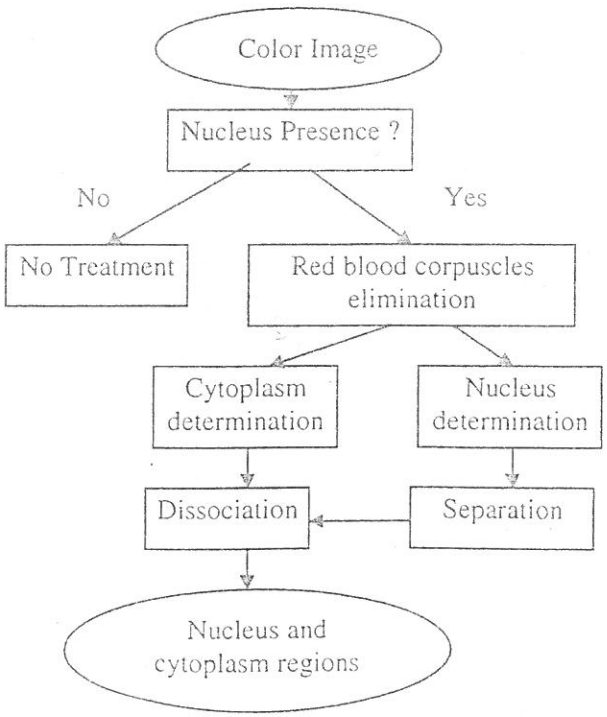


Figure 4 Segmentation strategy.

has been determined, the erythrocytes have to be eliminated. As before, the elements of which the hue characterizes red objects are selected. An opening operation is performed on those objects to refine their boundaries. The binary mask obtained defines the regions of the image having no erythrocytes. An example of this mask is given in Figure 5A. The use of such a mask reduces the number of cells to be segmented significantly. However, it does not interfere with the later recognition of nucleated cells since it does not eliminate the nucleated cells overlapped by erythrocytes.

Segmenting the Cytoplasm and Nuclei

In the next segmentation steps, a morphologic segmentation method is developed. It is based on the use of watersheds[1] applied to color images. The principle of watershed is twofold:

- Image simplification based on a set of markers extracted by procedures that take into account some knowledge of the objects, the image context and the image-processing operators being used.
- Building the watershed lines of a "potential function" using the precedent markers as local minima for optimal region growing. By imposing those local minima on the potential function, only one connected region will be obtained for each object.

The color information contained in the images is used for extracting the markers and building the watershed lines. For each segmentation step, the color space that gives the best results is determined and applied to the extraction of the regions.

Extracting Cytoplasmic Regions

To extract the cytoplasm regions, a color watershed is used. The cytoplasm has to be isolated for all cells, whether isolated or clustered. For the cytoplasm of clustered cells, further analysis is needed to dissociate cells of which the cytoplasm touches. To find the cytoplasmic markers, various color spaces were studied. Two color spaces were finally used: HSL and RGB. Markers are found by automatic thresholding of the luminance in the HSL color space. The use of the luminance is suggested by the difference in brightness in the different areas of an image: the background is brighter than the cytoplasm, and the transitions between both are well defined. This allows a reliable detection of markers for all the images. To minimize Gaussian noise (intrinsically contained in the images), which causes residual variations, exponential smoothing is performed before determining a threshold. The influence zones of the markers are obtained by multiple erosions of the markers and background. Once the markers are found, the watershed can be computed on a potential function. The color gradient in the

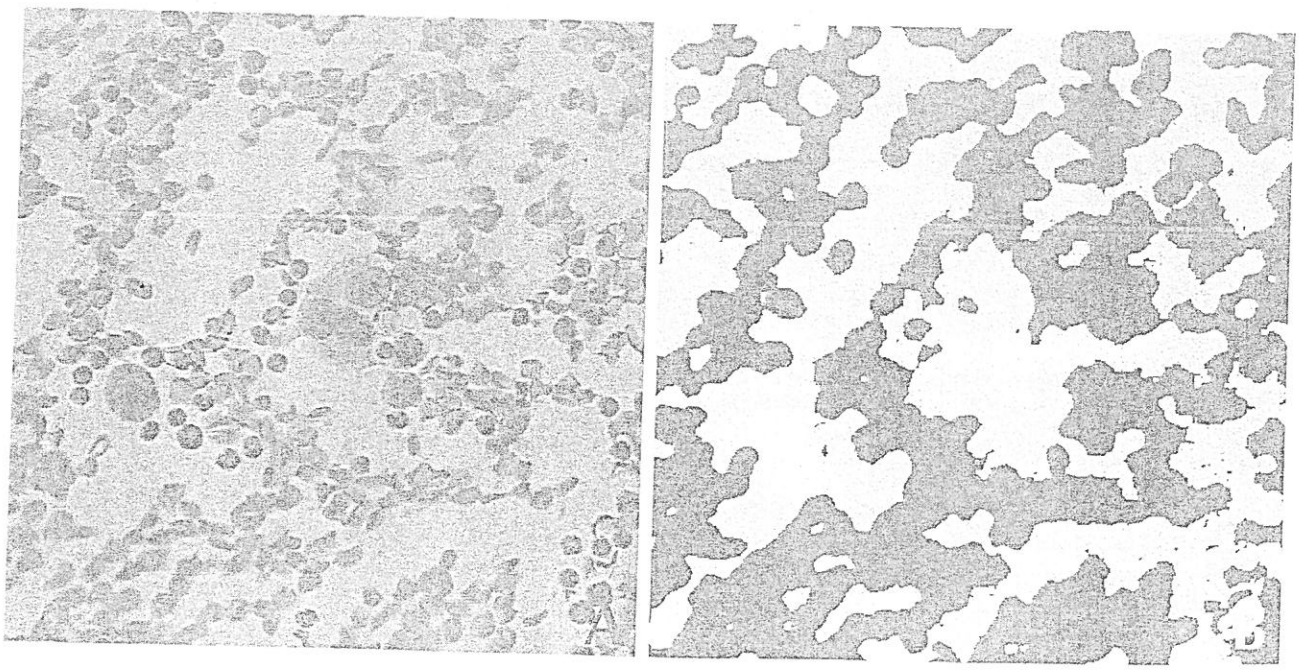


Figure 5 (A) One image with erythrocytes and (B) the corresponding mask.

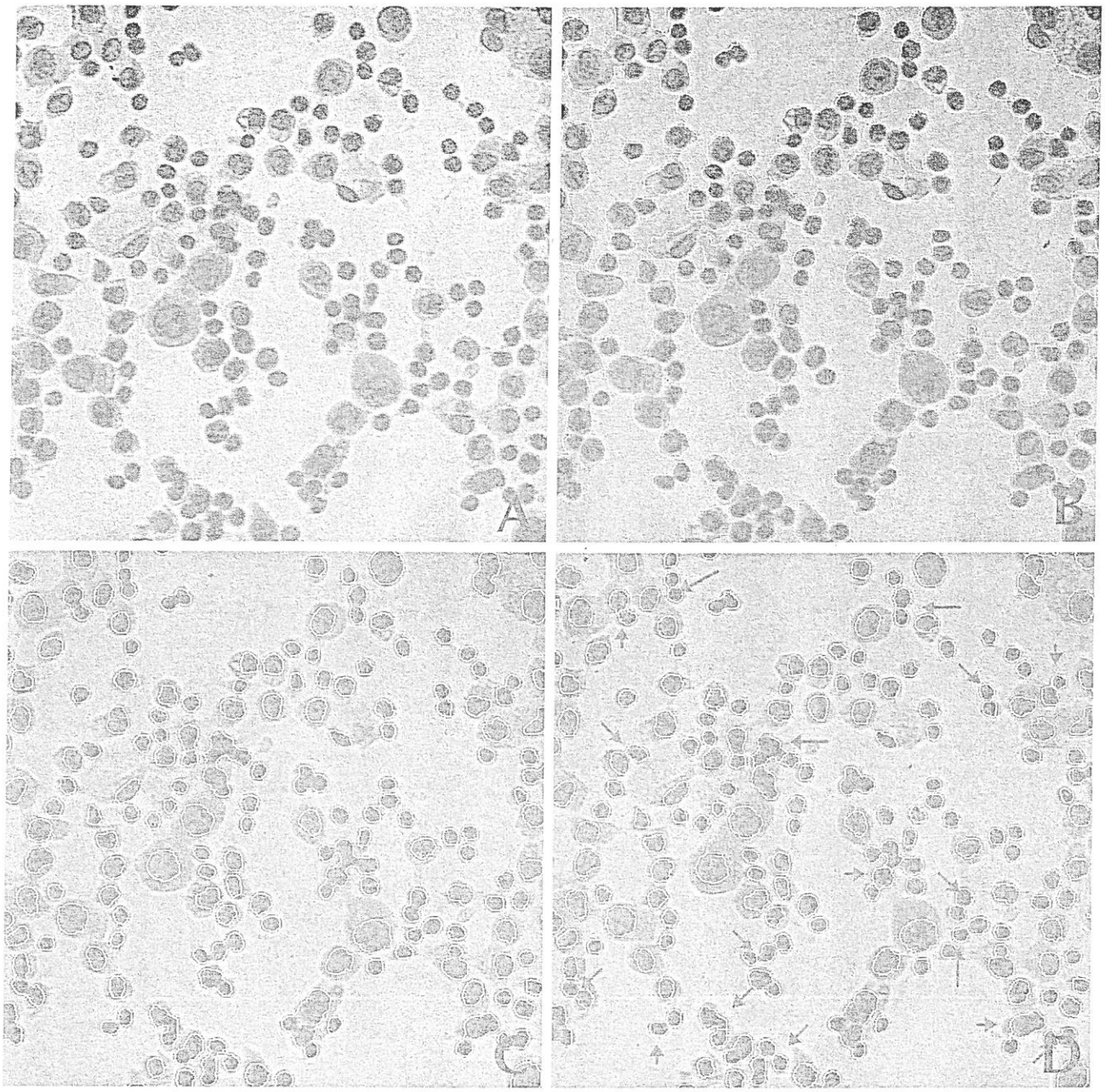
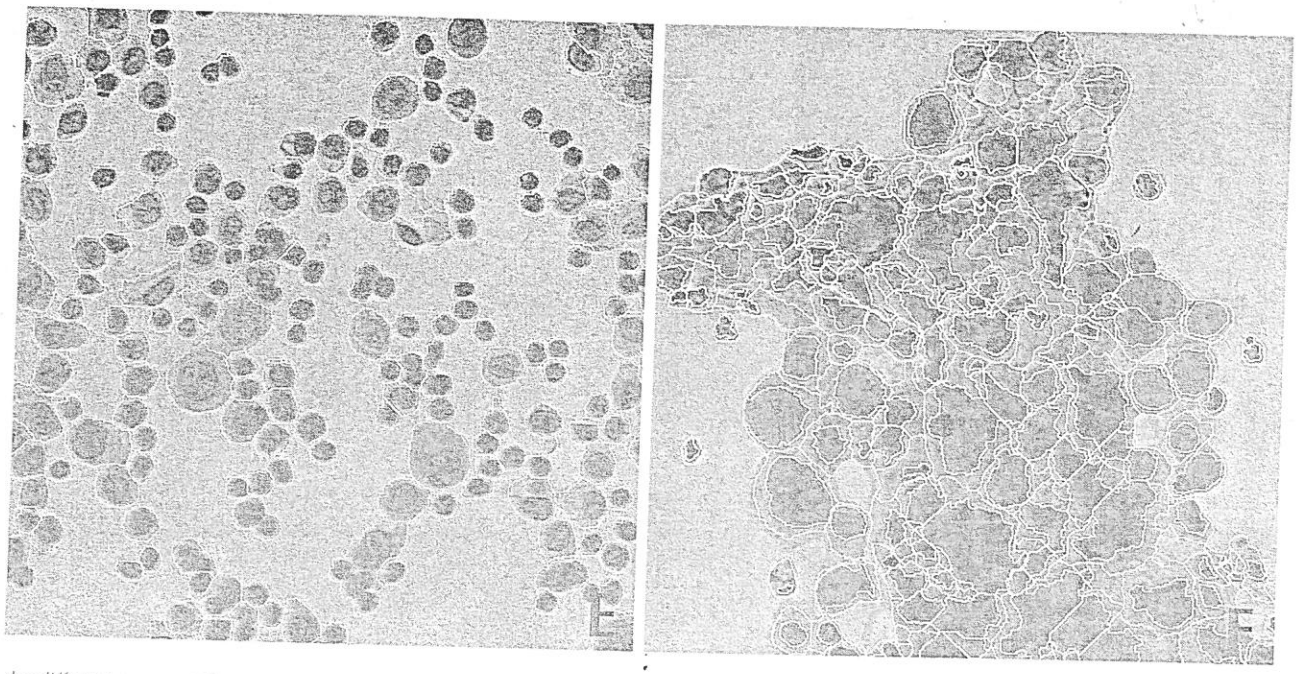


Figure 6 Images obtained during the different segmentation steps. (A) Initial image. (B) Initial image with highlighted cytoplasm. (C) Initial image with highlighted nuclei. (D) Nucleus after separation. Separated nuclei are shown with arrows. (E and F) Images obtained during

RGB color space is then used. The color gradient has the property to be higher for points where the intensity of the image presents high discontinuity and to be lower for points where the intensity of the image presents small variations. The gradient is really appropriate for color watersheds because the color transitions between background and cytoplasm are sharp. Those transitions define edges of large gradient magnitudes, which allow defining the cytoplasm regions precisely. An example is given in Figure 6B.

Extracting Nuclear Regions

The next step of the segmentation is to delineate nu-



the different segmentation steps (continued). (E) Cytoplasm after separation into that from only one nucleus. (F) Overlaid segmented cytoplasm and nuclei in white, from (7A).

clei within the cytoplasm. Several color spaces are used in this step. Due to chromatin irregularities, some nuclei appear to be very pale, although others appear very dark; that makes establishing good nuclear markers difficult. To find the latter in a reliable way, images from two different color channels are used.

The first image is computed to bring out the bright nuclei; it is the v* channel from the L*u*v* color space. (v* Is a chromatic channel of the image.) The second image, called I, is the subtracted blue minus green image. This image emphasizes the difference between nuclei and cytoplasm and levels the hue difference between all nuclei. The mean of those two images gives an I_M image, which emphasizes both the pale and dark nuclei, enabling the selection of all of them (Figure 7).

$$I_M = \frac{(B-G)+v^*}{2}.$$

To obtain a noise-reduced image, the color RGB image is smoothed prior to computing the I_M image. Extraction of the nuclear markers is done by automatic thresholding. The threshold is found from the first minimum to the left of the maximum peak of the histogram. This threshold separates the image in two regions, the nuclear pixels and the others. This threshold is determined so as to minimize the cytoplasmic area that could be included in the markers. (The boundary fits the nucleus better if it contains less cytoplasm—i.e., less green.) Since the image I levels differences between all the nuclei and emphasizes the difference between the nuclei and cytoplasm, the potential function is obtained from the gradient of I. Therefore, the edges of the gradient have high magnitudes for the blue-to-green transitions and allow better determination of the watershed. The watershed is performed on the gradient image of I with the I_M markers. In this manner the nuclear regions of isolated, clustered or overlapping nuclei (Figure 6C) are obtained. We found that the extraction of nuclei is more accurate with the use of color attributes. A segmentation based only on gray-level RGB images could not really achieve accurate nuclear regions in all cases because the intensity is not accurate enough for both pale and dark nuclei. The use of different gray-level channels clearly emphasizes the distinction between nuclei and cytoplasm.

Nuclear Separation

Certain arrangements of cell nuclei have high complexity; this includes touching and overlapping nuclei and nuclei that are so close together that even a

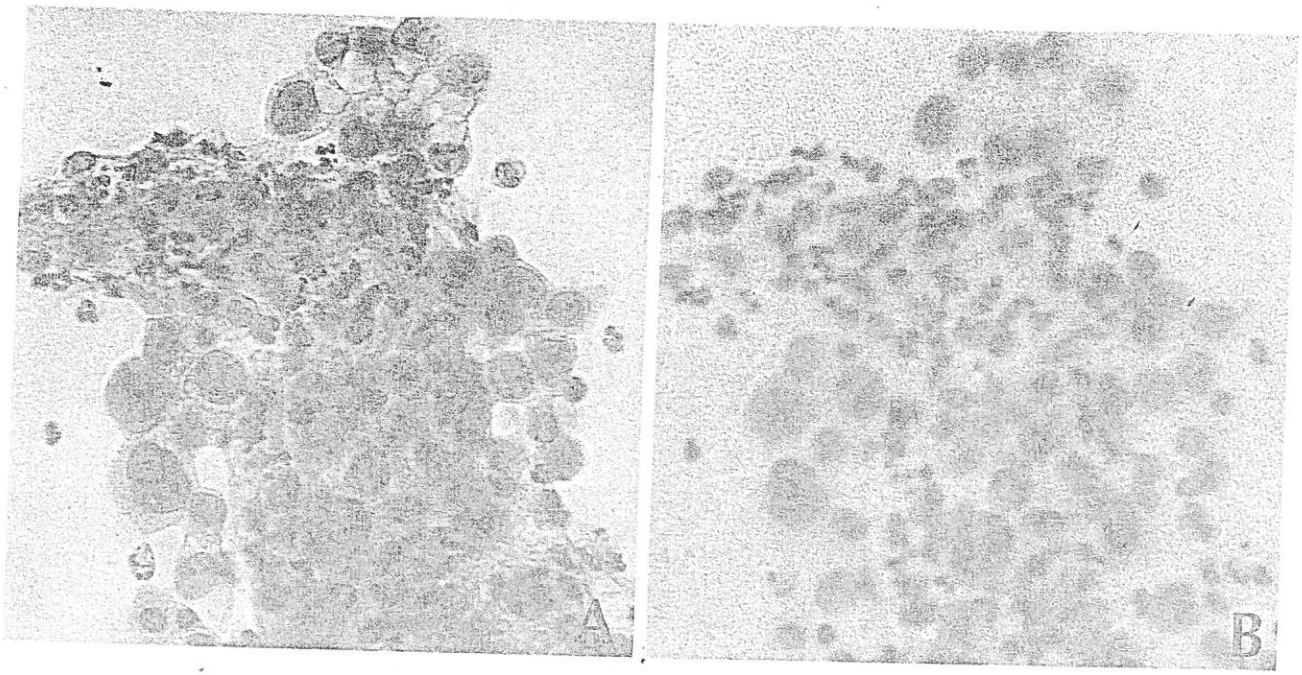


Figure 7 (A) Original image. (B) I_M corresponding image. It shows clearly a distinction between the nuclei and the rest of the image. The brightest nuclei are emphasized.

human observer would have difficulty segmenting them. Hence, only the nuclei that could be visually separated are actually separated. They can be recognized on the basis of concavities in the nuclear boundaries. The nuclear separation method uses those concavities to dissociate touching cells. For instance, let us take two touching cells and compute a distance function from the boundary for each point of the nucleus. When this function is visualized, maxima are obtained at the center of each distinct nucleus (Figure 8C); minima are the points at concavities of touching nuclei. One simple method of separating these touching nuclei is to use a watershed on the opposite of the distance function, using, as markers, the ultimate erode set of the distance function. With this method the different nuclei tend to be separated at the narrows. An example of separation is given in Figure 8. By this method a noticeable limit emerges where the cells touch one another (Figure 6D). The method works for all the types of nuclei described above. In other cases, nuclei cannot be separated because of the overlapping of multiple nuclei.

At this step all the nuclei are delineated; they are either isolated or overlapping clusters.

Cytoplasm Separation

In the final step of the segmentation strategy, the cytoplasm has to be separated so as to contain only one nucleus or cluster of nuclei. This is not always the case: touching cytoplasm can be found with several nuclei inside. The lines of separation are the points where the cytoplasms touch one another (Figure 9A). The major problem is to consider those lines for the separation of the cytoplasm. A gray-level method based on one color channel can be used but is problematic. When cytoplasm presents a cluster of nuclei with no visible line of separation between them, the separation cannot be satisfactory unless the position of each nucleus in the cytoplasm is used. Therefore, the separation method must take this into account. The way to achieve this is to use a watershed on one color channel with nuclear markers. To perform the watershed, the smoothed blue channel image is used as a potential function, and the nuclei (obtained from the nuclear separation step) are used as markers. This method has two advantages. The use of the nuclei as markers takes their spatial configuration and position into account in one cytoplasm. However, the use of the smoothed blue channel image enables considera-

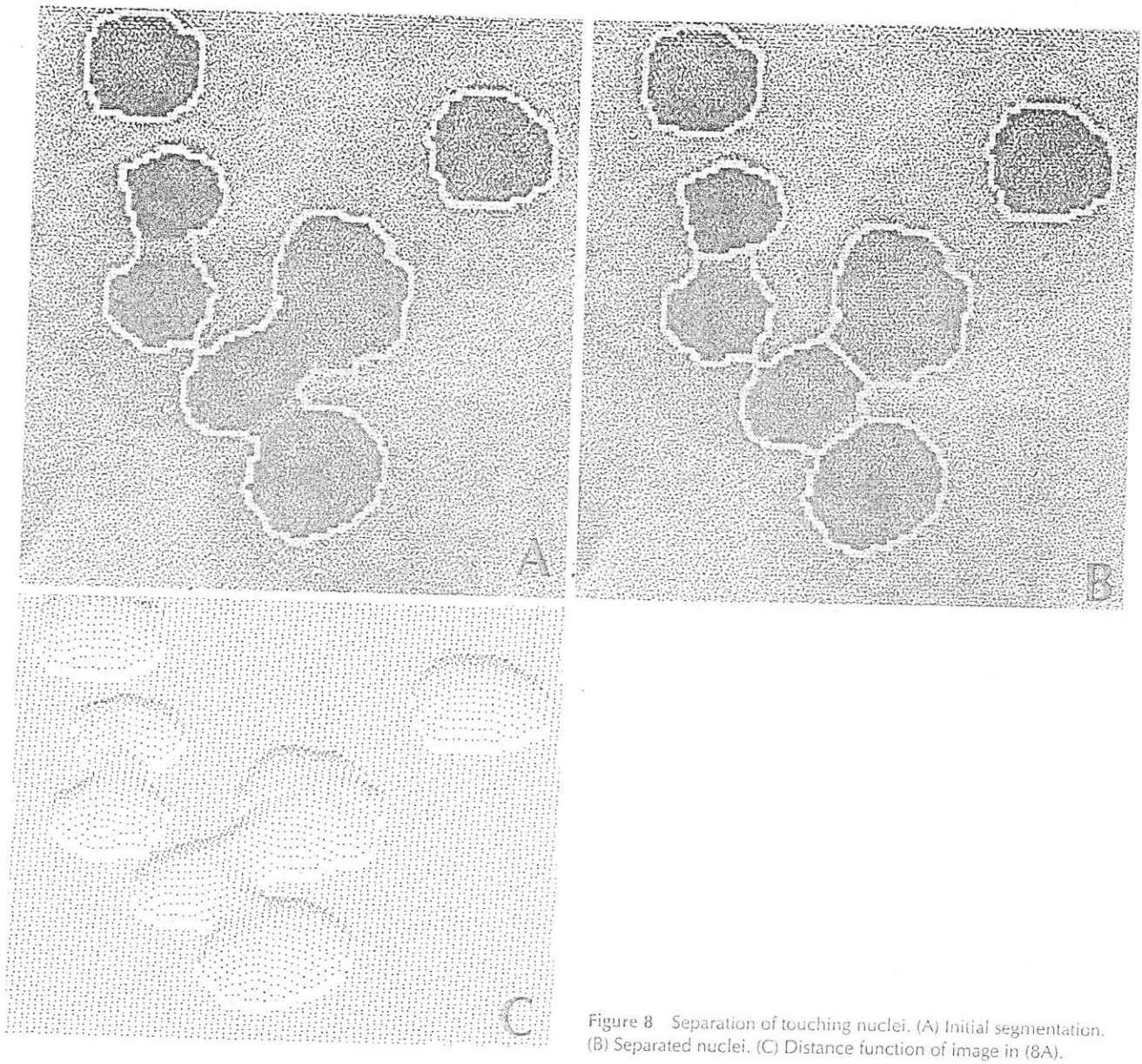


Figure 8 Separation of touching nuclei. (A) Initial segmentation. (B) Separated nuclei. (C) Distance function of image in (8A).

tion of the shape of the cytoplasm and of the lines of separation between touching cytoplasm. Separation of cytoplasm permits separation of cytoplasm with several nuclei inside for two configurations:

- Cytoplasm with concavities in their boundaries or with visible lines of separation between the touching cytoplasms.
- Cytoplasm without any semantic information related to their shapes or lines of separation (Figures 6 and 9).

This step ends the segmentation strategy. With this reliable and systematic method, all the distinct or clustered nuclei are accurately separated so as to be enclosed by one cytoplasm only. An example of all the distinct steps in the segmentation process is given in Figure 6.

Data Collection

To test the segmentation strategy of color images, approximately 2,000 cells from serous cytology slides were used. These slides were obtained from the Anatomy and Cytopathology Laboratory of the Louis Pasteur Hospital Center. The classification of the slides was known in advance since they had

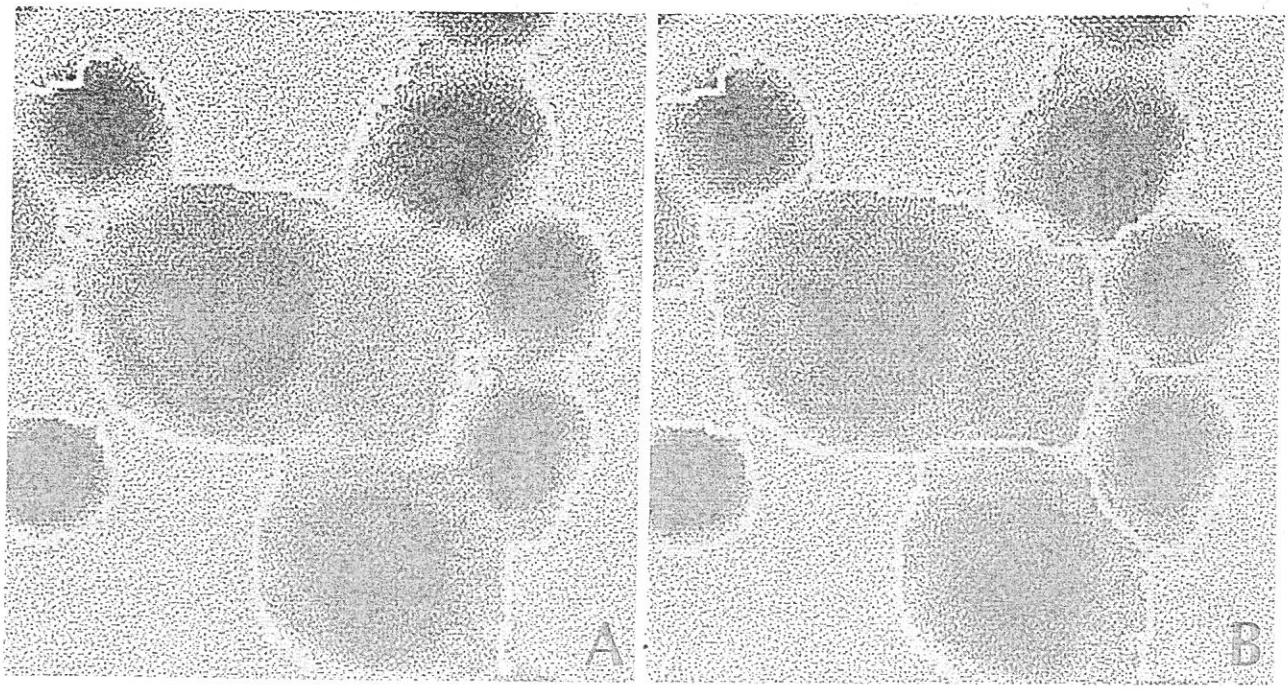


Figure 9 Example of cytoplasm separation. (A) Cytoplasm of touching cells. (B) Cytoplasm after separation into corresponding touching cytoplasm.

been prepared and screened by a pathologist. The images of the cells were taken at random from the slides. Each image was manually focused to obtain the greatest contrast in the image on the monitor (although automated focusing would have produced a better focus). Color images are acquired and processed by the ARCTIC segmentation program. The detected cells were overlaid on the original images and displayed on the monitor for visual inspection.

Visual Inspection

Visual inspection is a subjective method of identifying cell segmentation errors but is simple and easy to implement. A more-objective method of assessing the accuracy of a segmentation process is to determine manually the boundaries of all the cells (cytoplasm and nucleus) and to compare the results with segmentation by counting the misclassified pixels. This method is very tedious and time-consuming, so the pathologist visually inspects the boundaries of the cells overlaid on the monitor and chooses the imprecisely segmented cells or cells missed by the process only. For these cells he determines the correct boundaries for either the nuclei, cytoplasm or both. This method is used to determine the segmentation success rate. The percentage of correctly segmented cells is then obtained. All segmented cells, even clustered or overlapping ones, are examined by the pathologist. However, when an expert cannot determine the correct boundaries of cells because of too much overlap-

Table I *Segmentation Rate for Nucleated Cells*

Expert no.	Correct nuclei (%)	Correct cytoplasm (%)	Nuclear segmentation errors (%)	Cytoplasm segmentation errors (%)
1	96.3	97.1	3.7	2.9
2	94.1	89.5	5.9	10.5
3	87.2	86.7	12.8	13.3
Average	92.5	91.1	7.5	8.9

Table II *Rate of Segmentation Errors According to the Category of the Error*

Expert no.	Undersegmented nuclei (%)	Oversegmented nuclei (%)	Undersegmented cytoplasm (%)	Oversegmented cytoplasm (%)
1	12.5	87.5	26.9	73.1
2	19.1	80.9	37.2	62.2
3	27.3	72.7	35.1	64.9

ping, those cells are marked by the expert as nonsignificant cells. Those cells are not used in the evaluation of the segmentation rate.

Results

In the experiment described in this paper, segmented cells were inspected by three experts. All nucleated cells were detected correctly regardless of their spatial configuration (isolated, just touching, clustered or overlapping). Neither debris nor erythrocytes were retained in the examined images taken at random. All the final objects segmented by the segmentation strategy were cells. Besides the detection of cells, the important criterion in analyzing images is the accuracy in defining the regions of the cells. The correctness of segmentation is crucial in this process since the rest of the analysis is based on the defined regions. By visual inspection, the segmentation errors were divided into four categories: cytoplasm too large, nuclei too large, cytoplasm too small and nuclei too small. Larger errors are defined as oversegmentation and smaller errors as undersegmentation. The results of the segmentation process for each expert are presented in Table I. The success rate in segmenting the cells ranges from 87.2% to 96.3% for nuclei and from 86.7% to 97.1% for cytoplasm. The results can vary from one expert to another (nearly 10% difference). According to the difficulty in estimating minimal errors and ambiguous cases, this difference is considered unavoidable. Although visual inspection leads to a more-objective measure of the success rate in segmenting cells, the subjectivity in detecting cells by an expert remains. These limitations are similar to those encountered in manual screening.

Table III *Rate of Errors in the Segmentation of Nuclei*

Expert no.	Separation errors (%)	Other errors (%)
1	48.6	51.4
2	58.8	41.2
3	42.5	57.5

To understand the difference between the four types of error, the rate of imprecisely segmented cells is presented according to the type of segmentation error. The errors are mainly oversegmentation of cytoplasm and nuclei and can be verified, whoever the expert (Table II).

Oversegmentation of nuclei is chiefly due to nuclear separation, which is designed to divide only a certain type of just-touching cells: those having concavities in their nuclear boundaries. The problem with this method is that just-touching cells, having few or no concavities in their nuclear boundaries, are not separated. Approximately half the oversegmented nuclei come from this problem (Table III). One way to improve nuclear separation is to use a more-accurate distance function. In this study, a 3×3 Chamfer distance is used, but this was not accurate enough to avoid oversegmentation. Another reason for oversegmentation or undersegmentation is the lack of contrast between the nuclei and cytoplasm or imperfect focus; those errors of segmentation can be avoided by using an automatic focus.

Oversegmentation of cytoplasm is directly related to oversegmentation of nuclei: when touching nuclei without any concavities in their boundaries have not been separated, their cytoplasm remains the same. These errors can be avoided only if oversegmentation of nuclei was previously avoided as well. The second reason for oversegmentation is related to nearby mucus: since it has the same color and texture as cytoplasm, it is included in the segmented cytoplasm.

Undersegmentation of cytoplasm is related to two main configurations: nonlinearity in the definition of the cytoplasm (holes or tearing of part of it) or a cluster of overlapping cytoplasm without superposition of nuclei. In the last case, separation of the cytoplasm profits from the nuclear distribution and not from the lines of separation in the cytoplasm since they are hidden by the overlapping. The expert extrapolates because he is able to focus on different points of view.

Discussion

Systems for automated cell classification offer a powerful technique for isolating and recognizing cells in serous cytology slides. The use of color and mathematical morphology operators was demonstrated in this paper. Color is convenient for segmenting cytologic images. On the one hand, a simple thresholding operation is sufficient for extracting all the erythrocytes. On the other hand, the images obtained from different color spaces are useful for the segmentation of cytoplasm and nuclei. The method uses all these possibilities and couples them with a powerful mathematical morphology tool, watershed. The complete strategy of segmentation successfully detects all cells except erythrocytes.

Most segmentation errors are due to touching nuclei with no concavities in their boundaries. These errors can be corrected by using a better distance function. Other errors are related mainly to misfocusing the cells. As the image becomes less focused, the boundaries of nuclei and cytoplasm are blurred and lead to errors in segmenting the cells. Introducing an autofocusing algorithm will definitely improve the accuracy of the exact focus positioning and therefore segmentation of the cells.

To conclude, the success rate of segmentation depends highly on the cell population of the slide. For isolated or touching cells, the strategy is very efficient. For more complex configurations, the success rate cannot be evaluated since an expert cannot segment the cells in an accurate and reliable way. However, the system allows one to separate some of those complex clusters (Figure 6F). We are now studying a way to extract cells from these clusters.

References

1. Beucher S: Segmentation d'images et morphologie mathématique. PhD thesis, Ecole Nationale Supérieure des Mines, 1990
2. Clouard R, Elmoataz A, Angot F, Lezoray O: Pandore: Une bibliothèque et un environnement de programmation d'opérateurs de traitement d'images. Internal report of the GREYC. Caen, France, GREYC, 1997
3. Garbay C: Image structure representation and processing: A discussion of some methods in cytology. IEEE Trans PAMI 1986;8:140–146
4. Harms H, Günzer U, Aus HM: Combined local color and texture analysis of stained cells. Comput Vision Graph Image Processing 1986;33:364–376
5. Knesel EA Jr: Roche Image Analysis System. Acta Cytol 1996;40:60–66
6. Koss LG: The Papanicolaou test for cervical cancer detection: A triumph and a tragedy. JAMA 1989;261:737–743
7. Luong QT: La couleur en vision par ordinateur: 1. Une revue. Institut National de la Recherche en Informatique et Automatique, 1990
8. Papanicolaou GN: A new procedure for staining vaginal smears. Science 1942;95:432–438
9. Patten SF Jr, Lee SSJ, Nelson AC : NeoPath AutoPap 300 Automatic Pap Screener System. Acta Cytol 1996;40:45–52
10. Pauli H: Proposed extension to the CIE recommendations on uniform color spaces, color difference equations and metric color terms. J Opt Soc Am 1975;66:866–867
11. Revision of Laboratory requisitions for medicine, Medicaid and clinical laboratories improvement act of 1967, CLIA'67 programs. Fed Reg 1967;5:9600–9610
12. Rosenthal DL, Acosta D, Peters RK: Computer-assisted rescreening of clinically important false negative cervical smears using the PAPNET testing system. Acta Cytol 1996;40:120–126
13. Tanaka N, Ikeda H, Ueno T, Mukawa A, Watanabe S, Okamoto K, Hosoi S, Tsunekawa S: Fundamental study of cyto-screening for uterine cancer: New system for automated apparatus (CYBEST) utilizing pattern recognition method. Acta Cytol 1977;21:72–85
14. Tolles WE: The Cytoanalyser: An example of physics in medical research. Trans NY Acad Sci 1955;17:250–259
15. Wyszecki G, Styles WS: Color Science: Concepts and Methods, Quantitative Data and Formulae. Second edition. New York, John Wiley & Sons, 1982