Morphological Feature Extraction for the Classification of Digital Images of Cancerous Tissues

Jean-Philippe Thiran,* *Student Member, IEEE*, and Benoît Macq, *Member, IEEE*

Abstract— **This paper presents a new method for automatic recognition of cancerous tissues from an image of a microscopic section. Based on the shape and the size analysis of the observed cells, this method provides the physician with nonsubjective numerical values for four criteria of malignancy. This automatic approach is based on mathematical morphology, and more specifically on the use of Geodesy. This technique is used first to remove the background noise from the image and then to operate a segmentation of the nuclei of the cells and an analysis of their shape, their size and their texture. From the values of the extracted criteria, an automatic classification of the image (cancerous or not) is finally operated.**

I. Introduction

OBJECTIVE analysis of microscopic images of cells and tissues has been a goal of human pathology and cytology since the middle of the 19th century. Early work in this area consisted of simple manual measurements of cell and nuclear size, followed by calculation of cell and nuclear volume. But it is only in the past 30 or 40 years that quantitative cytopathology has become established as a useful method for detection of malignant and premalignant conditions, as well as for the diagnosis of infections, hormonal imbalances, and many other disorders. Historical developments of this research can be found in [1]–[3].

The recent spectacular developments in objective analysis of cells, cell components, and histologic tissue sections relate to the developments in computer sciences that allowed for automation of many functions previously requiring manual calculations. There are today different approaches to automated, computer-based methods of cell and tissue analysis, for instance flow cytometry [4], or image analysis, including morphometry [5] and texture analysis. This paper is related to this last category: image analysis. Many studies have been performed to identify and analyze isolated cells according to parameters measured by image analysing systems, such as cell size, cell shape, density of staining, color hue, etc. [6], [7]. However, for aggregate cells, problems remained for image segmentation and shape description and analysis. Some global parameters such as the cell occupied surface [8], [9] have been used in this case for approximate estimation of cell size and cell density, but these methods do not provide precise morphological information for individual cells.

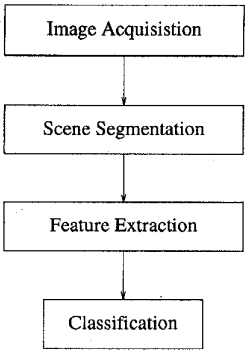


Fig. 1. General structure of the method.

In this paper, we present a new automatic method for the recognition of cancerous tissues, based on morphological segmentation and shape analysis of cells, even if they overlap one another. Our approach can be divided into several well-defined stages, presented in Fig. 1. After acquisition, the image is first segmented in order to isolate the interesting parts and remove noise and undesired components. Next, the feature extraction process is applied, to extract the useful information from the segmented objects, and finally the classification can be operated according to the characteristics extracted by the previous stage. Each of these steps will be presented in the next sections.

Before going into the details of the method, Section II will give some information about the major cytologic characteristics of normal and malignant cells, in order to specify the criteria of malignancy to consider. The acquisition, segmentation, feature extraction and classification tools developed will be described in Section III. Section IV will present and evaluate experimental results and Section V will discuss some specific points about this method. Finally, some conclusions will be drawn in the Section VI.

II. Cytologic Features of Normal and Malignant Cells

This section deals with the basic concepts of cytopathology. It is not the aim of this paper to expose in details the general pathology and cytology of benign and malignant tumors (more details can be found in [10]). The emphasis is more on the main morphological features characterizing cells derived from

Manuscript received July 15, 1994; revised May 2, 1996. The work of J.-P. Thiran was supported by the Belgian Fonds pour la formation à la Recherche dans l'Industrie et dans l'Agriculture (F.R.I.A.) under Grant Nr 930518. *Asterisk indicates corresponding author.*

*J.-P. Thiran is with the Laboratoire de Télécommunications et Télédétection of the Université Catholique de Louvain, B-1348 Louvain-la-Neuve, Belgium (e-mail: thiran@tele.ucl.ac.be).

B. Macq is with the Laboratoire de Télécommunications et Télédétection of the Université Catholique de Louvain, B-1348 Louvain-la-Neuve, Belgium.

Publisher Item Identifier S 0018-9294(96)07258-8.

normal and malignant tissues, in order to define some criteria of malignancy to be used in an automatic approach of the recognition process.

In contrast to normal and benign cells, which are typically uniform in appearance, malignant cells are characterized by *irregular morphology* that is reflected in several parameters. Most of the criteria of malignancy are seen in the nuclei of the cells. However, valuable information can be obtained from the cytoplasm. The nuclei of malignant cells are larger than those of benign cells, vary in size (*anisonucleosis*) and vary in shape (*nuclear deformity*). Within cells of the same type and age, the degree of nuclear enlargement is different from one cell to another. The metabolism of a cancer cell is geared predominantly to reproduction and duplication of nuclear material, rather than to cytoplasmic differentiation and function. Consequently, malignant cells tend to have scant cytoplasm. As the nuclear enlargement is not accompanied by similar enlargement of the cytoplasm, the ratio of the area of the nuclei to that of the cytoplasm of the cells, usually expressed as the *nucleocytoplasmic ratio*, is increased. In the internal structure of the nuclei, the chromatin is irregularly distributed, with areas of clumping and condensation (heterochromatin), whereas other areas show chromatin clearing (euchromatin). Consequently, the chromatin exhibits coarse granularity and irregular clumping, called *hyperchromasia*.

According to these considerations, it is possible to define several criteria of malignancy of a tissue. In this paper, the following criteria will be studied: *nucleocytoplasmic ratio, anisonucleosis, nuclear deformity, and hyperchromasia*. Fig. 2 shows the image of a malignant pulmonary tissue where these four criteria are visible. A general method for the extraction of a numerical value for each of these criteria of malignancy will be explained in the next section. With these values, it will be possible to build a classifier to automatically propose a diagnosis, in order to help the physician to decide whether the observed tissue is cancerous or not.

III. Materials and Methods

A. Origin and Acquisition of the Images

In this study, every image is obtained by biopsy, using the same tissue processing and acquisition method, briefly described below. We have tested the method presented in this article on a experimental set of 83 images, coming from two different organs: lungs and digestive tract.

Each biopsy specimen is transferred from the biopsy forceps to a glass slide with a fine needle and then delicately rolled onto the slide so that an imprint of all sides of the specimen is obtained. This technique is called *touch cytology* [11] and provides samples that are richer in cells than those obtained with aspiration or brushing. The samples are air-dried. A rapid fixation and staining method is used ($M + D +$ Quick, Merz-Dade AG, Dudingen, Switzerland). This method is similar to the classic May-Grünwald-Giemsa stain and takes 15 s. The slide is dipped for 5 s in three staining solutions, rinsed with water, and allowed to air-dry.

The technical equipment used for the image acquisition included a high quality optical microscope, a high-resolution CCD B/W camera and an acquisition hardware/software system, running on an UNIX workstation. Every image was acquired with 256 grey levels, with a magnifying factor of 60 on the microscope, under immersion. Manual corrections were operated on the contrast and intensity in such a way that the digitized image became visually acceptable for further manual classification by a specialist. No other filtering was operated. The final resolution of the images was 0.16 μm per pixel. Notice that in this range of magnification, our shape analysis process is independent of the magnification factor, since all the measured values will be normalized according to the size of the objects. This point will be detailed later.

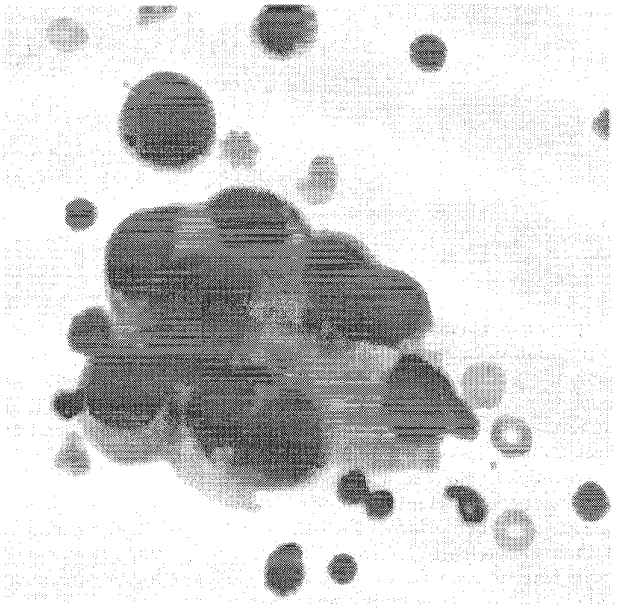


Fig. 2. Image of the microscopic section of a malignant pulmonary tissue. Four criteria of malignancy are visible: large nucleocytoplasmic ratio, anisonucleosis, nuclear deformity, and hyperchromasia.

B. Segmentation of the Nuclei

As indicated above, the major features of malignancy are related with the nuclei of the cells. It is therefore essential to operate a segmentation of the image, to isolate the nuclei from the rest of the image, i.e., from the cytoplasm, and from some other undesirable elements coming from the cytopreparation.

When looking at a typical section image (Fig. 2), it can be seen that the nuclei are large dark spots, when the cytoplasm is globally light, but with some small dark dots, excluding a simple thresholding to separate the nuclei from the rest of the cells. Moreover, there are other entities than just the nuclei and the cytoplasm, such as white or red corpuscles, or elements coming from the preparation of the microscopic section.

Several segmentation methods have been proposed in the literature, such as edge detection, hierarchic thresholding or two-dimensional (2-D) histogram threshold (see for example [12]–[15]), but they do not take a specific characteristic of the nuclei into account: their shape. We propose an other approach, better suited for shape analysis, which is mainly based on mathematical morphology. Morphological tools are first developed to remove the background noise from the image

and to operate a segmentation of the nuclei of the cells, next to analyze the shape and the size of these nuclei and then to evaluate their texture. In the next subsections, the main morphological operators will be briefly introduced, followed by a description of the segmentation process of the nuclei. In the next sections, the method for extracting a numerical value for the four concerned criteria of malignancy will be presented.

1) Mathematical Morphology: Mathematical morphology provides an approach to the processing of digital images which is based on shape. Appropriately used, Mathematical morphological operations tend to extract their essential shape characteristics and to eliminate irrelevancies [16]. The language of mathematical morphology is that of set theory [17]. Sets in mathematical morphology represent the shapes which are manifested on binary or grey tone images. The set of all the black pixels in a black and white image (a binary image) constitutes a complete description of that binary image. Sets in Euclidean 2-space denote foreground regions in binary images. Sets in Euclidean 3-space denote grey scale images. In this paper, we will only describe the grey scale morphology, since all the treated images are grey scale.

The four basic operations of mathematical morphology are dilation, erosion, opening and closing. Another useful operator is the morphological reconstruction. We will briefly define each of these operators hereafter. As these notions are generally well known, we will not go into the details of the definitions. A more complete theory of mathematical morphology can be found in [16]–[19].

The dilation of an image $f: F \rightarrow C$ by a structuring element $k: K \rightarrow C$, denoted $f \oplus k$, is defined at point x by $(f \oplus k)(x) = \max_{\substack{z \in K \\ x-z \in F}} \{f(x-z) + k(z)\}$.

The erosion of an image $f: F \rightarrow C$ by a structuring element $k: K \rightarrow C$, denoted $f \ominus k$, is defined at point x by $(f \ominus k)(x) = \min_{\substack{z \in K \\ x+z \in F}} \{f(x+z) - k(z)\}$.

The mathematical properties of these two operators are extensively presented in [16], as well as those of the following operators. Let us note that in many applications the structuring element is chosen flat, i.e., $k(z) = 0 \ \forall z \in K$.

In practice, dilations and erosions are usually employed in pairs, either dilation of an image followed by the erosion of the dilated result, or image erosion followed by dilation. These operations are respectively called closing and opening. In either case, the result of iteratively applied dilations and erosions is elimination of specific image details smaller than the structuring element used. For example, opening an image with a disk structuring element smooths the contours, breaks narrow isthmuses, and eliminates small islands and sharp peaks or capes. Closing an image with a disk structuring element smooths the contours, fuses narrow breaks and long thin gulfs, eliminates small holes, and fills gaps on the contour. The opening of an image f by a structuring element k, denoted $f \circ k$, is defined by $f \circ k = (f \ominus k) \oplus k$. The closing of an image f by a structuring element k, denoted $f \bullet k$, is defined by $f \bullet k = (f \oplus k) \ominus k$.

Of practical significance is the fact that image transformations employing iteratively applied dilations and erosions are *idempotent*, that is, their reapplication effects no further changes to the previously transformed result. Idempotence can be considered as a criterion for the accuracy of the representation and transcription of visual material. Moreover, the functionality of morphological closings and openings corresponds closely to the specifications of a signal by its bandwidth. Morphologically filtering an image by an opening or closing operation corresponds to the ideal nonrealizable bandpass filters of conventional linear filtering: further ideal filtering does not alter the result.

Fig. 3 illustrates the opening of a three-dimensional (3-D) function by a square structuring element. On this figure we can see the morphological filtering effect: all the objects smaller than the structuring element have been eliminated, while the large objects still remain. However, their shape has been altered, clearly showing the shape of the structuring element. The next operator, called *reconstruction*, will solve this problem.

As we have seen, the shape of the objects in an image is altered by an opening or by a closing. The aim of the reconstruction is to recover the original shape of these objects from the opened or closed images, while keeping the morphological filtering effect. The notion of reconstruction is closely related to the recent notion of Morphological Geodesy, first introduced by Lantuéjoul [20]–[22]. Let us first define a geodesic dilation.

The *1-size geodesic dilation* of an image f by a structuring element k, conditionally to an image Y, denoted $(f \oplus k)^{(1)}$, is defined by

$$(f \oplus k)^{(1)} = (f \oplus k) \cap Y. \tag{1}$$

The image Y is called the *geodesic mask* and k is a usually very small (for instance a 3×3 pixels square in a 8-connectivity digital space or a five-pixels rhomboid in a 4-connectivity digital space). When handling with grey scale images, the intersection operator is to be understood as the *minimum* operator.

The *n-size geodesic dilation* of an image f by a structuring element k, conditionally to an image Y, denoted $(f \oplus k)^{(n)}$, is defined by

$$(f \oplus k)^{(n)} = \underbrace{(((f \oplus k)^{(1)} \oplus k)^{(1)} \oplus \cdots \oplus k)^{(1)}}_{n \text{ times}}. \tag{2}$$

With these definitions, it is now possible to give a strict definition of the reconstruction.

The *reconstruction of an image f, conditionally to an image Y*, is the geodesic dilation until idempotence. Let i be the value from which the idempotence is obtained. Hence, the reconstructed image, denoted R_Y, is defined by

$$R_Y = (f \oplus k)^{(i)}, \quad \text{with } (f \oplus k)^{(i)} = (f \oplus k)^{(i-1)}. \tag{3}$$

As indicated above, when the original image is taken as the geodesic mask Y, the reconstruction allows to recover the exact shape of the objects, keeping the morphological filtering effect, as illustrated in Fig. 4.

These morphological operators will be the basic tools used in the method built to recognize cancerous cells in the image of a microscopic section. The next section describes this method.

2) Segmentation: In this paper, the following convention is adopted: as the background of the image is light and the

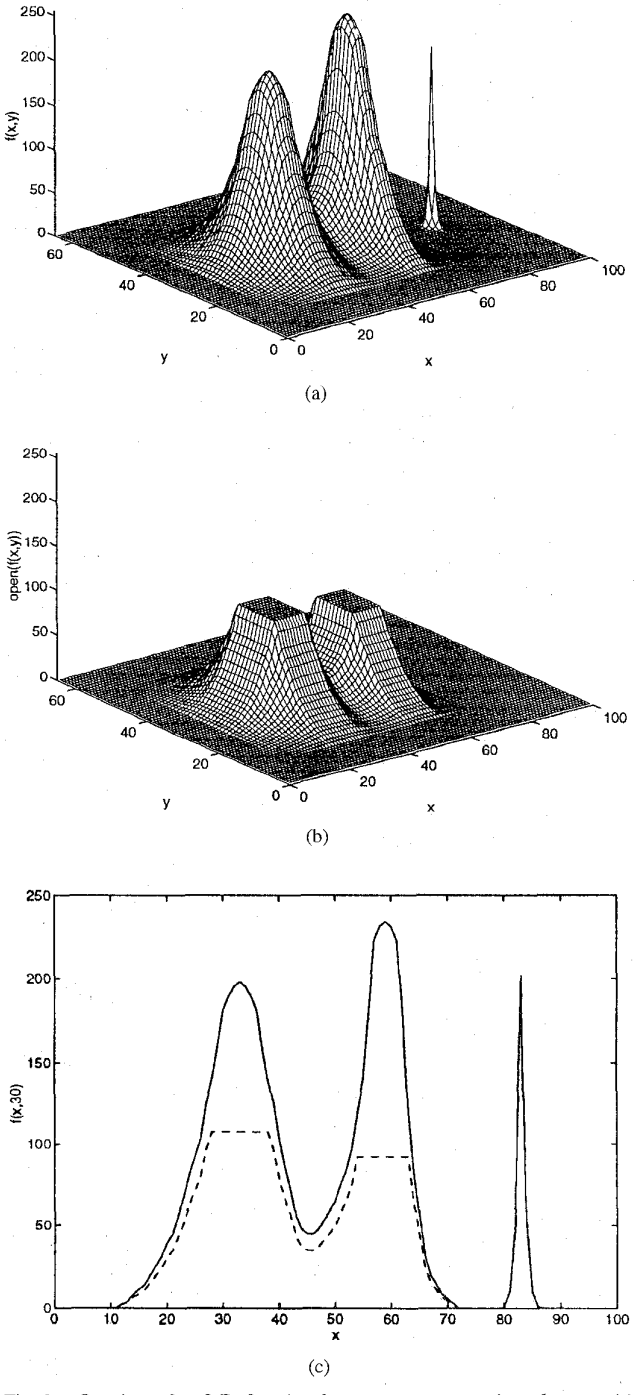


Fig. 3. Opening of a 3-D function by a square structuring element: (a) original function, (b) opened function, (c) cross sections along the plan $y = 30$: original function (solid line) and opened function (dashed line).

objects are darker, we consider that the grey level is zero for a white pixel and grows to the maximum value for a black pixel. All the following descriptions will be made according to this convention.

The segmentation process is presented in Fig. 5. The first operation is an opening of the image by a large flat square structuring element, close to the size of the nuclei. As indicated above, the result of this operation is a morphological filtering

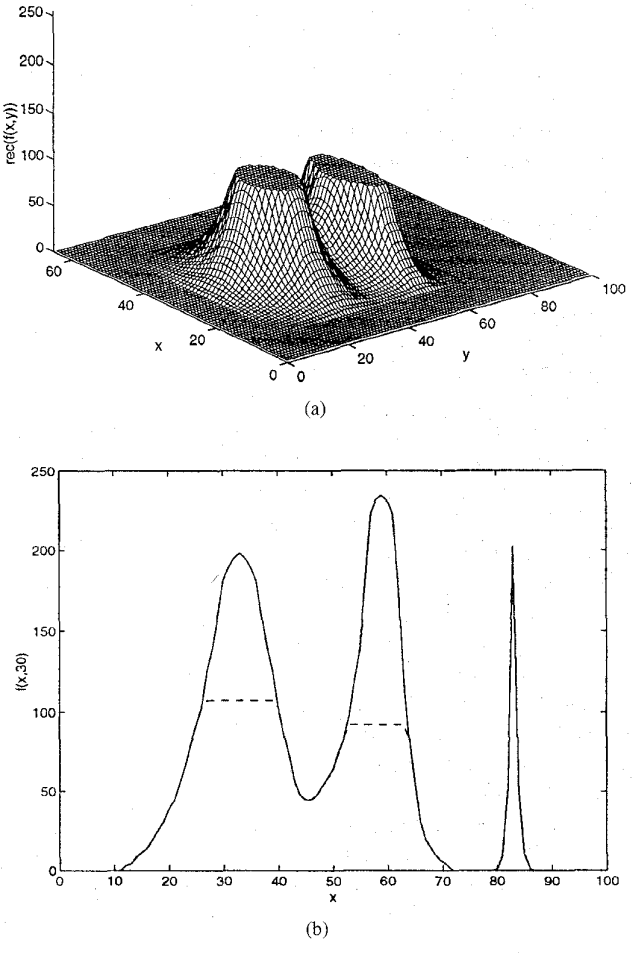


Fig. 4. Reconstruction of the opened function of Fig. 3 conditionally to the original function: (a) three-dimensional reconstructed, (b) cross sections along the plan $y = 30$: original function (solid line) and reconstructed function (dashed line).

effect: all the elements smaller than the structuring element disappear, or are roughly lighted. Thus its size has to be chosen according to the magnification factor used during the acquisition of the image. The only constraint is that it has to be large enough to eliminate all the nonnuclear elements. Fig. 6 shows the result of such an operation on the image of Fig. 2. Furthermore, the shape of the remaining objects has been altered by the opening, and the square shape of the structuring element comes into sight everywhere in the image. Another choice for the shape of the structuring element, such as a circle, should be more suitable, but is not so easy to implement in a digital space, and above all much slower to execute. Moreover, as it will be explained below, the reconstruction operation will solve this problem. The reconstruction of this image conditionally to the original image is the next step. It allows to recover the exact shape of all the remaining objets while keeping the morphological filtering effect. In Fig. 7, showing the result of this operation, we can see that the nuclei are completely reconstructed, and that the rest of the image is uniformly light. The next step is therefore evident: a simple thresholding of the reconstructed image is enough to isolate the

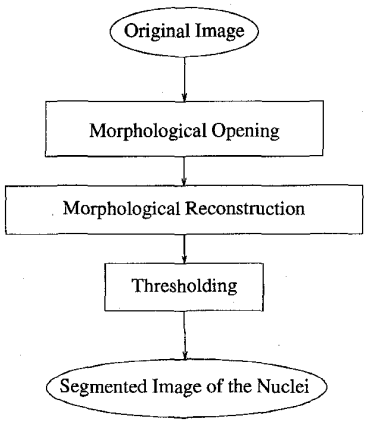


Fig. 5. General description of the segmentation process.

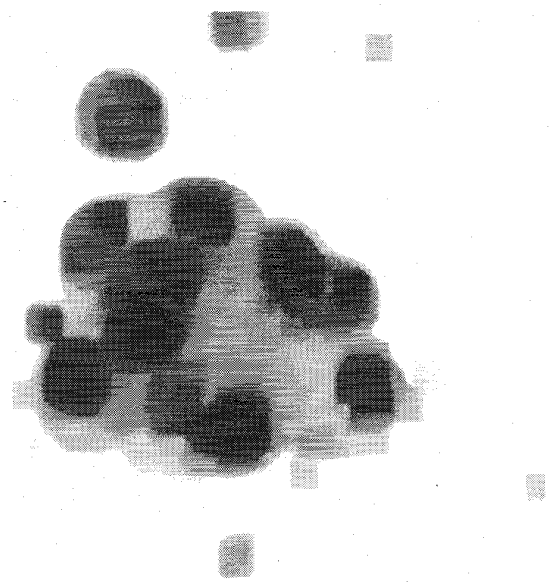


Fig. 6. Result of the opening of the original image of Fig. 2 by a square flat structuring element.

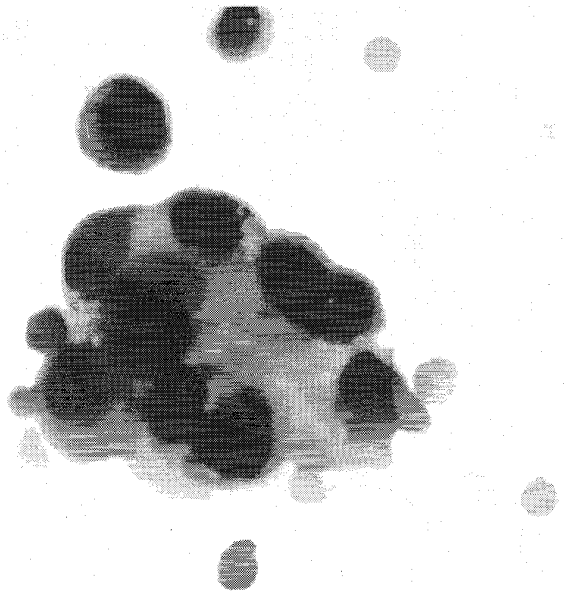


Fig. 7. Result of the reconstruction of the image of Fig. 6 conditionally to the original image of Fig. 2.

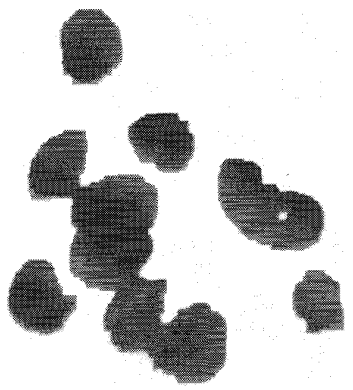


Fig. 8. Result of the thresholding of the reconstructed image: The nuclei are isolated.

nuclei, as shown in Fig. 8. The value of the threshold used here becomes very easy to choose, because of the morphological filtering effect: all the large objects (the nuclei) are uniformly dark, while all the other elements are light. A simple histogram method [13] allows an optimal choice of the threshold value. Let's notice that an opening followed by a reconstruction is equivalent to an erosion followed by the same reconstruction. In this paper, for reasons of clarity, our description was based on the opening because of the well-known morphological filtering effect of this operator.

C. Feature Extraction

1) Nucleocytoplasmic Ratio: The first criterion of malignancy is the value of *the nucleocytoplasmic ratio*: the ratio of the area of the nuclei to that of the cytoplasm. The evaluation of the nuclear area is simply done by counting the number of nonwhite pixels in the image of the nuclei. By this way, we obtain a measurement of the first part of the ratio. The next step will be the extraction of the cytoplasm from the original image. This part is accomplished as follows. We first operate a difference between the original image and the image of the nuclei, to obtain all the nonnuclear elements. Next, a double thresholding is executed, to remove the background light noise and the dark spots. The resulting image is closed by a small structuring element, in order to uniformize the cytoplasmic area. The result of these operations is an image where we assume that only the cytoplasm is present, without the nuclei and other components. The cytoplasmic area is then evaluated by the same way as we did for the nuclear area, by counting the nonwhite pixels. Finally, the *nucleocytoplasmic ratio* is obtained by the ratio of the nuclear area to the cytoplasmic area. We have thus a value for the first criterion of malignancy. The next section will deal with the second and the third criterion: the anisonucleosis and the nuclear deformity.

2) Anisonucleosis and Nuclear Deformity: To obtain a numerical value for these two criteria, the main tool is again mathematical morphology. Morphological operators will be

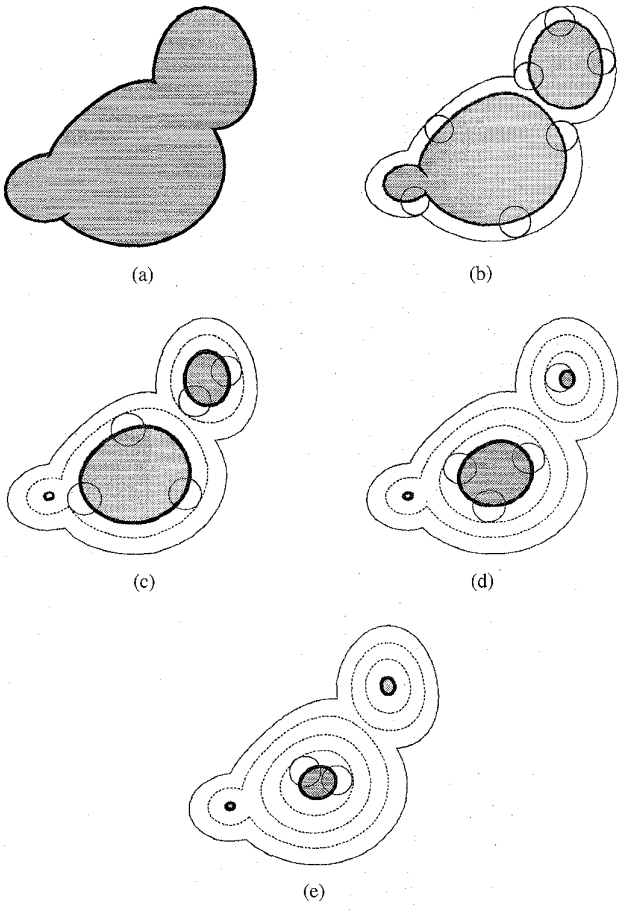


Fig. 9. (a)–(d) Successive erosions and (e) ultimate eroded.

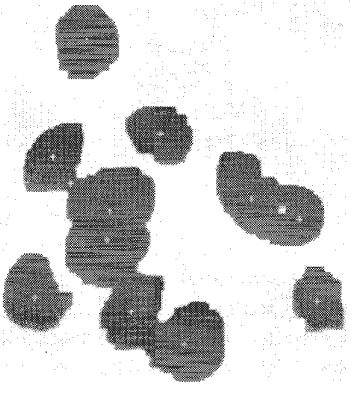


Fig. 10. Detected centers of gravity of the nuclei.

used to find the center of each nucleus, even if they overlap one another. With the position of these centers, the anisonucleosis and the nuclear deformity will be evaluated by calculating the distance from the center to the border of the nucleus, in several well-chosen directions.

The first step is thus the detection of the center of gravity of each nucleus. The problem of the partial superimposition of the nuclei is solved by the use of morphological erosions. The image of the nuclei is binarized and successive binary erosions by a very small structuring element are operated. By this way, the different parts of the image tend to separate from each other, becoming smaller and smaller, as illustrated in Fig. 9. These successive erosions are repeated in order to obtain the "ultimate eroded," which are the last parts of each object present on the image, before disappearing by the next erosion [22]. The center of gravity of each "ultimate eroded" is the center of each nucleus. Fig. 10 shows the image of the nuclei, with a cross on all the centers of gravity detected by this method. From the position of these detected centers, it is possible to evaluate the anisonucleosis and the nuclear deformity. First, we calculate the distance from the center to the border of each nucleus in several directions (8 or 16 cardinal directions for example). From these lengths, only the so-called *characteristic distances* are kept. These are defined as follows. A distance is called *characteristic* if it is smaller than the distance between the center of the considered nucleus and the closest center in a narrow angular sector $d\Theta$ around the considered direction, as shown in Fig. 11. In this figure, distance d_{i1} is not characteristic because it is bigger than the distance between the center cg_i and the center $cg_{i'}$, within the angular sector $d\Theta$. This technique allows to consider only the distances relative to the considered nucleus, and not the distances crossing several superimposed nuclei. With these distances, we evaluate the anisonucleosis by the variance, on all the nuclei, of the mean value of the characteristic distances of each nucleus, reported to a measure of the mean size of the nuclei. In the same way, the nuclear deformity is computed by the mean value, on all the nuclei, of the variance of the characteristic distances of each nucleus, reported to the mean size of the nuclei.

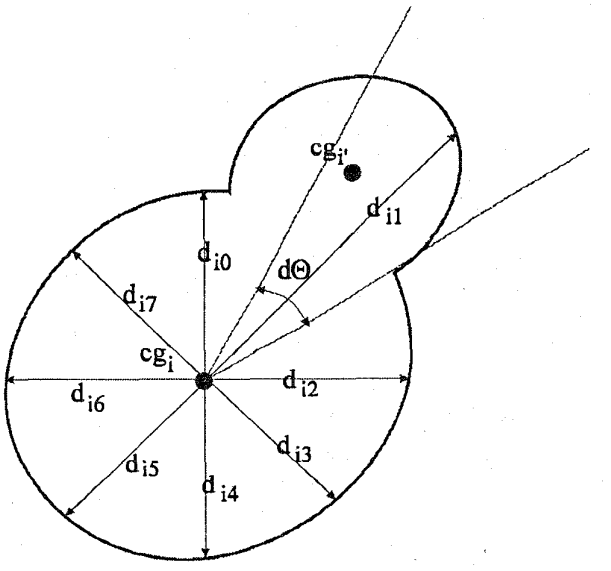


Fig. 11. Example of characteristic and noncharacteristic distances: All distances are characteristic, except d_{i1}.

Let d_{ij} be the characteristic distance of nucleus i in direction j

$$\begin{aligned} &\textit{The mean size of nucleus } i = ms_i = \underset{charact.\,j}{mean}\,(d_{ij}) \\ &\textit{The mean size of the nuclei} = ms = \underset{i}{mean}(ms_i) \\ &\Rightarrow \textit{The anisonucleosis} = \frac{\mathrm{var}_i(ms_i)}{ms^2} \end{aligned} \tag{4}$$

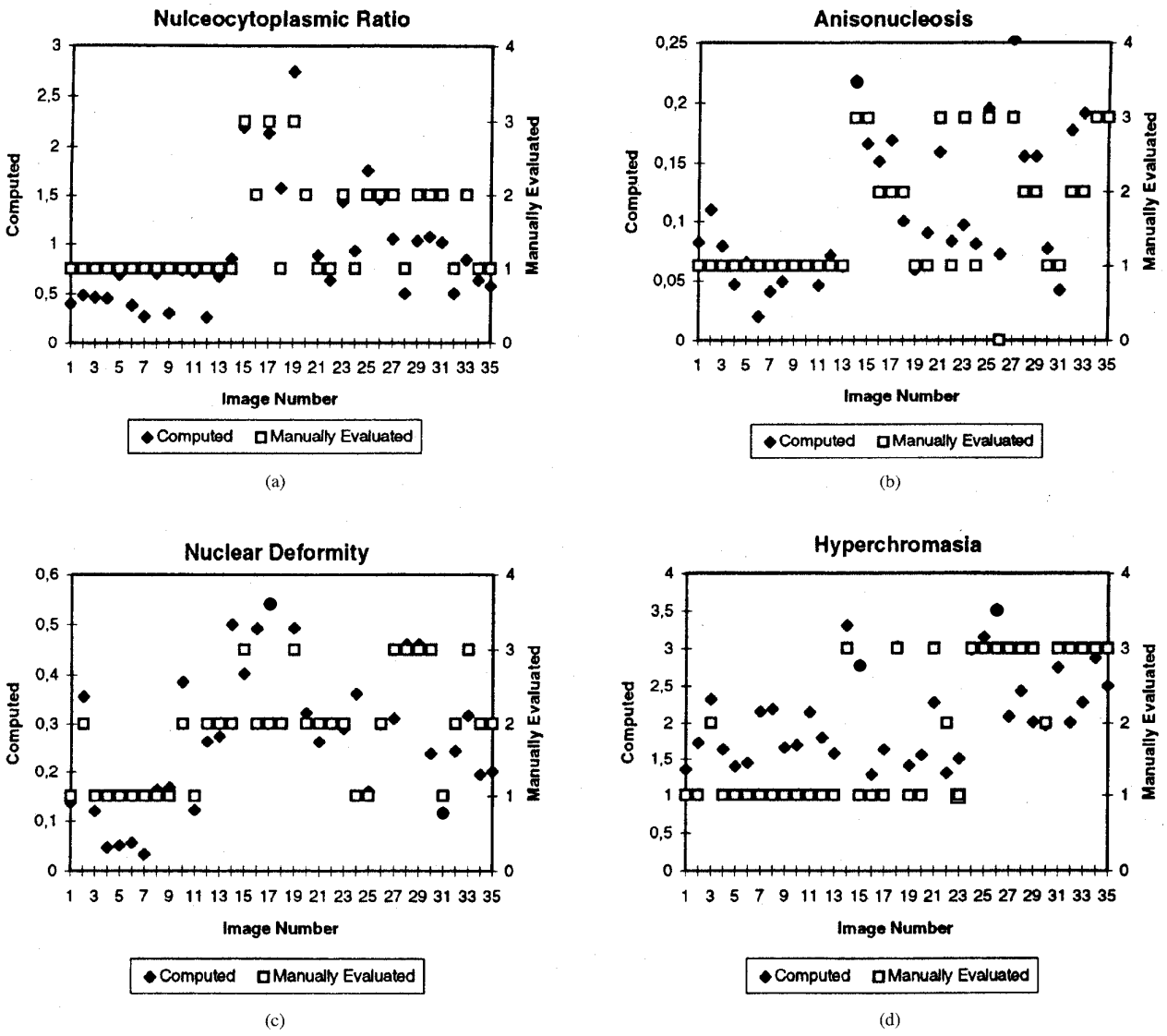


Fig. 12. Comparison between manual and automatic evaluation of the four criteria of malignancy on an experimental set of 35 images of lungs: (a) nucleocytoplasmic ratio, (b) anisonucleosis, (c) nuclear deformity, and (d) hyperchromasia.

$$\textit{The deformity of nucleus } i = df_i = \operatorname*{var}_{charact.j}(d_{ij})$$

$$\Rightarrow \textit{The nuclear deformity} = \frac{mean_i(df_i)}{ms^2}. \tag{5}$$

The last criterion of malignancy to be treated is the *hyperchromasia*. This will be done in the next section.

3) Hyperchromasia: The hyperchromasia can be easily extracted from the image of the nuclei by the well-known Morphological Top Hat Transform [17], [23], [24]. First, a morphological closing is operated, by a small structuring element. The effect is again a morphological filtering, eliminating the granular texture of the nuclei. This closed image is then subtracted from the original image of the nuclei. The result is an image only containing all the small "spikes" in the nuclei. The hyperchromasia is evaluated by quantifying the importance of these spikes in the resulting image, i.e., the mean value of the spikes over the nuclear area.

In summary, a set of tools has been presented in this paper, based on mathematical morphology. A new deformity criterion has been proposed, called *characteristic distances*. These techniques allow to extract a numerical value for the four concerned criteria of malignancy. The last step of this work is now to decide, from these four values, whether the tissue is cancerous or not. This problem related to multi-criteria decisions, is investigated in the next chapter.

D. Classification

From the value of these four criteria, we must classify the analyzed image, in order to decide whether the tissue is cancerous or not. Although it is possible to use sophisticated classificators like Vector Quantization [25]–[27] or Kohonen neural networks [28]–[31], we suggest here a more simple method, based on the evaluation of the *score* of the image, i.e., a number indicating the degree of malignancy or healthiness

of the considered tissue. This score, in percent, denoted S, is calculated by

$$S = \frac{100}{1+l} \tag{6}$$

with

$$l = \frac{1}{4}\left(\left(\frac{C_1}{C_{1m}}\right)^n + \left(\frac{C_2}{C_{2m}}\right)^n + \left(\frac{C_3}{C_{3m}}\right)^n + \left(\frac{C_4}{C_{4m}}\right)^n\right) \tag{7}$$

- C_i The value of the ith criterion of malignancy, higher values indicating a higher degree of malignancy.
- C_{im} The maximum value (threshold) of the ith criterion still indicating nonmalignancy. This value is obtained from the maximum value of the ith criterion on a set of training images of noncancerous tissues.
- n A value to be determined experimentally.

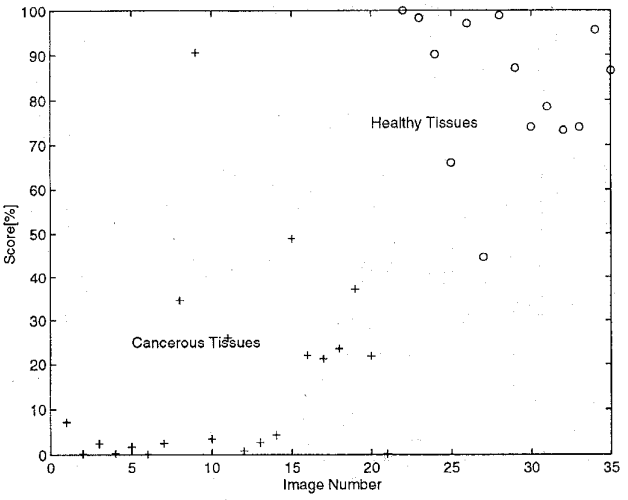
Notice that when one of the C_i is above the corresponding C_{im}, and for large values of n, l tends to huge values.

With this definition, a value of l lower than one indicates a noncancerous tissue, while a value of l larger than one indicates that at least one criterion is larger than the healthiness threshold. The score, S, varies from 0% for very malignant tissues to 100% for healthy cells. A score of 50% is the limit between cancerous and noncancerous tissues. The value of n has to be adapted to minimize the error rate in the proposed diagnostic. This value has to be determined experimentally. A large value allows a good separation of the two regions, but the scores are concentrated near 0% and 100%. A small value for n increases the number of classification errors. Notice that it could be possible to refine the classificator by chosing specific n values for each criterion. The interest of this score is that it provides an indication on the degree of malignancy of the observed tissue.

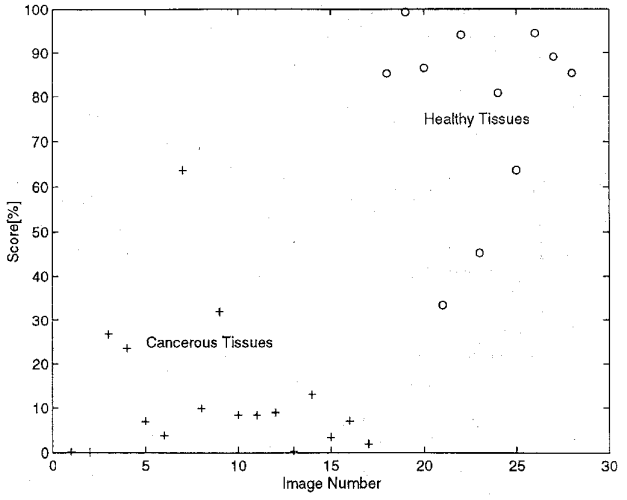
IV. Experimental Results

We have tested the method presented in this article on a experimental set of 83 images, coming from different biopsies and from two different organs: lungs and digestive tract. First, the results of the feature extraction procedure have been compared to the results of a manual evaluation of the different criteria of malignancy. In this study, an experienced specialist has manually evaluated the four criteria by giving a value from 1 to 3 for each of them (1 meaning that the criteria is not present in the image, 2 that it is slightly present, and 3 that it is very important). Comparisons of this evaluation with the automatically extracted values are presented in Fig. 12 for a set of 35 pulmonary images. We can see a good correlation between the manual and the automatic approach. And even if some errors can occur for an individual criterion, we will see that the classification based on the four criteria gives very good results.

For the classification, the values of n and C_{im} have been set according to this first data set of 35 pulmonary images. The value of $n = 7$ has been chosen because this value was the best compromise to minimize the error rate while still keeping a good identification of the degree of malignancy of the tissues.



(a)



(b)

Fig. 13. Experimental results of the classification with $n = 7$: (a) 35 images of lungs and (b) 28 images of digestive tract.

We have then compared the classification obtained by this method with the results of a manual diagnosis operated by a experienced specialist. On the set of 35 images of lungs, with $n = 7$, we only have two classification errors (image 9 and 27). Except for these, all the cancerous images of this set have a score lower than 50% and all the healthy image have a score higher than this limit. These results are presented in Fig. 13(a), where the crosses represent images of cancerous tissues and the circles represent images of noncancerous tissues. We can observe a good discrimination between cancerous and healthy cells.

To evaluate the reproductibility of this method, we have then operated the same classification, with the same values of n and C_{im}, on another set of 20 images of lungs and 28 images of digestive tract. The results show two classification errors in the first case and three classification errors in the second case, i.e., a recognition rate of 90% and 89%. Fig. 13(b) shows this last result.

V. Discussion

The method described here is fully automatic when the different parameters (thresholds) have been set to appropriate values. These values are constant for a given acquisition procedure, when the contrast and intensity corrections have been fixed. For the classification, the determination of the maximum values of each criterion still indicating nonmalignancy is done on a training set of images of noncancerous tissues.

This method has been implemented on classical workstations, and no special attention has been paid to the speed of the algorithms in our implementation. In these conditions, a complete processing, from the acquisition to the classification, takes about one minute.

In this context, it is very easy to extend the method to a classification based on multiple images of the same tissue, as it is normally done by manual methods: the features would be extracted from each image and presented to the classificator, which would have more inputs than described in Section III-D, but the principle of the method would still remain the same.

VI. Conclusion

The problem of shape analysis was investigated in the case of cancer detection. We have elaborated a method to extract from an image of a microscopic section a set of informations relative to the size and the shape of the objects in the image. The main tool used was mathematical morphology. The classification is then operated by calculating the *score*, which represents the degree of healthiness of the tissue. The experimental results show a good efficiency and reproductibility of this method.

References

[1] L. G. Koss, "Analytical and quantitative cytology. A historical perspective," *Anal. Quantum Cytol.*, vol. 4, pp. 251–256, 1982.

[2] ——, "High resolution automated microscopy," *Anal. Quantum Cytol.*, vol. 7, pp. 2–3, 1985.

[3] ——, "Automated cytology and histology. A historical perspective," *Anal. Quantum Cytol.*, vol. 9, pp. 369–374, 1987.

[4] L. G. Koss, B. Czerniak, F. Herz, and R. P. Wersto, "Flow cytometric measurements of DNA and other cell components in human tumors. A critical appraisal," *Human Pathol.*, vol. 20, pp. 528–548, 1989.

[5] J. P. A. Baak and J. Oort, *Morphometry in Diagnostic Pathology*. Berlin: Springer Verlag, 1983.

[6] C. Garbay, J. M. Chassery, and G. Brugal, "An interactive region-growing process for cell image segmentation based on local color similarity and global shape criteria," *Anal. Quantum Cytol. Histol.*, vol. 8, pp. 25–34, 1986.

[7] C. Gauvain, D. Seigneurin, and G. Brugal, "A quantitative analysis of the human bone marrow erythroblastic cell lineage using the SAMBA 200 cell image processor," *Anal. Quantum Cytol. Histol.*, vol. 9, pp. 253–262, 1987.

[8] J. C. Bisconte and S. Margules, "Real-time continous quantitative analysis of cultured living cells," *Mikroskopie*, vol. 37, pp. 204–208, 1980.

[9] P. Travo, P. Bodin, G. Burnstock, and J. C. Stoclet, "Quantitative method for study of morphological changes induced by vasoactive agents in single aortic myocytes grown in primary cultures," *Lab. Invest.*, vol. 56, pp. 335–343, 1987.

[10] L. G. Koss, *Diagnostic Cytology and Its Histopathologic Bases*, 4th ed, vol. 1. Philadelphia, PA: Lippincott, 1992.

[11] J. C. Debongnie, J. Mairesse, M. Donnay, and X. Dekoninck, "Touch cytology, a quick, simple, sensitive screening test in diagnosis of infections of the gastrointestinal mucosa," *Arch. Pathol. Lab. Med.*, vol. 118, pp. 1115–1118, 1994.

[12] R. C. Gonzalez and P. Wintz, *Digital Image Processing*, 2nd ed. Reading, MA: Addision-Wesley, 1987.

[13] M. Sonka, V. Hlavac, and R. Boyle, *Image Processing, Analysis and Machine Vision*. London, U.K.: Chapman & Hall, 1993.

[14] W. K. Pratt, *Digital Image Processing*. New York: Wiley, 1981.

[15] C. E. Liedtke, T. Gahm, F. Kappei, and B. Aeikens, "Segmentation of microscopic cell scenes," *Anal. Quantum Cytol.*, vol. 9, pp. 197–211, 1987.

[16] R. M. Haralick, S. R. Sternberg, and X. Zhuang, "Image analysis using mathematical morphology," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 9, no. 4, pp. 532–549, July 1987.

[17] J. Serra, *Image Analysis and Mathematical Morphology*. London: Academic, 1982.

[18] S. Sternberg, "Gray scale morphology," *Computer Vision, Graphics, Image Processing*, vol. 35, pp. 333–355, 1986.

[19] R. M. Haralick and L. G. Shapiro, *Computer and Robot Vision*, vol. 1. Reading, MA: Addison Wesley, 1992.

[20] C. Lantuéjoul and S. Beucher, "On the use of geodesic metric in image analysis," *J. Microsc.*, vol. 121, pp. 39–49, 1981.

[21] C. Lantuéjoul and F. Maisonneuve, "Geodesic methods in image analysis," *Pattern Recognition*, vol. 17, pp. 117–187, 1984.

[22] L. Vincent and E. R. Dougherty, "Morphological segmentation for textures and particles," in *Digital Image Processing Methods*, E. R. Dougherty, Ed. New York: Dekker, 1994, pp. 43–102.

[23] F. Meyer, "Contrast feature extraction," in *Quantitative Analysis of Microstructures in Material Sciences, Biology and Medicine*, J.-L. Chermant, Ed., Special issue of Practical Metallography Stuttgart, Germany: Riederer Verlag, 1978.

[24] A. Moragas, M. Garcia-Bonafe, I. de Torres, and M. Sana, "Textural analysis of lymphois cells in serous effusios, a mathematical morphologic approach," *Anal. Quantum Cytol.*, vol. 15, pp. 1656–1170, 1993.

[25] Y. Linde, A. Buzo, and R. M. Gray, "An algorithm for vector quantization design," *IEEE Trans. Communicat.*, vol. COM-28, no. 1, pp. 84–95, Jan. 1980.

[26] R. M. Gray, "Vector quantization," *IEEE Acoust., Speech, Signal Processing Mag.*, vol. 1, pp. 9–31, Apr. 1984.

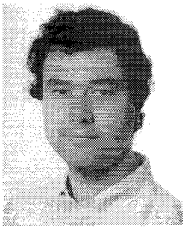
[27] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*. Boston, MA: Kluwer, 1992.

[28] R. P. Lippmann, "An introduction to computing with neural network," *IEEE Acoust., Speech, Signal Processing Mag.*, pp. 4–22, Apr. 1987.

[29] T. Kohonen, "The self-organizing map," in *Proc. IEEE*, vol. 78, no. 9, pp. 1464–1480, Sept. 1990.

[30] G. A. Carpenter and S. Grossberg, "A self-organizing neural network for supervised learning, recognition and prediction," *IEEE Communicat. Mag.*, pp. 38–48, Sept. 1992.

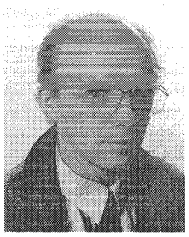
[31] J.-P. Thiran and M.-O. Becks, "Morphologie mathématique et quantification vectorielle pour la classification d'images médicales," Masters thesis, Catholic Univ. of Louvain, Telecommunications and Remote Sensing Laboratory, Louvain-la-Neuve, Belgium, June 1993.



Jean-Philippe Thiran (S'96) was born in Namur, Belgium, in 1970. He received the degree in electrical engineering from the Université Catholique de Louvain (UCL), Louvain-la-Neuve, Belgium, in 1993.

He is now a Ph.D. degree student with the Telecommunication Laboratory of the UCL, where his works are funded by a grant from the Belgian FRIA foundation. His current interests include 2-D and 3-D medical imaging, image segmentation, multimodal image registration and fusion, shape analysis, computer graphics, and telediagnosis in medical imaging.

Mr. Thiran received the 1993 AILv Award and won the 1993 UCL IEEE Local Paper Contest, both for his master thesis. He is an active member of the Organizing Committee of the 1996 IEEE International Conference on Image Processing (ICIP'96).



Benoît Macq (S'83–M'84) received the "Ingénieur Civil Electricien" and the "Docteur en Sciences Appliquées" degrees from the Université Catholique de Louvain (UCL), Louvain-la-Neuve, Belgium, in 1984 and 1989, respectively.

After his military service at the Ecole Royale Militaire, Brussels in 1984, he worked on telecommunication planning in the Tractionnel Society in 1985, and on video coding in the Telecommunication Laboratory of the UCL from 1986 to 1990. From 1990 to 1991, he was with the Philips Research Laboratory Belgium. He is now a Senior Researcher of the Belgian NSF ("Chercheur Qualifié" du FNRS) and Associate Professor with the Telecommunication Laboratory of the Université Catholique de Louvain. His current interests concern algorithms for image compression and video coding in broadband digital networks—based on wavelets, neural networks, perceptual quantization, entropy coding, and mathematical morphology—and multimedia access control and copyright protection for images.

Dr. Macq has participated to the RACE Hivits, to the COST 211 bis and ter, to the Belgian Broadband Experiment and to MPEG-2 and MPEG-4 groups. He has been Coordinator of the CEC RACE M1005—ACCOPI (Access Control and COpiright Protection for Images) project which has strongly influenced the DAVIC standardization security group. He is now Coordinator of the ACTS-OKAPI project. He is the leader of a 20-researcher team involved in many other projects related to image processing and multimedia telecommunications. He is reviewer for several IEEE Transactions, for image communication and signal processing. He is Co-Organizer of the ECMAST'96 conference and has been Session Chairman for several conferences. He received the 1990 Bell Telephone MC Award and the URSI Young Scientist Award in 1991. He is member of the SPIE.