(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 1 579 366 B1

(12) EUROPEAN PATENT SPECIFICATION

(45) Date of publication and mention of the grant of the patent:
02.05.2007 Bulletin 2007/18

(21) Application number: **03811419.5**

(22) Date of filing: **13.11.2003**

(51) Int Cl.:
G06T 7/00 *(2006.01)* ***G06T 7/60*** *(2006.01)*

(86) International application number:
PCT/GB2003/004910

(87) International publication number:
WO 2004/046994 (03.06.2004 Gazette 2004/23)

(54) HISTOLOGICAL ASSESSMENT OF NUCLEAR PLEOMORPHISM

HISTOLOGISCHE BEWERTUNG DES NUKLEARPLEOMORPHISMUS

EVALUATION HISTOLOGIQUE DE PLEOMORPHISME NUCLEAIRE

(84) Designated Contracting States:
AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IT LI LU MC NL PT RO SE SI SK TR

(30) Priority: **21.11.2002 GB 0227160**

(43) Date of publication of application:
28.09.2005 Bulletin 2005/39

(73) Proprietor: **QinetiQ Limited**
London, SW1E 6PD (GB)

(72) Inventors:
- **PETROU, Maria,**
  School of Elect. Computing & Math.
  Guildford GU2 7XH (GB)
- **KESIDIS, Anastasios**
  11524 Athens (GR)
- **VARGA, Margaret Jai,**
  Qinetiq Limid
  Malvern,
  Worcestershire WR14 3PS (GB)
- **DUCKSBURY, Paul Gerard,**
  Qinetiq Limited
  Malvern,
  Worcestershire WR14 3PS (GB)

(74) Representative: **Williams, Arthur Wyn Spencer et al**
QinetiQ Limited
Intellectual Property
Malvern Technology Centre
St Andrews Road
Malvern
Worcestershire WR14 3PS (GB)

(56) References cited:
WO-A-99/08091 **US-A- 5 978 497**

- **OTSU N: "A THRESHOLD SELECTION METHOD FROM GRAY-LEVEL HISTOGRAMS" IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS, IEEE, NEW YORK, NY, US, vol. 9, no. 1, 1979, pages 62-66, XP000617438 ISSN: 0018-9472 cited in the application**
- **DOUDKINE A ET AL: "Nuclear texture measurements in image cytometry" PATHOLOGICA, GENOVA, IT, vol. 87, 1995, pages 286-299, XP002079698 ISSN: 0031-2983**

EP 1 579 366 B1

Description

[0001] This invention relates to a method, an apparatus and a computer program for histological assessment of nuclear pleomorphism: it is particularly relevant (although not exclusively) to assessment of histological images to provide clinical information on cancers such as breast cancer. The method is also relevant to other cancers, e.g. colon and cervical cancer as well as breast cancer.

[0002] Breast cancer is a common form of female cancer, and it occurs also in the male albeit to a lesser extent: once a lesion indicative of breast cancer has been detected, tissue samples are taken and examined by a histopathologist to establish a diagnosis, prognosis and treatment plan. However, pathological analysis of tissue samples is a time consuming and subjective process. It entails interpretation of images by human eye: it can be characterised by inconsistencies in observations of the same samples by different observers and even by the same observer at different times. For example, two different observers assessing images of the same tissue samples may give different opinions for a number of the images. Different opinions may arise in as many as 30% of the images. The problem is exacerbated by heterogeneity, i.e. complexity of some tissue sample features.

[0003] WO 99/08091 A (Oncometrics Imaging Corp.) 18 February 1999 discloses a system and method for detecting diagnostic cells and cells having malignancy-associated changes. US-A-5 978 497 (Bannister Wendy R et al.) 2 November 1999 discloses an automated cytology system that identifies and classifies free-lying cells.

[0004] There is a need to provide an objective measurement of the degree of nuclear pleomorphism to support the pathologist's diagnosis and the patient's treatment.

[0005] The present invention provides a method of histological assessment of nuclear pleomorphism by identifying image regions potentially corresponding to cell nuclei in histological image data, characterised in that the method also includes:

a) decomposing colour image data into a set of sub-images each of which overlaps its neighbours and removing from each sub-image regions at sub-image edges potentially corresponding to cell nuclei, Principal Component Analysis (PCA) being used to transform each sub-image into greyscale image data comprising a first principal component obtained from a matrix which is a product of a matrix of colour image data and a matrix of eigenvectors of a covariance matrix, the greyscale image data providing improved image definition compared to an individual red green or blue plane in colour image data, and

b) thresholding the greyscale image data to render it binary, determining perimeters and areas of identified image regions, calculating image region shape factors from the perimeters and areas and assessing pleomorphism from the shape factors' statistical parameters.

[0006] The invention provides the advantage that it is an objective measurement to inform a pathologist's diagnosis and patient treatment.

[0007] The shape factors' statistical parameters may comprise at least one of their mean, weighted mean, median, mode, maximum and minimum. The step of thresholding the imaged data may be Otsu thresholding.

[0008] The step of assessing pleomorphism determines pleomorphism as being relatively low, moderate or high according to whether the mean or median of the shape factors is relatively low, moderate or high respectively. A shape factor S for an image region potentially corresponding to a cell nucleus may be given by $S = \frac{kP^2}{A}$, where k is a constant, P is image region perimeter and A is image region area, and a mean shape factor S_m for a set of image regions potentially corresponding to cell nuclei may be thresholded as $S_m \leq 30k$ (low), $30k < S_m \leq 35k$ (moderate) and $S_m > 35k$ (high) corresponding to pleomorphism being relatively low, moderate or high respectively.

[0009] The step of identifying image regions potentially corresponding to cell nuclei in histological image data includes filtering the image data to overwrite regions which are not of interest using a filtering process which does not appreciably affect image region perimeter. The overwriting step may include setting relatively small image regions to a background pixel value and setting hole pixels in relatively larger image regions to a non-hole image region pixel value.

[0010] Image regions potentially corresponding to cell nuclei in histological image data may be identified by a procedure including:

a) dividing the image data into overlapping sub-images,

b) applying PCA to each sub-image to provide a respective greyscale sub-image, and

c) removing from the greyscale sub-images:

i) image regions touching or intersecting sub-image boundaries,
ii) unsuitably small image regions, and
iii) holes in relatively large image regions, and

d) reassembling the sub-images.

[0011] In another aspect, the present invention provides an apparatus for histological assessment of nuclear pleomorphism by identifying image regions potentially corresponding to cell nuclei in histological image data, characterised in that the apparatus incorporates a computer programmed to:

a) decompose colour image data into a set of sub-images each of which overlaps its neighbours and remove from each sub-image regions at sub-image edges potentially corresponding to cell nuclei,
b) use PCA to transform colour image data in each sub-image into greyscale image data comprising a first principal component obtained from a matrix which is a product of a matrix of colour image data and a matrix of eigenvectors of a covariance matrix, the greyscale image data providing improved image definition compared to an individual red green or blue plane in colour image data, and
c) threshold the greyscale image data to render it binary, determine perimeters and areas of identified image regions, calculate image region shape factors from the perimeters and areas and assess pleomorphism from the shape factors' statistical parameters.

[0012] In a further aspect, the present invention provides computer software for use in histological assessment of nuclear pleomorphism and having instructions for controlling a computer to identify image regions potentially corresponding to cell nuclei in histological image data, characterised in that the software also has instructions for controlling a computer to:

a) decompose colour image data into a set of sub-images each of which overlaps its neighbours and remove from each sub-image regions at sub-image edges potentially corresponding to cell nuclei,
b) use PCA to transform colour image data in each sub-image into greyscale image data comprising a first principal component obtained from a matrix which is a product of a matrix of colour image data and a matrix of eigenvectors of a covariance matrix, the greyscale image data providing improved image definition compared to an individual red green or blue plane in colour image data, and
c) threshold the greyscale image data to render it binary, determine perimeters and areas of identified image regions, calculate image region shape factors from the perimeters and areas and assess pleomorphism from the shape factors' statistical parameters.

[0013] The computer apparatus and computer software aspects of the invention may have preferred features equivalent to corresponding method aspects of the invention.

[0014] In order that the invention might be more fully understood, embodiments thereof will now be described, by way of example only, with reference to the accompanying drawings, in which:-

Figure 1 is a block diagram of a procedure for measuring pleomorphism to assist in formulating diagnosis and treatment; and

Figure 2 is a block diagram showing in more detail pleomorphism feature detection of the invention for use in the procedure of Figure 1.

[0015] A procedure 10 for the assessment of tissue samples in the form of histopathological slides of carcinomas of the breast is shown in Figure 1. This drawing illustrates processes which measure degree of pleomorphism for use in assessment of patient condition.

[0016] The procedure 10 employs a database 12, which maintains digitised image data obtained from histological slides as will be described later. Sections are taken (cut) from breast tissue samples (biopsies) and placed on respective slides. Slides are stained using the staining agent haemotoxylin & eosin (H&E). H&E is a very common stain for delineating tissue and cellular structure. Tissue stained with H&E is used to assess pleomorphism.

[0017] Nuclear pleomorphism is a measurement of degree of variability in shape of the nucleus of tumour cells within tissue. In normal tissue, cell nuclei have a regular structure in terms of shape and size, whereas nuclei of cancer cells can become larger and irregularly shaped, with a marked variation in shape and size.

[0018] In a prior art manual procedure, a clinician places a slide under a microscope with magnification of 40X and examines a region of it (often referred to as a tile) for indications of the degree of pleomorphism. This manual procedure requires a pathologist subjectively to assess unusual size and shape of cell nuclei in a tissue sample. The values obtained

in this way are combined to give a single measurement for use in diagnosis. The process of the invention replaces the prior art subjective procedure with an objective procedure.

[0019] In the present example, image data were obtained from histological slides by a pathologist using Zeiss Axioskop microscope with a Jenoptiks Progres 3012 digital camera. Image data from each slide is a set of digital images obtained at a linear magnification of 40 (i.e. 40X linear, 1600X area), each image being an electronic equivalent of a tile. Images obtained by other microscopes and cameras can also be used.

[0020] To obtain images, a pathologist scans the microscope over a slide, and at 40X magnification selects regions of the slide which appear to be most promising in terms of pleomorphism assessment. Each of these regions is then photographed using the microscope and camera mentioned above, and this produces for each region a respective digitised image in three colours, red, green and blue (R, G & B). Three intensity values are obtained for each pixel in a pixel array to provide an image as a combination of R, G and B image planes. This image is stored temporarily at 12 for later use. Two tiles are required for pleomorphism measurement at 14 by a process 16: the results of the process 16 are converted into a measurement at 20 for input to a diagnostic report at 22.

[0021] Referring now to Figure 2, the pleomorphism feature detection process 16 is shown in more detail. It is carried out for each of the two tiles or digitised images (raw (RGB) input images) mentioned above, and will be described for one such image. At a first stage 30 the raw (RGB) input image is separated into overlapping windows of size 128x128 pixels. In each window, 64 pixels overlap with 64 pixels of respective neighbouring windows in both horizontal and vertical directions. For example, an image of 256x256 would give a 3x3 set of 128x128 overlapping windows: thus, each window overlaps half of each of its row and column neighbours. To each window a technique referred to as "Principal Component Analysis" (PCA, Karhunen-Loeve Transform) is applied. PCA is a standard mathematical technique described by Jollie I.T., 'Principal Component Analysis', Springer series in statistics, Springer Verlag, 1986. It is also described by Jackson J.E., 'A User Guide to Principal Components', John Wiley & Sons, 1991, pp 1-25. PCA is a standard mathematical technique for transforming a set of (possibly) correlated variables into a smaller number of uncorrelated variables called principal components. Of the principal components, a first principal component accounts for as much of the variables' variability as possible compared to other principal components. In this first principal component, image definition, contrast and characteristics are improved compared to R, G or B image planes in colour image data. PCA is a transformation into a domain or representation in which data is more easily exploited or understood (e.g. separable or classified into different classes or categories). There are other transformations (analysis techniques or filters) which could also be used, and which have particular characteristics (e.g. linear or non-linear, exploiting data characteristics, means, variances etc. in different ways). Neural networks/adaptive networks and related techniques are also used to implement such transformations.

[0022] PCA involves calculating a covariance matrix and solving for its eigenvalues and eigenvectors. The image is now treated as being arranged as an Nx3 matrix, i.e. having N pixels and three planes (R, G and B). The covariance matrix is calculated using a formula for its matrix elements $C_{i,j}$ as follows:

$$C_{i,j} = \frac{\sum_{k=1}^{N}(x_k - \mu_x)(y_k - \mu_y)}{N-1} \qquad (1)$$

where $C_{i,j}$ is the covariance of variable i with variable j, x_k and y_k are the *ith* and *jth* feature values of the *kth* object, μ_x is the mean of all N values of x_k, μ_y is the mean of all N values of y_k. The covariance matrix is 3x3 and PCA yields three eigenvectors: the eigenvectors are treated as a 3x3 matrix, which is used to multiply the transpose of the Nx3 image matrix to produce a product matrix. The product matrix has an Nx1 first column which is the first principal component, which may be considered as the most important component: it is the component with the maximum eigenvalue, and it provides a greyscale sub-image (one pixel value for each of N pixels) with a maximum range of information compared to equivalents associated with other components. PCA therefore produces monochromatic image data. It is carried out for each of the overlapping windows defined above, and each provides a respective first principal component and greyscale sub-image of size 128x128 pixels.

[0023] At 32, a thresholding technique referred to as "Otsu" is applied to each sub-image resulting from 30 to transform it into a respective binary sub-image. Otsu is a standard thresholding technique published by Otsu N., 'A thresholding selection method from grey level histograms', IEEE Trans Systems, Man & Cybernetics, vol. 9, 1979, pp 62-66. The Otsu threshold selection method aims to minimise for two classes a ratio of within-class variance to between-class variance: i.e. the higher the variance between classes the better the separation. Otsu is a particularly preferred thresholding technique. In the present example the two classes are a below-threshold class (pixel value 0) and an above-threshold class (pixel value 1), so Otsu thresholding converts each greyscale sub-image to a binary sub-image containing

a set of blobs: here blobs are image regions (objects in the image) each of which is a respective group of mutually contiguous pixels all having value 1. The blobs may have holes (pixel value 0) in them.

[0024] At 34 all blobs that touch or are intersected by sub-image boundaries are removed. Thus, if at any pixel a blob meets a border it is removed by setting its pixels to a background pixel value. This is because boundaries meeting blobs produce artificially straight blob edges which can give misleading results later. Because of sub-image overlap, a blob which appears partly in one image may appear wholly in another sub-image and so is not necessarily lost.

[0025] At 36 the sub-images from 34 are inverted (pixel value 0 changes to 1 and vice versa) and connected component labelling (CCL) is applied to remove holes in blobs. CCL is a known image processing technique (sometimes referred to as 'blob colouring') published by Klette R., Zamperoniu P., 'Handbook of Image Processing Operators', John Wiley & Sons, 1996, and Rosenfeld A., Kak A.C., 'Digital Picture Processing', vols. 1 & 2, Academic Press, New York, 1982. CCL gives a respective label to each group of contiguous pixels of pixel value 1. Because of the inversion, areas of pixel value 1 which are labelled in CCL are now holes within blobs together with background pixels. Holes within each blob are now removed (filled) by setting their pixels to the value of other pixels of the blob. Background pixels are left unchanged.

[0026] At 38, the sub-images from 36 are inverted once more and CCL is applied again: due to this second inversion, areas labelled by CCL are now filled blobs within each sub-image. CCL has a facility for removal of blobs with areas smaller than a user-defined minimum area threshold. In this example, using 40X magnification, the minimum area threshold is 400 pixels: all blobs with areas less than 400 pixels are rejected and merged into the background by setting their pixels to a background value (0). Remaining blobs with area of at least 400 pixels are accepted for further processing as set of labelled blobs. CCL also gives for each remaining blob its perimeter P and area A in numbers of pixels in each case.

[0027] After stages 30 to 38, each sub-image is clear of unwanted small blobs: all remaining blobs have been filled to remove holes within them, so they consist of pixels which are all the same value. The advantage of stages 30 to 38 is that they provide spatial filtering but do not appreciably affect perimeter shapes of blobs, which is important for subsequent processing. Such filtering is not essential but it is helpful to reduce processing burden.

[0028] At 40, the sub-images output from step 38 are reassembled into a new binary image: the new binary image has the same size as the original raw (RGB) input image, and contains only blobs that will be assessed in subsequent pleomorphism processing.

[0029] Steps 30 to 40 are pre-processing steps which identify a set of blobs within the original raw (RGB) input image: each of these blobs should now correspond to a cell nucleus. At 42 a statistical analysis is applied. A shape factor $S = \frac{P^2}{A}$ is calculated for each blob from blob perimeter P and area A obtained in CCL at 38. S could also be a multiple or fraction of $\frac{P^2}{A}$ were this to be convenient, and other functions of P and A could be used to indicate shape. The value of S increases with increasing irregularity of blob shape, it is 4π (~12.57) for a perfect circle. A mean value S_m of S is calculated for all blobs in the image by adding their S values together and dividing by their number. S_m is thresholded to derive a measure of the pleomorphism of the original raw (RGB) input image: S_m thresholds were derived from an analysis of a test set of 200 pleomorphism images. There are three threshold categories, $S_m \leq 30$ (low), $30 < S_m \leq 35$ (moderate) and $S_m > 35$ (high) with pleomorphism scores 1, 2 and 3 respectively as tabulated below. These thresholds are for $S = \frac{P^2}{A}$: other shape factor expressions may need different thresholds.

S_m : Mean Value of Blob Shape Factor	Pleomorphism Score
Low, $S_m \leq 30$	1
Moderate, $30 < S_m \leq 35$	2
High, $S_m > 35$	3

[0030] For $S = \frac{kP^2}{A}$, where k is a constant, the three threshold categories become $S_m \leq 30\,k$ (low), $30\,k < S_m \leq 35\,k$ (moderate) and $S_m > 35k$ (high).

[0031] S_m may be calculated by other procedures. It might simply be the median of an odd number of values of S, or the average of two central values of an even number of values of S. It is also possible to use the maximum or minimum of a series of values of S, or with a largely single mode distribution, the mode value. However, this will affect thresholds. A weighted mean S_{wm} could be calculated by multiplying each value of S by a weight factor w_i, adding the weighted values and dividing their sum by the sum of the weights; each weight would indicate the probability of the associated blob or cell nucleus giving a reliable indication of pleomorphism: i.e.

$$S_{wm} = \frac{\sum_i w_i S_i}{\sum_i w_i} \qquad (2)$$

[0032] The above discussion shows that there are a number of ways an overall shape factor can be determined from the statistical characteristics of a set of shape factors, e.g. their mean, weighted mean, median, mode, maximum and minimum may all be used individually.

[0033] At 20 the score for the nuclear pleomorphism measurement has a value 1, 2 or 3 and a respective score is obtained for each of the input tiles. The maximum of these two scores is taken as the overall pleomorphism score.

Measurement of Nuclear Pleomorphism	Meaning	Points
Uniform	Minimal increase in size and variation in size compared to normal cell nuclei, i.e. nuclei are relatively small and uniform in size.	1
Moderate Variation	Moderate increase and variation in size and shape with vesicular nuclei.	2
Marked Variation	Marked variation in size and shape with vesicular nuclei.	3

[0034] The measurement of nuclear pleomorphism may be combined with others obtained for mitosis and tubules by different methods to derive an overall grading referred to in medicine as a "Bloom and Richardson grading" or modified versions of this grading: it is used by clinicians as a measure of cancer status.

[0035] The examples given in the foregoing description for calculating results can clearly be evaluated by an appropriate computer program on a carrier medium and running on a conventional computer system. Such a program is straightforward for a skilled programmer to implement without requiring invention because the procedures are well known, and it will therefore not be described further.

Claims

1. A method of histological assessment of nuclear pleomorphism by identifying image regions potentially corresponding to cell nuclei in histological image data, the method also including:

   a) decomposing (30) colour image data into a set of sub-images each of which overlaps its neighbours and removing from each sub-image regions at sub-image edges potentially corresponding to cell nuclei, Principal Component Analysis (PCA) (30) being used to transform each sub-image into greyscale image data comprising a first principal component obtained from a matrix which is a product of a matrix of colour image data and a matrix of eigenvectors of a covariance matrix, the greyscale image data providing improved image definition compared to an individual red green or blue plane in colour image data, and
   b) thresholding (32) the greyscale image data to render it binary, determining perimeters and areas of identified image regions, calculating (42) image region shape factors from the perimeters and areas and assessing (44) pleomorphism from the shape factors' statistical parameters.

2. A method according to Claim 1 **characterised in that** PCA is used to transform (30) colour image data in each sub-image into greyscale image data by:

a) applying PCA to each sub-image to provide a respective greyscale sub-image and removing from the greyscale sub-images:

i) image regions touching or intersecting sub-image boundaries,
ii) unsuitably small image regions, and
iii) holes in relatively large image regions, and

b) reassembling the sub-images.

3. A method according to Claim 1 or 2 **characterised in that** the shape factors' statistical parameters comprise at least one of their mean, weighted mean, median, mode, maximum and minimum.

4. A method according to Claim 1 or 2 **characterised in that** the step of thresholding the imaged data is a step (32) of Otsu thresholding, in which threshold selection aims to minimise for two classes a ratio of within-class variance to between-class variance.

5. A method according to Claim 1 or 2 **characterised in that** the step (44) of assessing pleomorphism determines pleomorphism as having a first magnitude, a second magnitude or a third magnitude according to whether the mean or median of the shape factors is less than values in an intermediate range, or in the intermediate range, or greater than values in the intermediate range respectively.

6. A method according to Claim 5 **characterised in that** a shape factor S for an image region potentially corresponding to a cell nucleus is given by $S = \frac{kP^2}{A}$, where k is a constant, P is image region perimeter and A is image region area, and a mean shape factor S_m for a set of image regions potentially corresponding to cell nuclei is thresholded (44) as $S_m \leq 30k$, $30k < S_m \leq 35k$ and $S_m > 35k$ corresponding respectively to pleomorphism being below, in or above the intermediate range.

7. A method according to Claim 1 or 2 **characterised in that** the step of identifying image regions potentially corresponding to cell nuclei in histological image data includes filtering the image data to overwrite regions which are not of interest using a filtering process (38) which removes regions with areas smaller than a user-defined minimum.

8. A method according to Claim 7 **characterised in that** the step of filtering the image data to overwrite regions which are not of interest includes setting (38) first image regions to a background pixel value and setting (36) hole pixels in second image regions to a non-hole image region pixel value, wherein the first image regions are smaller than the second image regions.

9. Apparatus for histological assessment of nuclear pleomorphism by identifying image regions potentially corresponding to cell nuclei in histological image data, the apparatus incorporating a computer programmed to:

a) decompose (30) colour image data into a set of sub-images each of which overlaps its neighbours and remove from each sub-image regions at sub-image edges potentially corresponding to cell nuclei,
b) use PCA (30) to transform colour image data in each sub-image into greyscale image data comprising a first principal component obtained from a matrix which is a product of a matrix of colour image data and a matrix of eigenvectors of a covariance matrix, the greyscale image data providing improved image definition compared to an individual red green or blue plane in colour image data, and
c) threshold (32) the greyscale image data to render it binary, determine perimeters and areas of identified image regions, calculate (42) image region shape factors from the perimeters and areas and assess (44) pleomorphism from the shape factors' statistical parameters.

10. Apparatus according to Claim 9 **characterised in that** the computer is programmed to use PCA to:

a) transform (30) colour image data in each sub-image into greyscale image data by applying PCA to each sub-image to provide a respective greyscale sub-image,
b) remove from the greyscale sub-images:

i) image regions touching or intersecting sub-image boundaries,
ii) unsuitably small image regions, and
iii) holes in relatively large image regions, and

c) reassemble the sub-images.

11. Apparatus according to Claim 9 or 10 **characterised in that** the shape factors' statistical parameters comprise at least one of their mean, weighted mean, median, mode, maximum and minimum.

12. Apparatus according to Claim 9 or 10 **characterised in that** the computer is programmed to threshold the image data using Otsu thresholding in which threshold selection aims to minimise for two classes a ratio of within-class variance to between-class variance.

13. Apparatus according to Claim 9 or 10 **characterised in that** the computer is programmed to determine pleomorphism (44) as having a first magnitude, a second magnitude or a third magnitude according to whether the mean or median of the shape factors is less than values in an intermediate range, or in the intermediate range, or greater than values in the intermediate range respectively.

14. Apparatus according to Claim 9 or 10 **characterised in that** the computer is programmed to determine a shape factor S for an image region potentially corresponding to a cell nucleus, the shape factor S being given by $S = \frac{kP^2}{A}$, where k is a constant, P is image region perimeter and A is image region area, and the computer is also programmed to determine (42) a mean shape factor S_m for a set of image regions potentially corresponding to cell nuclei, to threshold the mean shape factor as $S_m \leq 30k$, $30k < S_m \leq 35k$ and $S_m > 35k$ corresponding respectively to pleomorphism being below, in or above the intermediate range.

15. Apparatus according to Claim 9 or 10 **characterised in that** the computer is programmed to set (38) first image regions to a background pixel value and to set (36) hole pixels in second image regions to a non-hole image region pixel value, wherein the first image regions are smaller than the second image regions.

16. Computer software for use in histological assessment of nuclear pleomorphism and having instructions for controlling a computer to identify image regions potentially corresponding to cell nuclei in histological image data, the software also having instructions for controlling a computer to:

a) decompose (30) colour image data into a set of sub-images each of which overlaps its neighbours and remove from each sub-image regions at sub-image edges potentially corresponding to cell nuclei,
b) use PCA (30) to transform colour image data in each sub-image into greyscale image data comprising a first principal component obtained from a matrix which is a product of a matrix of colour image data and a matrix of eigenvectors of a covariance matrix, the greyscale image data providing improved image definition compared to an individual red green or blue plane in colour image data, and
c) threshold (32) the greyscale image data to render it binary, determine perimeters and areas of identified image regions, calculate (42) image region shape factors from the perimeters and areas and assess (44) pleomorphism from the shape factors' statistical parameters.

17. Computer software according to Claim 16 **characterised in that** it has instructions for controlling a computer to use PCA to:

a) transform (30) colour image data in each sub-image into greyscale image data by applying PCA to each sub-image to provide a respective greyscale sub-image,
b) remove from the greyscale sub-images:

i) image regions touching or intersecting sub-image boundaries,
ii) unsuitably small image regions, and
iii) holes in relatively large image regions, and

c) reassemble the sub-images

18. Computer software according to Claim 16 or 17 **characterised in that** the shape factors' statistical parameters comprise at least one of their mean, weighted mean, median, mode, maximum and minimum.

19. Computer software according to Claim 16 or 17 **characterised in that** it has instructions for controlling a computer to threshold the imaged data using Otsu thresholding in which threshold selection aims to minimise for two classes a ratio of within-class variance to between-class variance.

20. Computer software according to Claim 16 or 17 **characterised in that** it has instructions for controlling a computer to determine (44) pleomorphism as having a first magnitude, a second magnitude or a third magnitude according to whether the mean or median of the shape factors is less than values in an intermediate range, or in the intermediate range, or greater than values in the intermediate range respectively.

21. Computer software according to Claim 20 **characterised in that** it has instructions for controlling a computer to:

a) determine (42) a shape factor S for an image region potentially corresponding to a cell nucleus, the shape factor S being given by $S = \frac{kP^2}{A}$, where k is a constant, P is image region perimeter and A is image region area,
b) threshold (44) a mean shape factor S_m for a set of image regions potentially corresponding to cell nuclei as $S_m \leq 30k$, $30k < S_m \leq 35k$ and $S_m > 35k$, and
c) indicate pleomorphism being relatively low, moderate or high respectively for a set of image regions potentially corresponding to cell nuclei thresholded (44) as $S_m \leq 30k$, $30k < S_m \leq 35k$ and $S_m > 35k$ corresponding respectively to pleomorphism being below, in or above the intermediate range.

22. Computer software according to Claim 16 or 17 **characterised in that** it includes instructions for setting (38) first image regions to a background pixel value and setting (36) hole pixels in second image regions to a non-hole image region pixel value, wherein the first image regions are smaller than the second image regions.

Patentansprüche

1. Verfahren zur histologischen Bewertung von nukleärem Pleomorphismus, indem Bildbereiche identifiziert werden, die potenziell Zellkernen in histologischen Bilddaten entsprechen, wobei das Verfahren außerdem enthält:

a) Zerlegen (30) von Farbbilddaten in eine Menge von Teilbildern, von denen jedes seine Nachbarn überlappt, und Entfernen von Bereichen an den Rändern der Teilbilder, die potenziell Zellkernen entsprechen, aus jedem Teilbild, wobei Hauptkomponentenanalyse (PCA, Principal Component Analysis) (30) verwendet wird, um jedes Teilbild in Bilddaten mit Graustufung umzuwandeln, die eine erste Hauptkomponente umfassen, die aus einer Matrix erhalten wird, die ein Produkt aus einer Matrix aus Farbbilddaten und einer Matrix aus Eigenvektoren einer Kovarianzmatrix ist, wobei die Bilddaten mit Graustufung eine verbesserte Bilddefinition im Vergleich zu einer einzelnen roten, grünen oder blauen Ebene in Farbbilddaten liefern, und
b) Vergleichen (32) der Bilddaten mit Graustufung mit einem Schwellwert, um sie binär zu machen, Feststellen der Umfänge und Flächen der identifizierten Bildbereiche, Berechnen (42) von Formfaktoren der Bildbereiche aus den Umfängen und Flächen, und Beurteilen (44) des Pleomorphismus an Hand der statistischen Parameter der Formfaktoren.

2. Verfahren nach Anspruch 1, **dadurch gekennzeichnet, dass** Hauptkomponentenanalyse verwendet wird, um Farbbilddaten in jedem Teilbild in Bilddaten mit Graustufung umzuwandeln (30), indem:

a) Hauptkomponentenanalyse auf jedes Teilbild angewendet wird, um ein jeweiliges Teilbild mit Graustufung bereitzustellen, und aus den Teilbildern mit Graustufung folgendes entfernt wird:

i) Bildbereiche, die Grenzen von Teilbildern berühren oder schneiden,
ii) ungeeignet kleine Bildbereiche, und
iii) Löcher in relativ großen Bildbereichen, und

b) die Teilbilder wieder zusammengesetzt werden.

3. Verfahren nach Anspruch 1 oder 2, **dadurch gekennzeichnet, dass** die statistischen Parameter der Formfaktoren wenigstens eine der Größen Mittelwert, gewichteter Mittelwert, Zentralwert, häufigster Wert, Maximum und Minimum umfassen.

4. Verfahren nach Anspruch 1 oder 2, **dadurch gekennzeichnet, dass** der Schritt des Vergleichens der Bilddaten mit einem Schwellwert ein Schritt (32) mit einem Schwellwertvergleich nach Otsu ist, in dem die Auswahl des Schwellwerts darauf abzielt, ein Verhältnis der Varianz innerhalb von Klassen zur Varianz zwischen Klassen für zwei Klassen zu minimieren.

5. Verfahren nach Anspruch 1 oder 2, **dadurch gekennzeichnet, dass** der Schritt (44) der Bewertung des Pleomorphismus Pleomorphismus mit einem ersten Ausmaß, einem zweiten Ausmaß oder einem dritten Ausmaß feststellt, je nach dem, ob der Mittelwert oder der Zentralwert der Formfaktoren kleiner als Werte in einem Zwischenbereich ist, oder in dem Zwischenbereich liegt oder größer als Werte in dem Zwischenbereich ist.

6. Verfahren nach Anspruch 5, **dadurch gekennzeichnet, dass** ein Formfaktor S für einen Bildbereich, der potenziell einem Zellkern entspricht, durch $S = \frac{kP^2}{A}$ angegeben wird, wobei k eine Konstante ist, P der Umfang des Bildbereichs und A die Fläche des Bildbereichs ist, und ein mittlerer Formfaktor S_m für eine Menge von Bildbereichen, die potenziellen Zellkernen entsprechen, durch $S_m \leq 30k$, $30k < S_m \leq 35k$ und $S_m > 35k$ mit Schwellwerten verglichen wird (44), was jeweils entspricht, dass der Pleomorphismus unterhalb, innerhalb oder oberhalb des Zwischenbereichs liegt.

7. Verfahren nach Anspruch 1 oder 2, **dadurch gekennzeichnet, dass** der Schritt des Identifizierens von Bildbereichen, die potenziell Zellkernen in den histologischen Bilddaten entsprechen, das Filtern der Bilddaten umfasst, um Bereiche zu überschreiben, die nicht von Interesse sind, wobei ein Filterprozess (38) verwendet wird, der Bereiche mit Flächen entfernt, die kleiner als ein benutzerdefiniertes Minimum sind.

8. Verfahren nach Anspruch 7, **dadurch gekennzeichnet, dass** der Schritt des Filterns der Bilddaten, um Bereiche zu überschreiben, die nicht von Interesse sind, das Setzen (38) von ersten Bildbereichen auf einen Wert von Hintergrundpixeln und das Setzen (36) von Pixeln von Löchern in zweiten Bildbereichen auf einen Pixelwert von Bildbereichen, die keine Löcher darstellen, enthält, wobei die ersten Bildbereiche kleiner als die zweiten Bildbereiche sind.

9. Vorrichtung zur histologischen Bewertung von nukleärem Pleomorphismus, indem Bildbereiche identifiziert werden, die potenziell Zellkernen in histologischen Bilddaten entsprechen, wobei die Vorrichtung einen Computer enthält, der für folgendes programmiert ist:

a) Zerlegen (30) von Farbbilddaten in eine Menge von Teilbildern, von denen jedes seine Nachbarn überlappt, und Entfernen von Bereichen an den Rändern der Teilbilder, die potenziell Zellkernen entsprechen, aus jedem Teilbild,
b) Verwenden der Hauptkomponentenanalyse (30), um Farbbilddaten in jedem Teilbild in Bilddaten mit Graustufung umzuwandeln, die eine erste Hauptkomponente umfassen, die aus einer Matrix erhalten wird, die ein Produkt aus einer Matrix aus Farbbilddaten und einer Matrix aus Eigenvektoren einer Kovarianzmatrix ist, wobei die Bilddaten mit Graustufung eine verbesserte Bilddefinition im Vergleich zu einer einzelnen roten, grünen oder blauen Ebene in den Farbbilddaten liefern, und
c) Vergleichen (32) der Bilddaten mit Graustufung mit einem Schwellwert, um sie binär zu machen, Feststellen der Umfänge und Flächen der identifizierten Bildbereiche, Berechnen (42) der Formfaktoren der Bildbereiche aus den Umfängen und Flächen, und Beurteilen (44) des Pleomorphismus an Hand der statistischen Parameter der Formfaktoren.

10. Vorrichtung nach Anspruch 9, **dadurch gekennzeichnet, dass** der Computer dazu programmiert ist, Hauptkomponentenanalyse zu verwenden, um

a) Farbbilddaten in jedem Teilbild in Bilddaten mit Graustufung umzuwandeln (30), indem Hauptkomponenten-

analyse auf jedes Teilbild angewendet wird, um ein jeweiliges Teilbild mit Graustufung bereitzustellen,
b) aus den Teilbildern mit Graustufung folgendes zu entfernen:

i) Bildbereiche, die Grenzen von Teilbildern berühren oder schneiden,
ii) ungeeignet kleine Bildbereiche, und
iii) Löcher in relativ großen Bildbereichen, und

c) die Teilbilder wieder zusammenzusetzen.

11. Vorrichtung nach Anspruch 9 oder 10, **dadurch gekennzeichnet, dass** die statistischen Parameter der Formfaktoren wenigstens eine der Größen Mittelwert, gewichteter Mittelwert, Zentralwert, häufigster Wert, Maximum und Minimum umfassen.

12. Vorrichtung nach Anspruch 9 oder 10, **dadurch gekennzeichnet, dass** der Computer dazu programmiert ist, die Bilddaten mit einem Schwellwert zu vergleichen, wobei ein Schwellwertvergleich nach Otsu verwendet wird, in dem die Auswahl des Schwellwerts darauf abzielt, ein Verhältnis der Varianz innerhalb von Klassen zur Varianz zwischen Klassen für zwei Klassen zu minimieren.

13. Vorrichtung nach Anspruch 9 oder 10, **dadurch gekennzeichnet, dass** der Computer dazu programmiert ist, Pleomorphismus als Pleomorphismus mit einem ersten Ausmaß, einem zweiten Ausmaß oder einem dritten Ausmaß festzustellen (44), je nach dem, ob der Mittelwert oder der Zentralwert der Formfaktoren kleiner als Werte in einem Zwischenbereich ist, oder in dem Zwischenbereich liegt, oder größer als Werte in dem Zwischenbereich ist.

14. Vorrichtung nach Anspruch 9 oder 10, **dadurch gekennzeichnet, dass** der Computer dazu programmiert ist, einen Formfaktor S für einen Bildbereich zu ermitteln, der potenziell einem Zellkern entspricht, wobei der Formfaktor S als $S = \frac{kP^2}{A}$ angegeben wird, wobei k eine Konstante ist, P der Umfang des Bildbereichs und A die Fläche des Bildbereichs ist, und der Computer außerdem dazu programmiert ist, einen mittleren Formfaktor S_m für eine Menge von Bildbereichen zu ermitteln (42), die potenziell Zellkernen entsprechen, den mittleren Formfaktor durch $S_m \leq 30k$, $30k < S_m \leq 35k$ und $S_m > 35k$ mit Schwellwerten zu vergleichen, was jeweils entspricht, dass der Pleomorphismus unterhalb, innerhalb oder oberhalb des Zwischenbereiches liegt.

15. Vorrichtung nach Anspruch 9 oder 10, **dadurch gekennzeichnet, dass** der Computer dazu programmiert ist, erste Bildbereiche auf einen Wert von Hintergrundpixeln zu setzen (38), und Pixel von Löchern in zweiten Bildbereichen auf einen Wert von Pixeln in einem Bildbereich, der kein Loch darstellt, zu setzen (36), wobei die ersten Bildbereiche kleiner als die zweiten Bildbereiche sind.

16. Computersoftware für die Verwendung bei der histologischen Bewertung von nukleärem Pleomorphismus mit Anweisungen zur Steuerung eines Computers, um Bildbereiche zu identifizieren, die potenziell Zellkernen in histologischen Bilddaten entsprechen, wobei die Software außerdem Anweisungen zur Steuerung eines Computers aufweist, um

a) Farbbilddaten in eine Menge von Teilbildern zu zerlegen (30), von denen jedes seine Nachbarn überlappt, und Bereiche an den Rändern der Teilbilder, die potenziell Zellkernen entsprechen, aus jedem Teilbild zu entfernen,
b) Hauptkomponentenanalyse zu verwenden (30), um jedes Teilbild in Bilddaten mit Graustufung umzuwandeln, die eine erste Hauptkomponente umfassen, die aus einer Matrix erhalten wird, die ein Produkt aus einer Matrix aus Farbbilddaten und einer Matrix aus Eigenvektoren einer Kovarianzmatrix ist, wobei die Bilddaten mit Graustufung eine verbesserte Bilddefinition im Vergleich zu einer einzelnen roten, grünen oder blauen Ebene in Farbbilddaten liefern, und
c) die Bilddaten mit Graustufung mit einem Schwellwert zu vergleichen (32), um sie binär zu machen, Umfänge und Flächen der identifizierten Bildbereichen festzustellen, die Formfaktoren der Bildbereiche aus den Umfängen und Flächen zu berechnen (42), und den Pleomorphismus an Hand der statistischen Parameter der Formfaktoren zu beurteilen (44).

17. Computersoftware nach Anspruch 16, **dadurch gekennzeichnet, dass** sie Anweisungen zur Steuerung eines

Computers aufweist, Hauptkomponentenanalyse zu verwenden, um

a) Farbbilddaten in jedem Teilbild in Bilddaten mit Graustufung umzuwandeln (30), indem Hauptkomponentenanalyse auf jedes Teilbild angewendet wird, um ein jeweiliges Teilbild mit Graustufung bereitzustellen,
b) aus den Teilbildern mit Graustufung folgendes zu entfernen:

i) Bildbereiche, die Grenzen von Teilbildern berühren oder schneiden,
ii) ungeeignet kleine Bildbereiche, und
iii) Löcher in relativ großen Bildbereichen, und

c) die Teilbilder wieder zusammenzusetzen.

18. Computersoftware nach Anspruch 16 oder 17, **dadurch gekennzeichnet, dass** die statistischen Parameter der Formfaktoren wenigstens eine ihrer Größen Mittelwert, gewichteter Mittelwert, Zentralwert, häufigster Wert, Maximum und Minimum umfassen.

19. Computersoftware nach Anspruch 16 oder 17, **dadurch gekennzeichnet, dass** sie Anweisungen zur Steuerung eines Computers aufweist, um die Bilddaten mit einem Schwellwert zu vergleichen, wobei ein Schwellwertvergleich nach Otsu verwendet wird, in dem die Auswahl des Schwellwerts darauf abzielt, ein Verhältnis der Varianz innerhalb von Klassen zur Varianz zwischen Klassen für zwei Klassen zu minimieren.

20. Computersoftware nach Anspruch 16 oder 17, **dadurch gekennzeichnet, dass** sie Anweisungen zur Steuerung eines Computers aufweist, um Pleomorphismus als Pleomorphismus mit einem ersten Ausmaß, einem zweiten Ausmaß oder einem dritten Ausmaß festzustellen (44), je nach dem, ob der Mittelwert oder der Zentralwert der Formfaktoren kleiner als Werte in einem Zwischenbereich ist, oder in dem Zwischenbereich liegt, oder größer als Werte in dem Zwischenbereich ist.

21. Computersoftware nach Anspruch 20, **dadurch gekennzeichnet, dass** sie Anweisungen zur Steuerung eines Computers aufweist, um

a) einen Formfaktor S für einen Bildbereich zu ermitteln (42), der potenziell einem Zellkern entspricht, wobei der Formfaktor durch $S = \frac{kP^2}{A}$ angegeben wird, wobei k eine Konstante ist, P der Umfang des Bildbereichs und A die Fläche des Bildbereichs ist,
b) einen mittleren Formfaktor S_m für eine Menge von Bildbereichen, die potenziellen Zellkernen entsprechen, durch $S_m \leq 30k$, $30k < S_m \leq 35k$ und $S_m > 35k$ mit Schwellwerten zu vergleichen, was jeweils entspricht, dass der Pleomorphismus unterhalb, innerhalb oder oberhalb des Zwischenbereiches liegt.

22. Computersoftware nach Anspruch 16 oder 17, **dadurch gekennzeichnet, dass** sie Anweisungen zur Steuerung eines Computers aufweist, um erste Bildbereiche auf einen Wert von Hintergrundpixeln zu setzen (38), und Pixel von Löchern in zweiten Bildbereichen auf einen Wert von Pixeln in einem Bildbereich, der kein Loch darstellt, zu setzen (36), wobei die ersten Bildbereiche kleiner als die zweiten Bildbereiche sind.

Revendications

1. Procédé d'évaluation histologique du pléomorphisme nucléaire par identification de régions d'images correspondant potentiellement à des noyaux cellulaires dans des données d'image histologique, le procédé incluant également :

a) la décomposition (30) des données d'image couleur en un jeu de sous-images, chacune d'elles recouvrant ses voisines et l'élimination dans chacune des sous-images des régions aux bordures de la sous-image correspondant potentiellement à des noyaux cellulaires, l'Analyse en Composantes Principales (ACP) (30) étant utilisée pour transformer chacune des sous-images en données d'image en niveaux de gris comprenant un premier composant principal obtenu à partir d'une matrice, lequel est le produit d'une matrice de données d'image couleur et d'une matrice des vecteurs propres d'une matrice de covariance, les données d'image en niveaux de gris fournissant une définition d'image améliorée par comparaison à un plan individuel rouge, vert

ou bleu des données d'image couleur, et
b) le seuillage (32) des données d'image en niveaux de gris pour la rendre binaire, la détermination des périmètres et des surfaces des régions d'image identifiées, le calcul (42) des facteurs de forme des régions d'image à partir des périmètres et des surfaces et l'évaluation (44) du pléomorphisme à partir des paramètres statistiques des facteurs de forme.

2. Procédé selon la revendication 1, **caractérisé en ce que** l'ACP est utilisée pour transformer (30) les données d'image couleur de chacune des sous-images en données d'image en niveaux de gris par :

a) l'application de l'ACP à chacune des sous-images pour fournir une sous-image respective en niveaux de gris et l'élimination à partir des sous-images en niveaux de gris :

i) des régions d'image touchant ou coupant les limites de la sous-image,
ii) des régions d'image trop petites pour convenir, et
iii) des vides dans les régions d'image relativement grandes, et

b) le réassemblage des sous-images.

3. Procédé selon les revendications 1 ou 2, **caractérisé en ce que** les paramètres statistiques des facteurs de forme comprennent au moins un élément parmi la moyenne, la moyenne pondérée, la médiane, le mode, le maximum et le minimum.

4. Procédé selon les revendications 1 ou 2, **caractérisé en ce que** l'étape de seuillage des données de l'image est une étape (32) de seuillage Otsu, dans laquelle le choix du seuil a pour but de minimiser pour deux classes un rapport d'une variance intra-classe à une variance inter-classes.

5. Procédé selon les revendications 1 ou 2, **caractérisé en ce que** l'étape (44) d'évaluation du pléomorphisme détermine un pléomorphisme comme prenant une première valeur, une deuxième valeur ou une troisième valeur selon que la moyenne ou la médiane des facteurs de forme est respectivement inférieure à des valeurs d'une plage intermédiaire ou dans la plage intermédiaire ou supérieure à des valeurs de la plage intermédiaire.

6. Procédé selon la revendication 5, **caractérisé en ce qu'**un facteur de forme S pour une région d'image correspondant potentiellement à un noyau cellulaire est donné par $S=\frac{kP^2}{A}$, où k est une constante, P est le périmètre d'une région d'image et A est la surface de la région d'image, et un facteur de forme moyen S_m pour un ensemble de régions d'image correspondant potentiellement à un noyau cellulaire est seuillé (44) par $S_m \leq 30k$, $30k < S_m \leq 35k$ et $S_m > 35k$ correspondant respectivement à un pléomorphisme étant au-dessus, dans ou en dessous de la plage intermédiaire.

7. Procédé selon les revendications 1 ou 2, **caractérisé en ce que** l'étape d'identification des régions d'image correspondant potentiellement à un noyau cellulaire dans les données d'image histologiques inclut le filtrage des données d'image pour écrire sur les régions ne présentant pas d'intérêt en utilisant un processus de filtrage (38) qui supprime les régions dont la surface est inférieure au minimum défini par l'utilisateur.

8. Procédé selon la revendication 7, **caractérisé en ce que** l'étape de filtrage des données d'image pour écrire sur les régions ne présentant pas d'intérêt inclut l'attribution (38) aux premières régions d'image une valeur de pixel d'arrière-plan et l'attribution (36) aux pixels vides dans la seconde région d'image d'une valeur de pixel de région d'image non-vides, dans lesquelles les premières régions d'images sont plus petites que les secondes régions d'image.

9. Appareil pour l'évaluation histologique du pléomorphisme nucléaire par identification de régions d'image correspondant potentiellement à des noyaux cellulaires dans les données d'images histologiques, l'appareil comprenant un ordinateur programmé pour :

a) décomposer (30) les données d'image couleur en un ensemble de sous-images, chacune d'elles recouvrant ses voisines et éliminer de chacune des sous-images les régions aux bordures de la sous-image correspondant

potentiellement à des noyaux cellulaires,
b) utiliser l'ACP (30) pour transformer les données d'image couleur de chacune des sous-images en données d'image en niveaux de gris comprenant une première composante principale obtenue à partir d'une matrice, laquelle est le produit d'une matrice de données d'image couleur et une matrice de vecteurs propres d'une matrice de covariance, les données d'image en niveaux de gris fournissant une définition d'image améliorée par comparaison à un plan individuel rouge vert ou bleu des données d'image couleur, et
c) seuiller (32) les données d'image en niveaux de gris pour la rendre binaire, déterminer les périmètres et les surfaces des régions d'image identifiées, calculer (42) les facteurs de forme des régions d'image à partir des périmètres et des surfaces et évaluer (44) le pléomorphisme à partir des paramètres statistiques des facteurs de forme.

10. Appareil selon la revendication 9 **caractérisé en ce que** l'ordinateur est programmé pour utiliser l'ACP pour :

a) transformer (30) les données d'image couleur de chacune des sous-images en données d'image en niveaux de gris en appliquant l'ACP à chacune des sous-images pour fournir une sous-image respective en niveaux de gris,
b) supprimer des sous-images en niveaux de gris :

i) des régions d'image touchant ou coupant les limites de la sous-image,
ii) des régions d'image trop petites pour convenir, et
iii) des vides dans des régions d'image relativement grandes, et

c) réassembler les sous-images.

11. Appareil selon les revendications 9 ou 10, **caractérisé en ce que** les paramètres statistiques des facteurs de forme comprennent au moins un élément parmi la moyenne, la moyenne pondérée, la médiane, le mode, le maximum et le minimum.

12. Appareil selon les revendications 9 ou 10, **caractérisé en ce que** l'ordinateur est programmé pour seuiller les données d'image en utilisant le seuillage Otsu dans lequel le choix du seuil a pour but de minimiser pour deux classes un rapport d'une variance intra-classe à une variance inter-classes.

13. Appareil selon les revendications 9 ou 10, **caractérisé en ce que** l'ordinateur est programmé pour déterminer un pléomorphisme (44) comme prenant une première valeur, une deuxième valeur ou une troisième valeur selon que la moyenne ou la médiane des facteurs de forme est respectivement inférieure à des valeurs d'une plage intermédiaire ou dans la plage intermédiaire ou supérieure à des valeurs de la plage intermédiaire.

14. Appareil selon les revendications 9 ou 10 **caractérisé en ce que** l'ordinateur est programmé pour déterminer un facteur de forme S pour une région d'image correspondant potentiellement à un noyau cellulaire, le facteur de forme S étant donné par $S=\frac{kP^2}{A}$, où k est une constante, P est le périmètre d'une région d'image et A est la surface de la région d'image, et l'ordinateur est également programmé pour déterminer (42) un facteur de forme moyen S_m pour un jeu de régions d'image correspondant potentiellement à un noyau cellulaire, pour seuiller le facteur de forme moyen par $S_m \leq 30k$, $30k < S_m \leq 35k$ et $S_m > 35k$ correspondant respectivement à un pléomorphisme étant au-dessus, dans ou en dessous de la plage intermédiaire.

15. Appareil selon les revendications 9 ou 10, **caractérisé en ce que** l'ordinateur est programmé pour attribuer (38) à des premières régions d'image une valeur de pixel d'arrière-plan et d'attribuer (36) à des pixels vides dans des secondes régions d'image une valeur de pixel de région d'image non-vide, dans laquelle les premières régions d'image sont plus petites que les secondes régions d'image.

16. Programme informatique pour utilisation en évaluation histologique de pléomorphisme nucléaire et ayant des instructions pour le contrôle d'un ordinateur pour identifier des régions d'image correspondant potentiellement à des noyaux cellulaires dans des données d'image histologiques, le logiciel ayant également des instructions pour le contrôle d'un ordinateur pour :

a) décomposer (30) les données d'image couleur en un jeu de sous-image chacune d'elles recouvrant ses voisines et éliminer de chacune des sous-images des régions aux bordures de la sous-image correspondant potentiellement à des noyaux cellulaires,
b) utiliser l'ACP (30) pour transformer les données d'image couleur de chacune des sous-images en données d'image en niveaux de gris comprenant une première composante principale obtenue à partir d'une matrice laquelle est le produit d'une matrice de données d'image couleur et une matrice de vecteurs propres d'une matrice de covariance, les données d'image en niveaux de gris fournissant une définition d'image améliorée par comparaison à un plan individuel rouge vert ou bleu des données d'image couleur, et
c) seuiller (32) les données d'image en niveaux de gris pour la rendre binaire, déterminer les périmètres et les surfaces des régions d'image identifiées, calculer (42) les facteurs de forme des régions d'image à partir des périmètres et des surfaces et évaluer (44) le pléomorphisme à partir des paramètres statistiques des facteurs de forme.

17. Programme informatique selon la revendication 16 **caractérisé en ce qu'**il comprend des instructions pour le contrôle d'un ordinateur pour utiliser l'ACP pour :

a) transformer (30) les données d'image couleur de chacune des sous-images en données d'image en niveaux de gris en appliquant l'ACP à chacune des sous-images pour fournir une sous-image respective en niveaux de gris,
b) éliminer des sous-images en niveaux de gris :

i) des régions d'image touchant ou coupant les limites de la sous-image,
ii) des régions d'image trop petites pour convenir, et
iii) des vides dans les régions d'image relativement grandes, et

c) réassembler les sous images.

18. Programme informatique selon les revendications 16 ou 17, **caractérisé en ce que** les paramètres statistiques des facteurs de forme comprennent au moins un élément parmi la moyenne, la moyenne pondérée, la médiane, le mode, le maximum et le minimum.

19. Programme informatique selon les revendications 16 ou 17, **caractérisé en ce qu'**il comprend des informations pour contrôler un ordinateur pour seuiller les données d'image en utilisant le seuillage Otsu dans lequel le choix du seuil a pour but de minimiser pour deux classes un rapport d'une variance intra-classe à une variance inter-classes.

20. Programme informatique selon les revendications 16 ou 17, **caractérisé en ce qu'**il a des instructions pour contrôler un ordinateur pour déterminer (44) un pléomorphisme comme ayant une première valeur, une deuxième valeur ou une troisième valeur selon que la moyenne ou la médiane des facteurs de forme est respectivement inférieure à des valeurs d'une plage intermédiaire ou dans la plage intermédiaire ou supérieure à des valeurs de la plage intermédiaire.

21. Programme informatique selon la revendication 20 **caractérisé en ce qu'**il comprend des instructions pour contrôler un ordinateur pour :

a) déterminer (42) un facteur de forme S pour une région d'image correspondant potentiellement à un noyau cellulaire, le facteur de forme S étant donné par $S=\frac{kP^2}{A}$, où k est une constante, P est le périmètre d'une région d'image et A est la surface de la région d'image,
b) seuiller (44) un facteur de forme moyen pour S_m pour un jeu de régions d'image correspondant potentiellement à un noyau cellulaire par $S_m \leq 30k$, $30k < S_m \leq 35k$ et $S_m > 35k$, et
c) indiquer un pléomorphisme comme respectivement relativement faible, modéré ou élevé pour un jeu de régions d'image correspondant potentiellement à des noyaux cellulaires seuillés (44) par $S_m \leq 30k$, $30k < S_m \leq 35k$ et $S_m > 35k$ correspondant respectivement à un pléomorphisme étant au-dessus, dans ou en dessous de la plage intermédiaire.

22. Programme informatique selon les revendications 16 ou 17 **caractérisé en ce qu'**il inclut des instructions pour attribuer (38) à des premières régions d'image une valeur de pixel d'arrière-plan et attribuer (36) à des pixels vides

dans des secondes régions d'image une valeur de pixel de région d'image non-vide, dans laquelle les premières régions d'image sont plus petites que les secondes régions d'image.

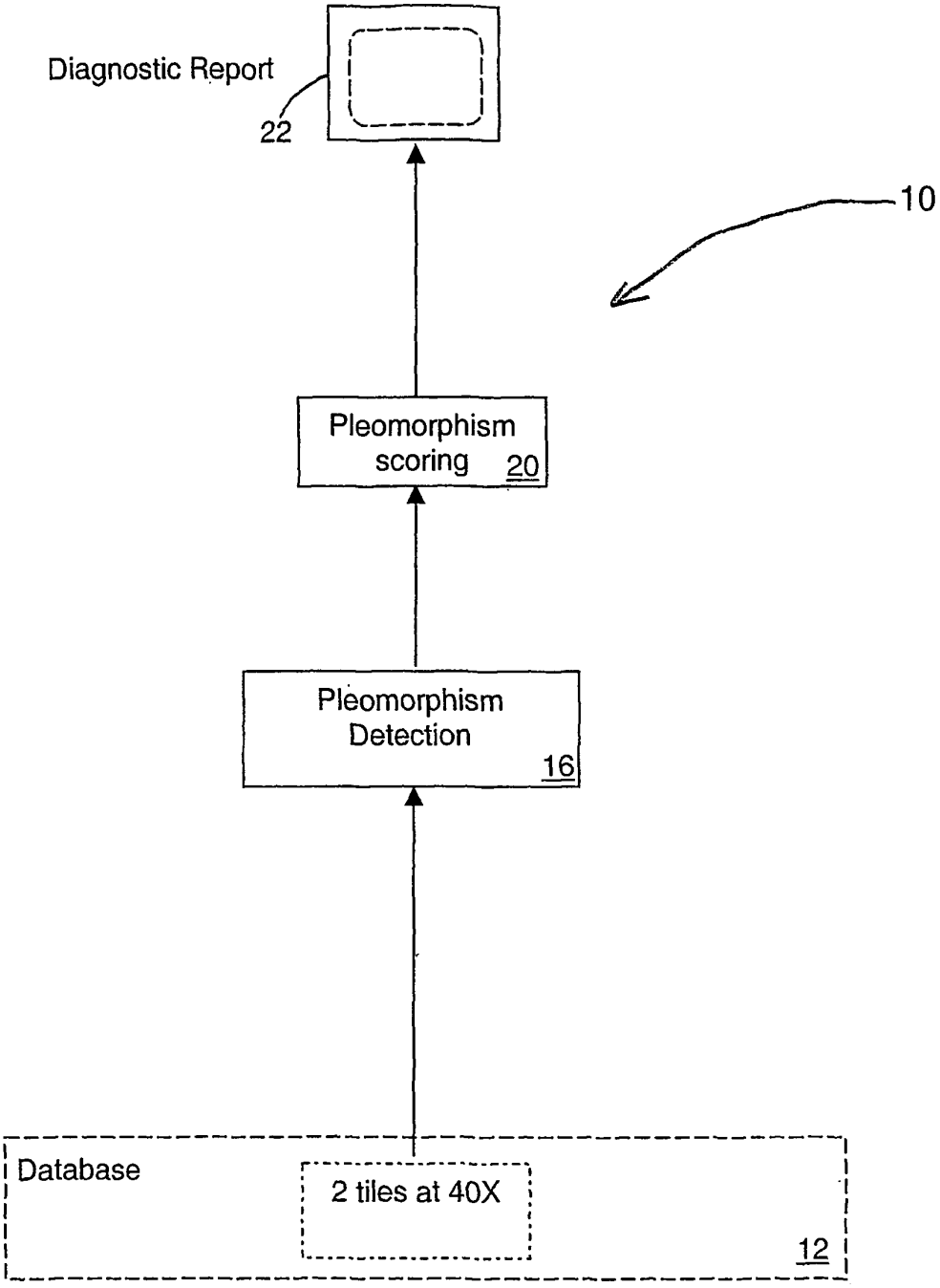
Diagnostic Report
22
10
Pleomorphism scoring 20
Pleomorphism Detection 16
Database
2 tiles at 40X
12

Figure 1

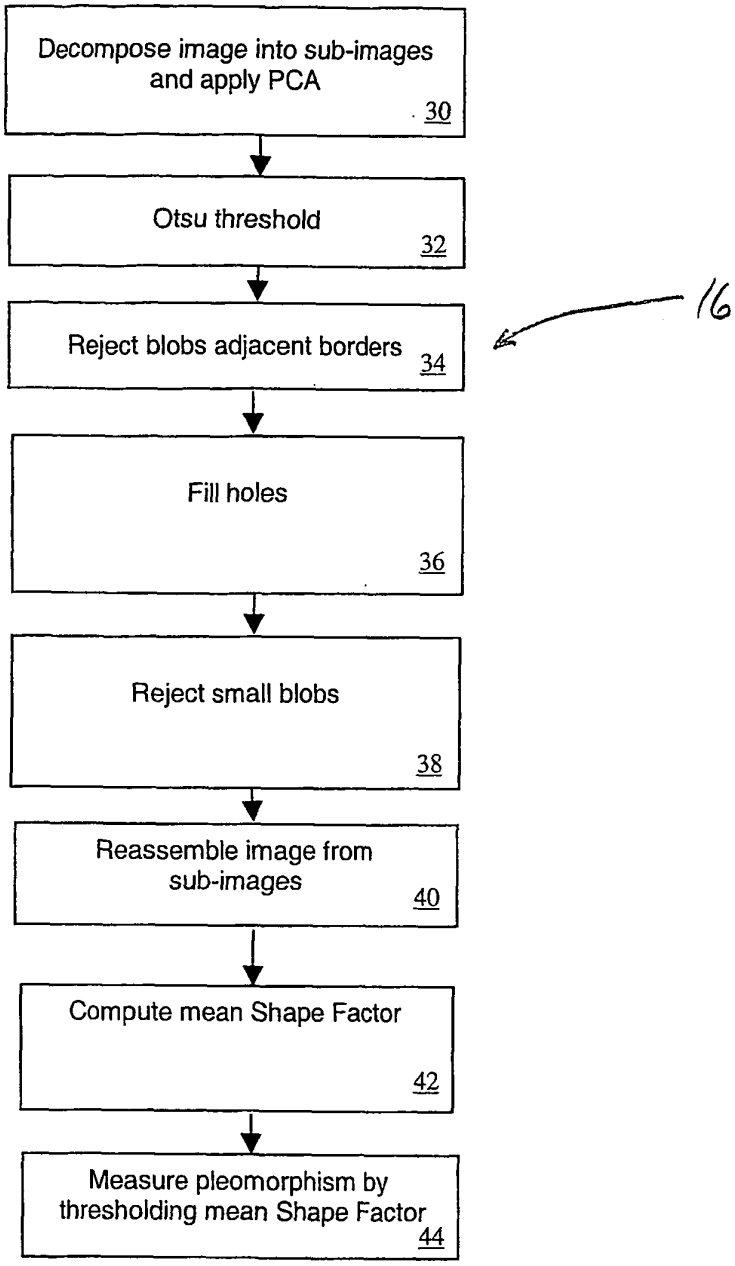

Figure 2 Nuclear Pleomorphism Feature Detection