Office de la Propriété Intellectuelle du Canada
Un organisme d'Industrie Canada

Canadian Intellectual Property Office
An agency of Industry Canada

CA 2086785 C 2008/10/07

(11)(21) 2 086 785

(12) BREVET CANADIEN
CANADIAN PATENT

(13) C

(22) Date de dépôt/Filing Date: 1993/01/06

(41) Mise à la disp. pub./Open to Public Insp.: 1994/04/15

(45) Date de délivrance/Issue Date: 2008/10/07

(30) Priorité/Priority: 1992/10/14 (US07/961,596)

(51) Cl.Int./Int.Cl. *G01N 15/14* (2006.01), *G01N 33/50* (2006.01), *G01N 1/30* (2006.01)

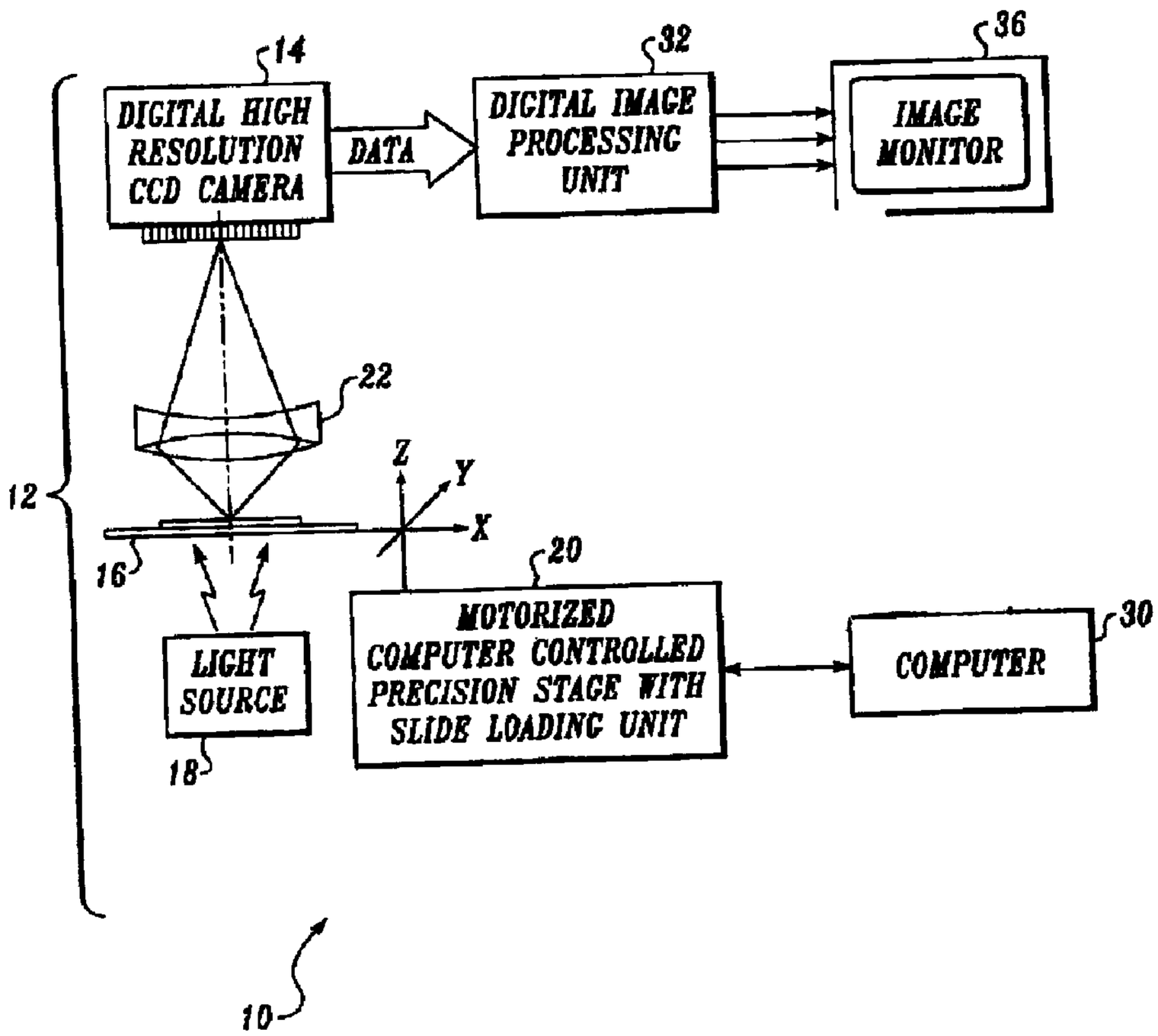
(72) Inventeurs/Inventors:
JAGGI, BRUNO, CA;
HARRISON, ALAN, CA;
PALCIC, BRANKO, CA;
GARNER, DAVID, CA;
MACAULAY, CALUM, CA

(73) Propriétaire/Owner:
BRITISH COLUMBIA CANCER AGENCY BRANCH, CA

(74) Agent: SMART & BIGGAR

(54) Titre : DETECTION AUTOMATIQUE DU TISSU CANCEREUX OU PRECANCEREUX A L'AIDE D'UNE MESURE DES CHANGEMENTS ASSOCIES AUX TUMEURS MALIGNES

(54) Title: AUTOMATED DETECTION OF CANCEROUS OR PRECANCEROUS TISSUE BY MEASURING MALIGNANCY ASSOCIATED CHANGES (MACs)



(57) Abrégé/Abstract:
An apparatus and a method for detecting malignancy associated changes relies on use of the microscope having a high sampling density light transducer. Images of the cell nuclei are obtained in precise focus. The images are segmented using a relocation

CA 2086785 C 2008/10/07

(11)(21) 2 086 785

(13) C

(57) Abrégé(suite)/Abstract(continued):
algorithm for precisely locating the edge of the nucleus. The features of the images including DNA distribution are analyzed using multivariate analysis to detect malignancy associated changes.

ABSTRACT

AUTOMATED DETECTION OF CANCEROUS OR PRECANCEROUS TISSUE BY MEASURING MALIGNANCY ASSOCIATED CHANGES (MACs)

An apparatus and a method for detecting malignancy associated changes relies on use of the microscope having a high sampling density light transducer. Images of the cell nuclei are obtained in precise focus. The images are segmented using a relocation algorithm for precisely locating the edge of the nucleus. The features of the images including DNA distribution are analyzed using multivariate analysis to detect malignancy associated changes.

TITLE OF THE INVENTION

AUTOMATED DETECTION OF CANCEROUS OR PRECANCEROUS TISSUE BY MEASURING MALIGNANCY ASSOCIATED CHANGES (MACs)

FIELD OF THE INVENTION

The invention relates to apparatus and a method for automatically detecting the presence of a cancerous or precancerous lesion in a tissue by measuring the characteristics of the cell nuclei of ostensibly normal tissue.

BACKGROUND OF THE INVENTION

For the last five decades, attempts have been made to develop a device which can automatically detect cancerous or precancerous cells deposited on a microscopy slide. For example, research and development has been undertaken by numerous groups who have been attempting to develop a fully automated prescreening device for cervical samples. In all cases, these devices attempt to perform tasks similar to those performed by human experts, i.e. scanning the entire slide at coarse and fine spatial resolutions and seeking and characterizing cancerous or precancerous cells which are exfoliated from a cancerous or precancerous lesion.

Prior art devices use nuclear and cytoplasmic features such as nuclear size, nuclear shape, nuclear to cytoplasmic ratio, DNA amount in the nucleus and DNA distribution in the nucleus to identify cancerous or precancerous cells. Many different approaches have been tried. To date no effective

system is commercially available although several groups have claimed to have achieved a semi-automated version with good to excellent results. Such systems are typically composed of an automated image cytometry device having a microscope equipped with a light transducer, often a video camera, a motorized x, y, z stage under computer control, a computer with imaging board(s), and peripheral devices such as video monitors, printers and input devices. Algorithms have also been previously developed for the automated recognition of cells and nuclei, for segmentation of areas of interest, automated focus and other functions. There are a variety of approaches by which such systems operate to recognize cancerous or precancerous cells including multivariate analysis (e.g. discriminant function analysis), decision trees, and neural networks.

All systems available or under development today rely on the detection and exact characterization of at least one fully cancerous or precancerous cell. This requires very high precision, sensitivity and specificity. The known systems are plagued with the problem of inadequacy of the artifact rejection algorithms developed to date and the only way around the problem is to use human experts to differentiate between true cancerous or precancerous cells and artifacts resulting in only a semi-automated approach.

These efforts have therefore not resulted in reliable detection systems. It has been well established, for example, that the false negative rate for cervical samples is between 7 -

10% even at the best cytology screening laboratories. Additional false negative results arise even where neither the screening cytotechnicians and cytopathologists nor their equipment are at fault. This arises when the sample to be examined is taken elsewhere than from the lesion itself, for example from an area slightly removed from the spot where cancerous or precancerous growth was present. The false negative rate due to such sampling error is even greater and has been estimated to be between 10 - 20%.

It has been reported for several tissues (e.g. cervix, colon) that very careful measurements of nuclear features of ostensibly normal cells growing in the vicinity of the cancerous growth, show slight differences from the nuclear features of truly normal cells, i.e. from the cells of a normal, healthy individual without cancerous growth in that tissue. The changed nuclear features of ostensibly normal cells growing in the vicinity of the cancerous or precancerous lesion compared to those of normal cells from the same type of tissue of a normal, healthy individual are referred to in the literature as Malignancy Associated Changes (MACs). MACs particularly manifest in those features, particularly those describing the distribution of the genetic material in the nuclei of these cells. Although no individual feature is sufficiently discriminating to be able to distinguish between healthy individuals and those harbouring the cancerous growth, a combination of many features in a multivariate analysis has been found to provide adequate separation between such individuals.

However, no fully automated system has been developed which can perform such detection. The data (images) of cells must be obtained at the highest spatial and photometric resolution, at the precise focus and exact segmentation of the areas of interest, i.e. the nucleus. It has often been said that malignancy associated changes could never be used for automated prescreening due to these requirements. It has been necessary for a highly trained expert to identify individual cells or nuclei to be examined, due to poor artifact rejection and the limited segmentation capabilities of prior art systems. In addition, a very large number of such cell images must be analyzed to achieve reasonable results, typically over 200 cell images per sample.

Such a procedure is not only time consuming and tedious, it is also impractical for prescreening as it takes typically up to several hours to accumulate, capture and analyze sufficient numbers of cell images.

This invention describes an apparatus and a method for measuring MACs in a fully automated way, using only images of the nuclei of ostensibly normal cells such that several hundred cell images are used in the analysis lasting only a few minutes.

SUMMARY OF THE INVENTION

Accordingly, the present invention provides a method of detecting malignancy-associated changes in a cell sample, comprising the steps of:

obtaining a cell sample;

staining the sample to identify cell nuclei within the sample;

obtaining an image of the cell sample with a microscope of the type that includes a slide stage;

focusing the image;

identifying objects in the image, each of the objects having an edge that separates the object from the background

calculating a set of feature values for each object; and

analyzing the feature values to determine whether each object is a cell nucleus having malignancy-associated changes.

The present invention also provides a method for detecting malignancy-associated cells in a cell sample, comprising the steps of:

obtaining a cell sample;

fixing the cells of the cell sample;

staining the cells to identify cell nuclei in the cell sample;

illuminating the sample and obtaining an image of the sample with a microscope;

analyzing the image to detect objects of interest;

determining a focus setting for each object of interest and obtaining an image of each object of interest;

calculating an edge that bounds each object of interest;

calculating a set of feature values for each object of interest; and

comparing the set of calculated feature values to a known feature values for malignancy-associated and normal objects.

BRIEF DESCRIPTION OF THE DRAWINGS

Aspects of the invention are illustrated, merely by way of example, in the enclosed Figures in which:

Figure 1 is a block diagram of the MAC detection system according to the present invention.

Figure 2 is a flow diagram of an example method according to the present invention.

Figure 3 is a flow diagram related to an example comparison of object feature values.

DETAILED DESCRIPTION OF THE INVENTION

The invention may be more fully appreciated by reference to the preferred embodiment thereof which relates to detection of malignancy associated changes in cervical samples.

A block diagram of the MAC detection system according to the present invention is shown in FIG. 1.

The flow diagram of Figure 2 represents a method of using the system 10 to detect cancerous or precancerous tissue by measuring malignancy associated changes.

The samples are first treated by stoichiometric staining of the nuclear (DNA) material. We achieved this by developing a modified Feulgen procedure using Thionin although other stains such as stoichiometric absorbance stains (e.g. gallocyanin, azure-A, etc.) or stoichiometric fluorescence stains (e.g.) DAPI, propidium iodid, etc.) can also be used.

Typically, a 40x objective would be used to view cells of cervical samples. However, we have found that a 40x objective results in difficulties in focusing and therefore in segmenting images of the nuclei due to the distortion effect in the lenses. In addition, the 40x objective is bulky and risks hitting the slide cover slip as well as having a shallow depth of focus. Accordingly, the preferred embodiment of the invention uses a 20x/.75 objective together with a high sampling density sensor mounted on a microscope.

A computer is provided to control a motorized microscope stage and to perform all software and output functions.

A light transducer is used, consisting of a scientific Charge Coupled Device with a 100% fill factor, over 500 gray level photometric resolution with a pixel size of 0.3 μm or less. The transducer is mounted in the primary image plane or in another camera port of the microscope providing negligible image distortion, and an imaging board is provided for capturing and processing the images of cell nuclei.

In order to detect MACs, the device must be capable of capturing images in precise and reproducible focus. This focus is achieved by maximizing contrast of the nuclear material as a function of focal position. Given the pixel size and spatial resolution, the image of a typical nucleus comprises several hundred individual pixels of 3 - 7 μm in diameter. This is achieved with the 20x objective with the transducer placed in the primary image plane of the objective.

In addition, all images must be segmented exactly such that all pixels covering the nucleus belong to the mask. The segmentation is an important step. A simple thresholding on an image is obtained from a calibrated image (corrected for lens, illumination, dark current of camera, and other imperfections) and is corrected by the absorbance of material around the nucleus (i.e. cytoplasm) which is assumed to be present in about equal

amounts over the nucleus. The thresholded mask edge represents the first approximation of the real edge of the nucleus. The latter is obtained by an edge relocation algorithm which operates by dilating the approximate edge and then step by step removing the pixels with the least gradient such as not to break a continuous contour around the nucleus.

Algorithms selecting only images of the nuclei of normal cells are used to ensure that only a few percent (typically 1 - 2%) of artifacts are present in the population of the ostensibly normal cells. This is achieved by discriminant function analysis and a decision tree process, but could possibly also be achieved by other statistical or neural network procedures, as will be appreciated by those skilled in the art of automated image classification.

Various discriminant features of the segmented digital images are then analyzed by the computer.

For the MAC approach to work, depending on the cell type many nuclear features must be employed, but particularly the DNA distribution of the cell nuclei. We typically measure over 100 nuclear features of which about 30 are then used in the multivariate analysis. For best results, the nuclear mask is adjusted depending on features. For example, for features describing the frequency content of the optical density in the nucleus the pixels covering the edge of the nucleus must be removed as otherwise these features lose (or are diminished in) their discriminant power. Conversely, other features require all edge pixels, even those catching only a fraction of the material in the

edge, to be included (e.g. integrated optical density). Therefore, the nuclear mask is adjusted individually for all critical features by either appropriate erosion or dilation algorithm of the best (highest gradient) edge.

Multivariate analysis may then be performed by the computer on the features of the segmented images. In the preferred embodiment, the multivariate analysis comprises cluster analysis leading to a decision tree made up of thresholds and discriminant functions. Alternatively, it can consist of a neural network. We have found that, using the above approach, reliable characterization of MACs has been achieved.

When all the above is achieved, MAC values of ostensibly normal cells are obtained which are significantly different form tissues of those individuals who are harbouring cancerous (such as carcinoma in situ, microinvasive or invasive cancer) or precancerous lesion (such as moderate and severe dysplasia). Several tissues have been examined by this approach (e.g. cervix, lung) and very high sensitivity and specificity can be achieved approaching or even surpassing those achieved by human experts using atypical cell detection approach.

It will be appreciated by those skilled in the art that modifications and variations from the preferred embodiment may be practised without the parting form the principals of the invention.

The embodiments of the invention in which an exclusive property or privilege is claimed are as follows:

1. A method of detecting malignancy-associated changes in a cell sample, comprising the steps of:

obtaining a cell sample;

staining the sample to identify cell nuclei within the sample;

obtaining an image of the cell sample with a microscope of the type that includes a slide stage;

focusing the image;

identifying objects in the image, each of the objects having an edge that separates the object from the background;

calculating a set of feature values for each object; and

analyzing the feature values to determine whether each object is a cell nucleus having malignancy-associated changes.

2. A method for detecting malignancy-associated cells in a cell sample, comprising the steps of:

obtaining a cell sample;

fixing the cells of the cell sample;

staining the cells to identify cell nuclei in the cell sample;

illuminating the sample and obtaining an image of the sample with a microscope;

analyzing the image to detect objects of interest;

determining a focus setting for each object of interest and obtaining an image of each object of interest;

calculating an edge that bounds each object of interest;

calculating a set of feature values for each object of interest; and

comparing the set of calculated feature values to a known feature values for malignancy-associated and normal objects.

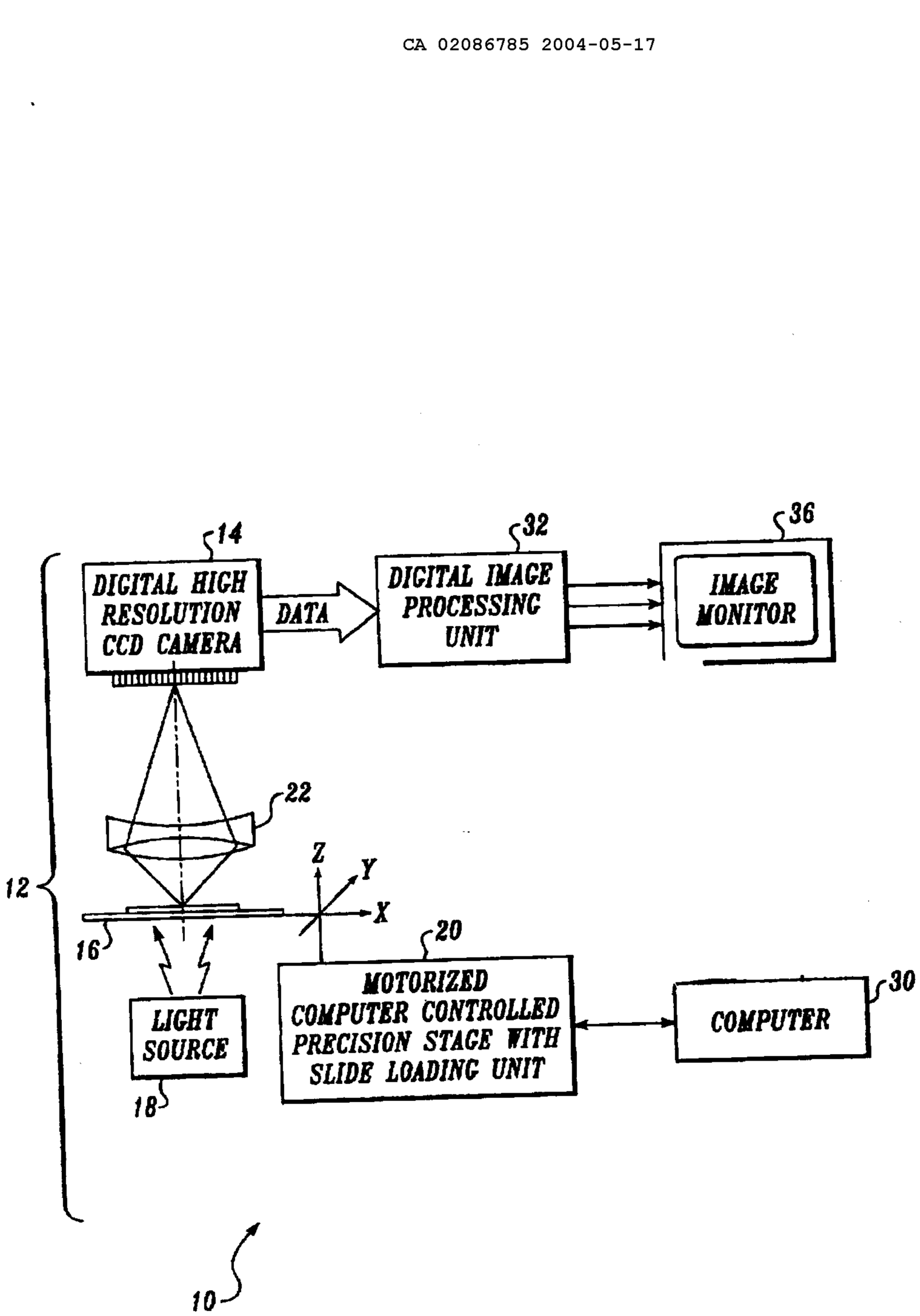

14
DIGITAL HIGH RESOLUTION CCD CAMERA
DATA
32
DIGITAL IMAGE PROCESSING UNIT
36
IMAGE MONITOR
22
12
Z
Y
X
16
20
MOTORIZED COMPUTER CONTROLLED PRECISION STAGE WITH SLIDE LOADING UNIT
COMPUTER
30
LIGHT SOURCE
18
10

Fig. 1

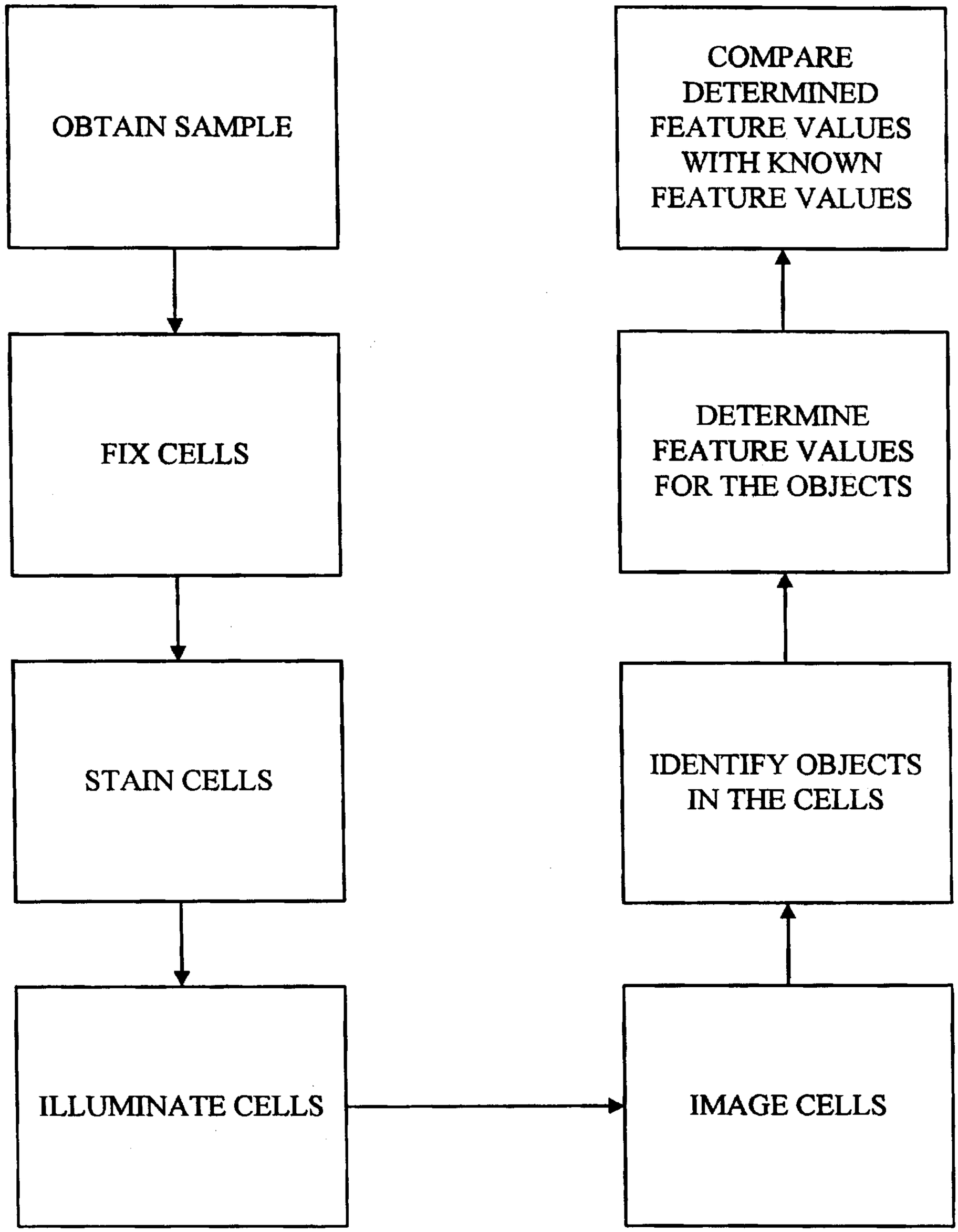

OBTAIN SAMPLE
FIX CELLS
STAIN CELLS
ILLUMINATE CELLS
IMAGE CELLS
IDENTIFY OBJECTS IN THE CELLS
DETERMINE FEATURE VALUES FOR THE OBJECTS
COMPARE DETERMINED FEATURE VALUES WITH KNOWN FEATURE VALUES

Fig. 2

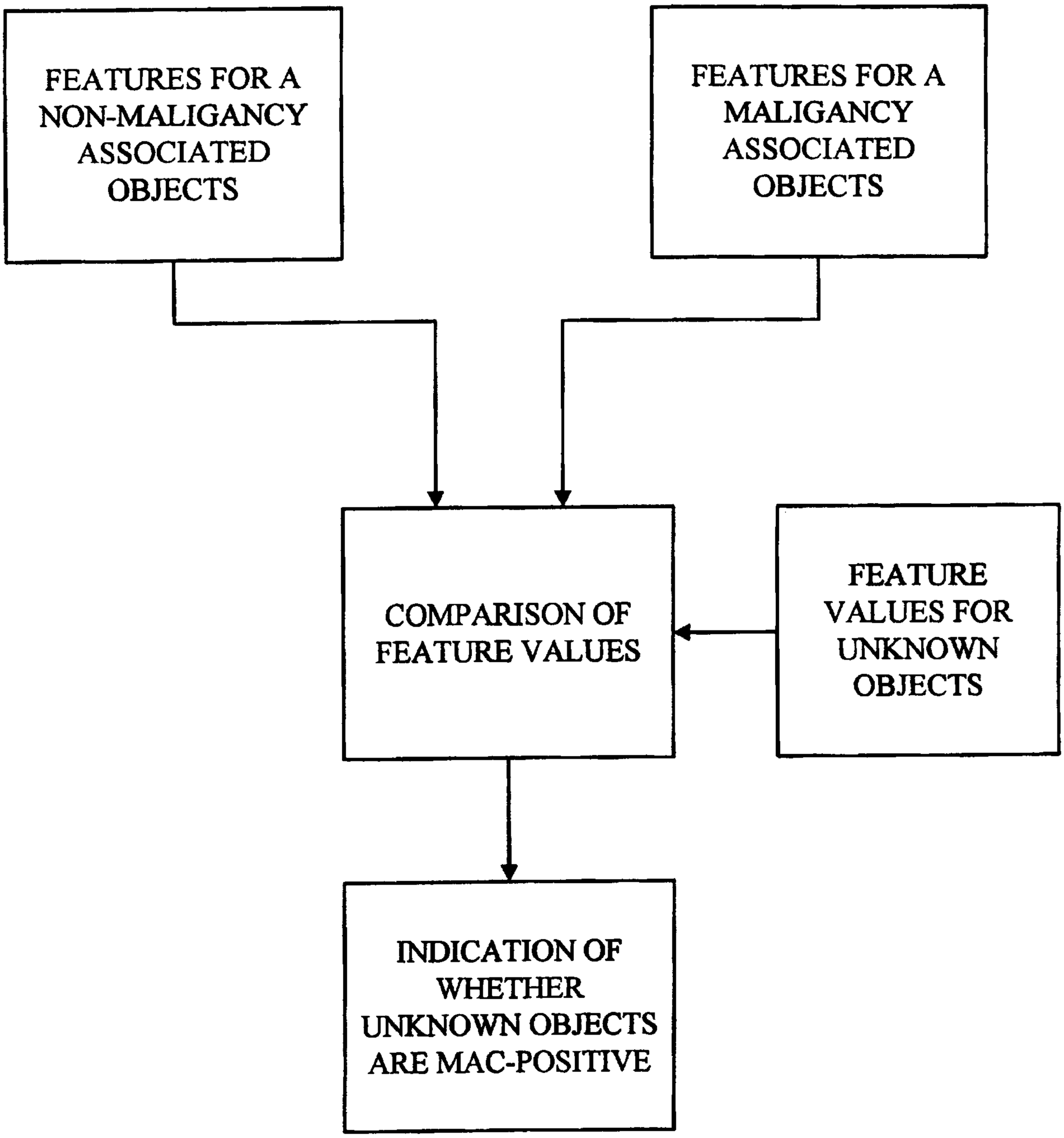
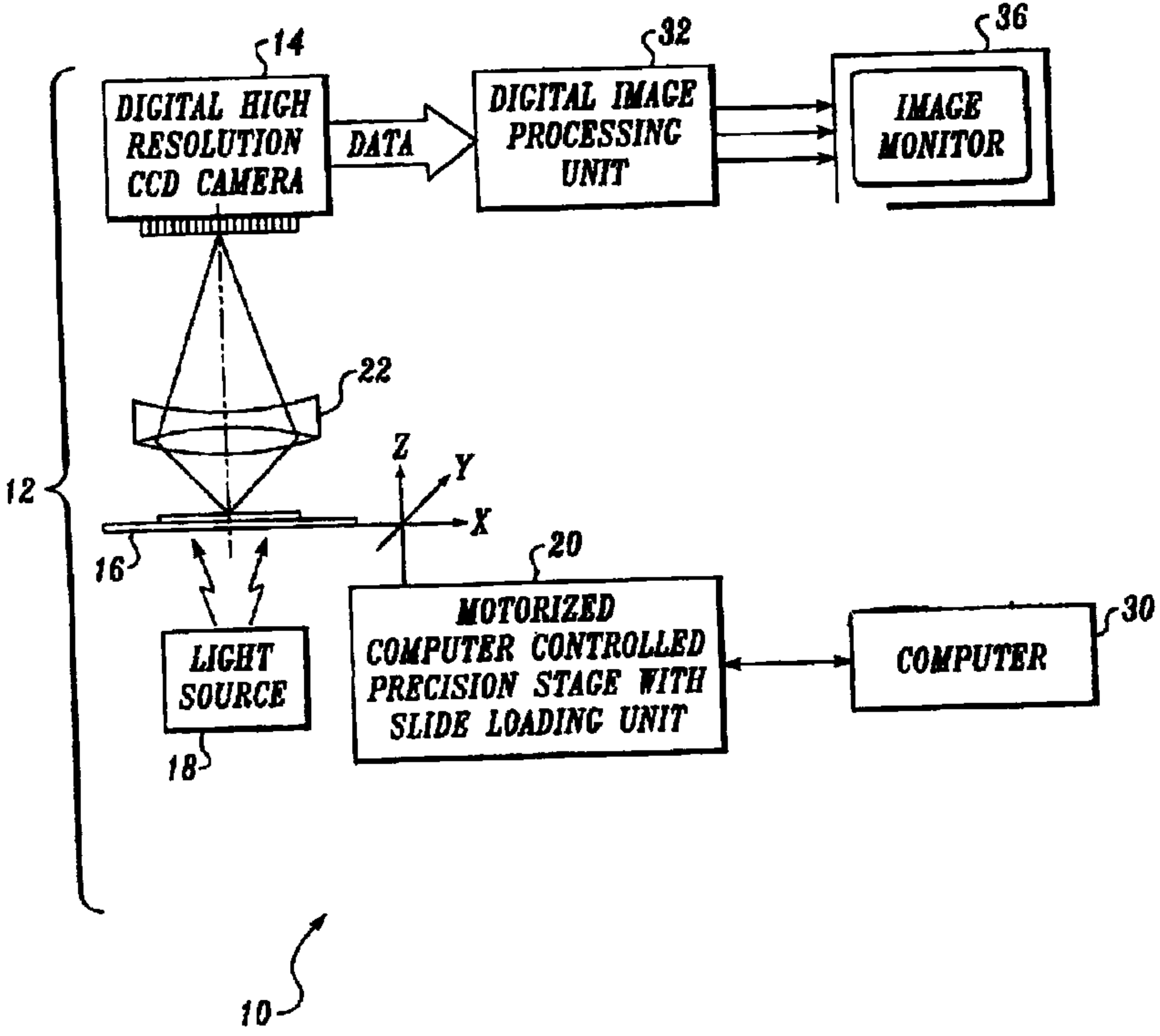

FEATURES FOR A NON-MALIGANCY ASSOCIATED OBJECTS
FEATURES FOR A MALIGANCY ASSOCIATED OBJECTS
COMPARISON OF FEATURE VALUES
FEATURE VALUES FOR UNKNOWN OBJECTS
INDICATION OF WHETHER UNKNOWN OBJECTS ARE MAC-POSITIVE

Fig. 3


14
DIGITAL HIGH RESOLUTION CCD CAMERA
DATA
32
DIGITAL IMAGE PROCESSING UNIT
36
IMAGE MONITOR
22
Z
Y
X
12
16
20
LIGHT SOURCE
18
MOTORIZED COMPUTER CONTROLLED PRECISION STAGE WITH SLIDE LOADING UNIT
COMPUTER
30
10