(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
28 August 2014 (28.08.2014)

WIPO | PCT

(10) International Publication Number
WO 2014/131013 A1

(51) International Patent Classification:
G01N 1/28 (2006.01) *G06K 9/36* (2006.01)
G01N 1/30 (2006.01) *G06K 9/54* (2006.01)
G01N 1/42 (2006.01) *G06K 9/62* (2006.01)
G06K 9/68 (2006.01)

(21) International Application Number:
PCT/US2014/018322

(22) International Filing Date:
25 February 2014 (25.02.2014)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/769,049 25 February 2013 (25.02.2013) US
13/934,237 3 July 2013 (03.07.2013) US

(71) Applicant: **FLAGSHIP BIOSCIENCES LLC** [US/US]; 2995 Wilderness Place, Ste 105, Boulder, Colorado 80301 (US).

(72) Inventors: **LANGE, Holger**; Enger (DE). **KRUEGER, Joseph**; Boulder, Colorado (US). **YOUNG, G. David**; Boulder, Colorado (US). **JOHNSON, Trevor**; San Diego, California (US). **VOELKER, Frank**; Boulder, Colorado (US). **POTTS, Steven**; San Diego, California (US).

(74) Agents: **COASTAL PATENT LAW GROUP, P.C.** et al.; P.O. Box 131299, Carlsbad, California 92013 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: CELL-BASED TISSUE ANALYSIS

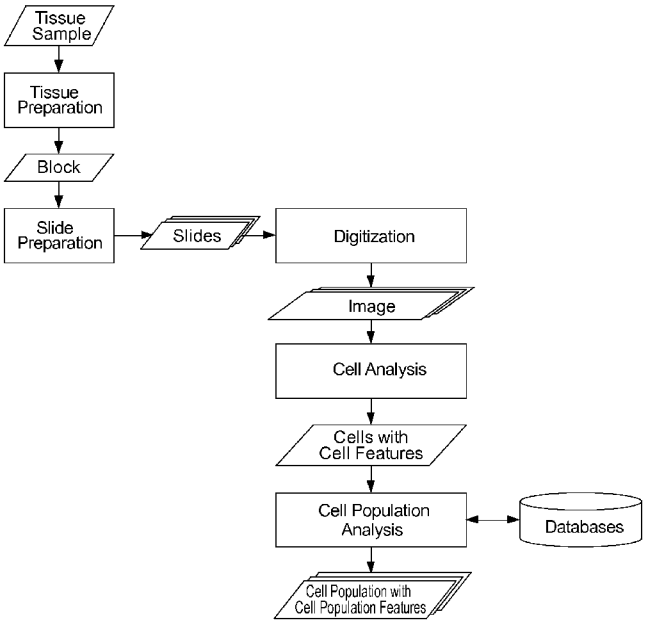


Fig. 1

(57) Abstract: Methods for cell-based tissue analysis utilize modern trends in digital microscopy to obtain, process, calibrate, and analyze digital images of tissue sections to quantify cell-based data for improved histological analysis. Using data from multiple images of a common tissue section, or data from images of multiple tissue sections, additional degrees of freedom are realized and the resulting analysis provides added depth to histological analysis of tissue samples. With computerized analytical methods, speed and accuracy of histological analysis is greatly improved.

WO 2014/131013 A1

Published:

— *with international search report (Art. 21(3))*

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

CELL-BASED TISSUE ANALYSIS

TECHNICAL FIELD

[0001] This invention relates generally to medical imaging; and more particularly, to methods for computerized analysis of microscopic images of tissue sections.

BACKGROUND ART

[0002] As before in radiology, now with the digitization of pathology, with more precise imaging of histology slides, new computer-assisted methods can be used that go far beyond the ability of human evaluation and interpretation of a glass slide using optical microscopy.

[0003] Sophisticated image analysis programs can be used to detect and characterize cells on partial or entire tissue sections and define and characterize different cell populations of a tissue specimen. Tissue specimens can then be examined and evaluated based on those cell populations in the context of other tissue specimens (study data set, target patient population data sets, etc.). Ultimately, a pathologist or tissue analyst can use the data from this approach in making assessments on endpoint determinants of the evaluation.

[0004] Similar concepts are used in flow cytometry, where multi-dimensional information is captured on every cell identified, and the dimensions are gated and compared to each other to define specific attributes of cell features. In contrast to flow cytometry, which relies on the physical capture and special instrumentation for the analysis of dissociated cells, image-based cell sorting does not rely on the physical capture and analysis of each cell as it is done digitally. Also, analysis of cells in their tissue section context has the advantage of maintaining the information of structural morphology of the tissue and context of the cells, while not requiring a special instrument.

[0005] There is a present and continuing need for systems and methods to build an image-based tissue analysis tool, which can be used to study multiple cell dimensions within a tissue to identify relationships between cells within and between tissue sections.

SUMMARY OF THE INVENTION

Technical Problem

[0006] Conventional methods for examining tissue specimens based on histology slides are limited to the ability of human evaluation and interpretation using a microscope.

Solution to Problem

[0007] In accordance with the embodiments herein, novel methods for examining tissue specimens based on histology slides are provided that go far beyond the ability of human evaluation and interpretation using an optical microscope, the methods generally comprise: detecting and characterizing cells on partial or entire tissue sections, defining and characterizing different cell populations of a tissue specimen, and examining tissue specimens based on cell population characteristics.

BRIEF DESCRIPTION OF THE DRAWINGS

[0008] Certain embodiments are described with reference to the following figures, wherein:

[0009] Fig. 1 shows a flowchart of the entire cell-based tissue analysis process from a tissue sample to the cell population analysis.

[0010] Fig. 2 illustrates the color calibration profile generation from a stain specific color calibration target.

[0011] Fig. 3 shows the data hierarchy associated with the cell analysis.

[0012] Fig. 4 illustrates the display of cell representations with intensity-coded cell features overlaid on an image.

[0013] Fig. 5 illustrates the display of cell feature data plots.

[0014] Fig. 6 illustrates the display of cell feature data tables.

[0015] Fig. 7 illustrates the display of cell-based heatmaps overlaid on an image.

[0016] Fig. 8(A-B) illustrates the interactive linkage of the cells between different data representations.

[0017] Fig. 9 illustrates the color calibration of images taken from different instruments using a color calibration profile.

[0018] Fig. 10 illustrates the calibration of the cell analysis to the staining process using tissue-based or equivalent controls.

[0019] Fig. 11 shows the data hierarchy associated with a multiplexed cell analysis.

[0020] Fig. 12(A-D) illustrates the mapping of cell features across different tissue sections using a common grid.

[0021] Fig. 13(A-B) illustrates the mapping of cell features across different tissue sections using virtual cells.

[0022] Fig. 14 shows the data hierarchy associated with the cell population analysis.

[0023] Fig. 15 illustrates a linear regression analysis of two cell features.

[0024] Fig. 16 illustrates an analysis based on cells from different tissue samples.

[0025] Fig. 17 illustrates the creation of a cell annotation.

[0026] Fig. 18 illustrates the creation of a new cell feature.

[0027] Fig. 19(A-C) illustrates the gate definition process using representative cells for the different cell populations.

[0028] Fig. 20(A-B) illustrates the pattern recognition process using representative cells for the different cell populations.

[0029] Fig. 21 illustrates a heatmap for tissue samples from a database.

[0030] Fig. 22 illustrates quality assessment based on the comparison of cell population feature distributions from different groups.

DESCRIPTION OF EMBODIMENTS

[0031] In the following description, for purposes of explanation and not limitation, details and descriptions are set forth in order to provide a thorough understanding of the present invention. However, it will be apparent to those skilled in the art that the present invention may be practiced in other embodiments that depart from these details and descriptions without departing from the spirit and scope of the invention.

[0032] Certain methods and systems are described for examining tissue specimens based on histology slides using simultaneous multi-parametric analysis of the characteristics of the cells and cell populations from partial or entire tissue sections.

[0033] In an illustrative embodiment, the method may generally comprise five consecutive processing steps, including: (1) tissue preparation; (2) slide preparation; (3) digitization; (4) cell analysis; and (5) cell population analysis. Fig. 1 shows a flowchart of the entire cell-based tissue analysis process from a tissue sample to the cell population analysis in accordance with a general embodiment.

Tissue Preparation

[0034] The tissue preparation process comprises: collection/acquisition of a tissue specimen or sample (e.g. biopsy, excision, surgical specimen), fixation of the tissue sample (e.g. using a fixative such as formalin), transporting the sample to a histology laboratory, creating a tissue block in which the tissue is embedded in a specified media. For purposes herein, a "tissue specimen" or "sample under investigation" may be referred to as a "tissue sample". A similar process is followed in the collection and preparation of frozen tissue samples, with the exception of freezing media is utilized instead of fixation media resulting in

a frozen tissue sample which is processed to a tissue slide using standard and accepted histological procedures.

[0035] This tissue preparation process can have a considerable effect on how the cell features of interest will be expressed in the tissue sections. Careful control needs to be applied to standardize this process.

Slide Preparation

[0036] The slide preparation process comprises standard and accepted histological procedures: the cutting of the tissue block into tissue sections that are placed on glass slides (aka histology slides) and subsequently the staining of the slides (e.g. Hematoxylin and Eosin - H&E, Immunohistochemistry – IHC, etc.) to make the specific cell features of interest detectable by the cell analysis.

[0037] Multiple tissue sections can be cut from a single tissue block whereby each tissue section is typically stained for a specific purpose (e.g. H&E for tissue and cell morphology, IHC-HER2 for the quantification of the protein expression of the Human Epidermal growth Receptor 2, IHC-ER for the quantification of the protein expression of the estrogen receptor, IHC-PR for the quantification of the protein expression of progesterone receptor).

[0038] The cutting of the tissue block (e.g. depth) and the staining process itself can have a considerable effect on how the cell features of interest are expressed in the tissue sections. Careful control needs to be applied to standardize this process.

[0039] Standardized tissue-based or equivalent (e.g. cell lines) controls with different established expression levels (e.g. 0%, 25%, 50% and 75%) of a cell feature (e.g. cell expressions, cell neighborhood characteristics, cell morphology) can be used as part of each staining batch (either on a separate slide or on all slides) to allow the cell analysis programs to automatic calibrated the cell feature measurements. The key is that the controls exhibit similar cell features as contained in the tissue samples.

[0040] One possible choice for a tissue-based control is to use samples from the target tissue. The key is to use samples with different cell feature expression levels and to establish the proper target values for those samples. Establishing the target values for new controls can be done using existing controls. Using this method, special care needs to be taken not to introduce a bias that accumulates from one control to the next.

[0041] In the past, the staining of histology slides had been designed for human interpretation using optical microscopes. Additional information may be obtained using

specialized analytical tools, such as electron microscopy and confocal microscopy, but these are generally not utilized in a high throughput or general laboratory/histology environment. Using sophisticated image analysis programs to analyze the images of histology slides imposes different requirements on the staining. Special care needs to be taken in the choice of stains and their staining procedures. The choice of stains will be driven by how well color separation programs (e.g. color deconvolution) can distinguish them from each other and how well they can be resolved when the stains are colocalized. Determining the right staining procedure will be driven by providing a consistent and detectable staining across the target area (e.g. nucleus). Since different procedures can result in different color hues and intensities for the same descriptive staining process (such as Hematoxylin counterstaining), a standardized staining procedure needs to be followed to avoid variability in interpretation of the staining. In some cases it may be appropriate to provide two histology slides, one stained for human interpretation using standard optical microscopes and one stained for the use of image analysis programs. In either case, the purpose of the standardization is to ensure for reliable subject interpretation by a human or consistent objective interpretation by an image analysis program.

Digitization

[0042] Histology slides can be digitized using commercially available digital microscopes and/or slide scanners (e.g. Aperio, Cri, Hamamatsu, Leica, Omnyx, Philips, Ventana and 3DHistech). Different imaging acquisition techniques (e.g. brightfield, fluorescence, multi-spectral, polarized) can be used to create a digital image of a histology slide. In some cases, different image acquisition techniques can be applied to the same histology slide resulting into multiple images for a single slide. The digitization of a slide can have a considerable effect on how the cell features of interest are imaged. Careful control needs to be applied to standardize this process.

[0043] Assuming that the specific image acquisition characteristics of the instruments are well controlled by each of the manufacturers, still instruments from different manufacturers exhibit sometimes considerable variations in their image acquisition characteristics.

[0044] Color is one of the most important image acquisition characteristics for the cell analysis of stained tissue sections. Standardized color calibration slides with established color target values that are representative for the color spectrum of the used stains or dyes can be used as part of a periodically repeated (e.g daily) instrument calibration procedure. The

difference of the measured color values and the established target values can be used to define a color calibration profile (e.g. ICC profile) that can be applied to the images prior to the cell analysis to normalize all images to a standard color space (e.g. sRGB) for the image analysis.

[0045] One possible choice for a color calibration target is to use existing standard color calibration targets (e.g. IT8). The problem is that those are not representative for the color spectrum of the used stains or dyes. A better way is to measure the color characteristics of the stains or dyes and to create specific color calibration targets.

[0046] Fig. 2 illustrates the color calibration profile generation from a stain specific color calibration target. A brightfield RGB image acquisition and the use of the sRGB color space as the standard color space for image analysis is assumed. The color calibration target shown includes three stains with single stain color patches of different staining intensities and gray scale patches of different intensities for all of which the target colors have been established in sRGB color space. A color calibration profiler program analyses the different patches in the image of the color calibration target and determines, based on the differences between the measured RGB color values and the established target sRGB color values a color calibration profile in form of a look-up table.

[0047] One possible application would be to include Hematoxylin, Eosin and DAB as the stains for the stain-specific color calibration target.

[0048] Important image acquisition characteristics that cannot be calibrated, including the spatial resolution, need to be measured and monitored using special calibration slides as part of a periodically repeated (e.g daily) instrument calibration verification procedure. It is important to make sure that the instrument is operating properly and that the required assumptions for the subsequent cell analysis hold true.

Cell Analysis

[0049] The cell analysis program includes the detection of the cells and the calculation of the cell features in the images from the different tissue sections. The cell analysis is typically application-specific to species (e.g. human), the tissue type (e.g. round cells in breast tissue vs. elongated cells in gastrointestinal tissue), the cell compartments being stained (nucleus, membrane and cytoplasm), the staining (e.g. Hematoxylin, Eosin, DAB) and the image acquisition (e.g. brightfield, fluorescence, multi-spectral, polarized).

[0050] A key idea of the cell-based tissue analysis is that it can include the structural and contextual tissue morphology of the cells in the analysis. The cell features typically include the cell morphology, features that represent a cell's physical presentation on the slide

(e.g. cell size), expressions of biomarkers (e.g. protein, gene and mRNA), and the staining (e.g. Hematoxylin counterstaining). However, cell features are also measured from the cells in a defined neighborhood around the cell, which includes the measurements of the combined groupings of cells to define tissue characteristics (e.g. cell density).

[0051] Fig. 3 shows the data hierarchy associated with the cell analysis. The cells (black circles) are detected in an image from a tissue section. The cell features (e.g. cell morphology, cell neighborhood characteristics, cell expressions) are calculated based on measurements of the cell itself as well as the cells in a neighborhood around the cell and are depicted as distinct greyscale patterns. The neighborhood of a cell is illustrated by the circle with a radius around a cell.

Visualization

[0052] The cells with their cell features can be visualized in many different ways, including intensity/color-coded cell representations (e.g. bitmap of detected cells and cell compartments, circles, ellipses) and heatmaps overlaid on the images from the different tissue sections, 1D, 2D and 3D data plots and sortable and searchable data tables.

[0053] Fig. 4 illustrates the display of cell representations with intensity-coded cell features overlaid on an image. Cells with two cell features "X" and "Y" are shown. The top image shows a cell representation in form of a circle, which is intensity-coded based on cell feature "X". Multiple cell features can be visualized at the same time. The bottom image shows a cell representation in form of two nested circles. The inner circle is intensity-coded based on cell feature "X". The outer circle is intensity-coded based on cell feature "Y".

[0054] Fig. 5 illustrates the display of cell feature data plots. The left data plot shows a 1D histogram of the values from cell feature "X". The right data plot shows a 2D scatter plot of the values from cell features "X" and "Y".

[0055] Fig. 6 illustrates the display of cell feature data tables. The cells and the cell features are displayed as a list in a table with a row for each cell and the cell features "X" and "Y" as the columns. The cells are sorted by the cell feature "X".

[0056] A cell-based heatmap provides high-resolution data in regions of high cell density. The cell feature values are propagated in a certain distance around the cells. The propagation is stopped at the maximum distance or equal distance to another cell.

[0057] Fig. 7 illustrates the display of cell-based heatmaps overlaid on an image.

[0058] The cell analysis can provide an interactive linkage of the cells between the cells overlaid in the images, the cells in the data plots and the cells in the data tables. Cells can

be selected in one of those data representations and those cells will be identified automatically in the other data representations.

[0059] Fig. 8 illustrates the interactive linkage of the cells between different data representations. Fig. 8A shows that when a cell **10** is selected in an image, then the same cell **10** is identified in the data plots and the data tables. Fig. 8B shows that when cells are selected **20** in a data plot, then the same cells are identified **20** in the image and the data tables.

[0060] The cell analysis program allows exporting of cells and the cell features for use with third party data analysis tools.

Calibration

[0061] The cell analysis program provides an automated calibration of the images acquired from different instruments to adjust for the variations in color. The color calibration profile defined for each instrument is applied to all pixels in an image prior to any analysis being performed.

[0062] Fig. 9 illustrates the color calibration of images taken from different instruments using a color calibration profile. A color calibration profile, in form of a look-up table, is used by a color calibration program to transform the RGB values of all pixels in the images from the instrument's RGB color space to a standard color space for the cell analysis, here the sRGB color space.

[0063] The cell analysis program provides an automatic calibration to adjust for the variations in the staining process by using standardized tissue-based or equivalent controls as part of the staining process. One way to use the controls for the calibration is to run the cell analysis program on the controls and then to adjust the data (e.g. cell feature measurements) or the parameters (e.g. cell feature detection thresholds, cell classification thresholds) based on the difference of the measured values and the established target values for the controls.

[0064] Fig. 10 illustrates the calibration of the cell analysis to the staining process using tissue-based or equivalent controls. The image of a slide contains the tissue section and the control. The different expression levels of a cell feature are shown as different intensities in the tissue section and the control. The tissue section can include the full range of a cell feature's expression levels. The control consists of four different samples with different established cell feature expression levels, here with the following target values of 0%, 25%, 50% and 75%. The cell analysis program is run on the control and the resulting histogram of the different control samples are shown with their associated target values. There are differences between the measured values and target values, which can be measured and used

to define a calibration function. The cell analysis program is run on the tissue section and the resulting histogram of the cell feature is shown. Using the calibration function the cell feature measurements are adjusted and shown in the calibrated cell feature histogram of the tissue section.

Multiplexing

[0065] The standard for multiplexing cell features (e.g. IHC-HER2, IHC-ER and IHC-PR) has been at a histology slide level, combining the scores of the entire slides. The cell analysis program can provide multiplexing at the cell level.

[0066] Fig. 11 shows the data hierarchy associated with a multiplexed cell analysis. The multiplexed cell features from the images of two different tissue sections are aligned and associated with the cells.

[0067] Cell features can be combined from images across different tissue sections using image registration techniques. The image registration provides a transform that calculates for any x, y coordinate in the image from one tissue section a new x, y coordinate in the image from another tissue section. Using the transforms, the image from one tissue section can be warped (i.e. spatially transformed) to align with the image from another tissue section, and the location and form of the cells detected in the image from one tissue section can be mapped to the image from another tissue section. In many cases, it makes sense to warp the images and map the cells with their cell features from all tissue sections (e.g. IHC-HER2. IHC-ER and IHC-PR) to a selected common tissue section (e.g. H&E).

[0068] Cell features can be combined from different images taken from the same tissue section using different imaging acquisition techniques (e.g. fluorescence and brightfield) without the use of sophisticated image registration techniques as those images are already perfectly aligned.

[0069] A visual linkage of the cell features from the images across the different tissue sections and/or different images from the same tissue sections can be provided by displaying the aligned warped images of the tissue sections together, by displaying a fused image from the aligned warped images of the tissue sections that can highlight certain cell features (e.g. brightfield H&E with fluorescence FISH-HER2), and by displaying the mapped cells with their cell features from the tissue sections overlaid on the image from a selected tissue section.

[0070] Multiplexing can be provided by combining the cell features across the different tissue sections. Different tissue sections do not necessarily include the same cells, so a 1:1 mapping of the cells and their cell features cannot be expected. But different tissue

sections can be assumed to include the same tissue type regions with very similar cell features that will allow a reasonable alignment of the cell features. The tolerability for maximum distances from cell-to-cell across tissue sections which can be transformed into a single regional measurement carried across slides is dependent on the nature of the tissue architecture and the biological process represented by the analyte.

[0071] A common data structure is needed for the data alignment to map the cell features properly across different tissue sections. Appropriate area-based data structures include a common grid and cell-based data structures include virtual cells.

[0072] A common grid consists of surface areas (e.g squares, octagons) of a certain size that cover the entire tissue sections. The aligned cells from the different tissue sections are associated to individual surface areas based on their location.

[0073] Virtual cells are created based on the cells from the different tissue sections. The aligned cells from the different tissue sections are associated to individual virtual cells based on their proximity.

[0074] The common data structure inherits the cell features from the associated cells and vice versa. Multiple associations can be resolved by averaging the cell features. Smoothing is important for providing a means to reduce the variations in the cell features measurements. The cell features can be averaged using the cells in a certain cell neighborhood. Propagation is important for providing a means to compensate for the inaccuracies in the image registration and the variability of cell locations in different tissue sections. Using a common grid, the cell features of the individual cells can be propagated in the common grid in a certain distance around the cells. Using virtual cells, the propagation is implemented via proximity criteria.

[0075] The cell features from the cells detected in one tissue section can now be mapped to the cells detected in another tissue section using the common data structure. This way all the cell features from the different tissue sections can be mapped to the cells detected in a selected tissue section.

[0076] Fig. 12 illustrates the mapping of cell features across different tissue sections using a common grid. Fig. 12A illustrates the cells detected in the images from three different tissue sections. The cells are shown as circles with different patterns each representing colors (red, green, blue) illustrating the cell features from different tissue sections. Fig. 12B illustrates the mapping of the different cells and cell features from the different tissue sections to the squares in the common grid. Fig. 12C shows the common grid with all the different cell features. Fig. 12D illustrates how cells from one of the tissue sections or virtual cells **30** can

be mapped to the common grid and the cell features from the common grid get associated with the cells.

[0077] Fig. 13 illustrates the mapping of cell features across different tissue sections using virtual cells. Fig. 13A illustrates the cells detected in the images from three different tissue sections. Fig. 13B illustrates how virtual cells **30** are created from the cells detected in the different tissue sections and how they inherit their cell features.

[0078] The common data structure with the cell features from the different tissue sections can be visualized in many different ways, including intensity/color-coded common structure representations and heatmaps overlaid on the image from a selected tissue section, 1D, 2D and 3D data plots and sortable and searchable data tables. The cell analysis program can allow exporting the common data structure with all the cell features for use with third party data analysis tools.

Cell Population Analysis

[0079] The cell population analysis program provides the examination of tissue samples based on the identification, characterization and interpretation of cell populations. The cell populations are used in order to facilitate the interpretation of a biological process. A cell population is defined as all cells that share a particular function of certain cell features. Tissue samples include one or multiple cell populations. Specific cell population features are calculated from the cells included in a cell population. The meta-data (e.g. species, tissue type, age, sex, conditions, drug dosage, clinical outcome) associated with a tissue sample varies depending on the purpose for the analysis.

[0080] Fig. 14 shows the data hierarchy associated with the cell population analysis. Based on all the detected cells and the cell features in a tissue sample, one or multiple cell populations can be identified and specific cell population features can be calculated. The meta-data of a tissue sample is associated to all its cell populations.

[0081] The cell populations can be defined in many ways, which ultimately are specific to the intended outcome of the analysis. The definition of the cell population, which meets the target cell criteria, may be specific, statistical, representational, combinations thereof, or overall collections of defined subpopulations.

[0082] A specific definition results in a definition of a cell based on the specific characteristics of certain cell features. An example of a specific definition would be cells that are classified ER-, PR-, and HER2+, by defining thresholds for positivity for the biomarkers and including the cells that meet all three criteria.

[0083] A statistical definition is one that identifies a cell as being significant by its association with a specific tissue attribute or group of cells in the tissue. An example of a statistical definition may be the inclusion of a cell into the target population if it is defined as being in the stromal compartment, or in an area involved with heavy inflammation. This definition is mainly based on the characteristics of the cell neighborhood features that provide the structural and contextual tissue morphology of the cells in a defined area. The cells that meet the criteria, could be defined individually on each tissue section, and then combined across the tissue sections. The likelihood of each individual cell meeting the criteria could be defined by a statistical confidence.

[0084] A representational population may be one that is characterized by one or more histogram profiles of a measured cell feature. The histogram profile represents a unique population profile with information about the mean, standard deviation, and the presence or absence of subpopulations of cells. An example of a histogram profile is defining cells which are positive for a specific cell feature outside of one standard deviation from the mean, and thus represent a subpopulation of cells different from the expected profile of average cells across a tissue, regardless of the magnitude of measurement.

[0085] An example of a cell population definition, which relies on a combination of these approaches, could be described as "inflammatory cells present in the highest HER2 expressing tumor nests". Here, a cell would be described specifically as inflammatory by a nuclear size threshold defining a small nuclei characteristic of an inflammatory cell; and described representationally as being a cell with a high HER2 expression, and described statistically as being in a tumor nest if the cells within a certain area were statistically comprised of tumor cells.

[0086] One possible cell population classification scheme that would facilitate a standardization of the data sets in a database is based on biology motivated cell types (e.g. invasive tumor cells, carcinoma in-situ cells, stroma cells, normal cells) in combination with the species (e.g. human) and tissue type (e.g. breast tissue) information from the meta-data of the tissue samples. Special cell populations required for the specific analyses could then be implemented as subpopulations of those cell populations or as part of the cell population features.

Identification

[0087] The key to identifying cell populations is to identify the important cell features and define the appropriate function of those cell features that define the cell population.

Uni-, Bi- and Multivariate Analysis

[0088] It is important to understand the relationship between the different cell features (e.g. membrane staining intensity vs. membrane completeness) and their relevance for the identification and/or quantification of certain cell characteristics that are meaningful for the interpretation of tissue samples (e.g. different cell types or cells with a certain protein expression profiles). The different cell features can be thought of as statistical variables.

[0089] The cell population analysis program can provide uni-, bi- and multivariate analysis (incl. probability distribution statistics) and visualization applied to statistical variables. Linear and non-linear component analysis and dimensionality reduction techniques (e.g. Pearson's product-moment coefficient, linear regression) can be used to understand the relationships between variables and their relevance to the actual problem under investigation. Unsupervised and supervised machine learning techniques can be used to identify relationships between the variables.

[0090] Fig. 15 illustrates a linear regression analysis of two cell features. The detected cells in the image from a tissue section are illustrated as a grid of circles. The cells are shown with two cell features "X" and "Y". The circles are intensity-code based on the value of the cell feature. A 2D scatter plot and linear regression of the cell feature "X" and "Y" are shown.

[0091] Alternatively to a cell-based analysis, the common grid (even when using only a single tissue section) can be used to create an area-based analysis and map the data back and forth between an area-based and a cell-based representation.

[0092] The analysis can include data from multiple tissue samples (or different blocks from the same tissue sample), in which case the tissue samples' meta-data can be thought of as additional statistical variables. The cells from different tissue samples can be visualized distinctively (e.g. color-coded) in the analysis. The cells and cell features from multiple tissue samples as well as the images of the tissue sections can be conveniently stored and accessed from a database.

[0093] Fig. 16 illustrates an analysis based on cells from different tissue samples. The cells from two tissue samples "A" and "B" are shown with the same cell features "X" and "Y". A 2D scatter plot of the cell features "X" and "Y" shows the cells from the two tissue sections in different colors (black – tissue sample A, and white – tissue sample B).

[0094] The intuitive or empirical knowledge about the target cells can be provided through cell annotations that can be thought of as additional statistical variables. Cell annotations are created by identifying examples of cells and associating them with certain labels or values.

[0095] Fig. 17 illustrates the creation of a cell annotation. Multiple regions-of-examples can be defined for a cell annotation in an image from a tissue section and provides for each region an associated value. The regions-of-examples are intensity-coded based on the associated value. A new cell feature is created for the cell annotation. Now the cell annotation can be used in the same way as the measured cell features, as shown with the 2D scatter plot and linear regression of the cell feature "X" and the cell annotation.

[0096] The cell population analysis program can provide the capability to define new cell features that are calculated from existing cell features. Those new cell features can be used to identify and quantify certain cell characteristics that are meaningful for the interpretation of tissue samples (e.g. certain cell protein expression profiles). New cell classification, scoring and interpretation schemes can be created based on or including those new cell features.

[0097] How a new cell feature is calculated is defined using different calculation models, including linear and non-linear functions, template similarities (e.g. normalized sum of absolute differences) and any algorithm (e.g. program code implementation) that calculates a new value from a set of values.

[0098] Supervised machine learning techniques can assist in the calculation of a new cell feature, which is determined automatically based on user-provided cell annotations.

[0099] Fig. 18 illustrates the creation of a new cell feature. A calculation model is provided that calculates a new cell feature based on the cell features "X", "Y" and "Z". A new cell feature is created. Now the new calculated cell feature can be used in the same way as the measured cell features, as shown with the visualization of the new cell feature in an image from a tissue section.

[0100] A possible application would be to implement the 5-year survival statistics based on the HER2, ER and PR expression levels published in [1] as a new cell feature. The cell feature can be calculated based on the data points from the publication (ER+/PR+/HER2 - 96.4%, ER+/PR-/HER2 - 91.9%, ER+/PR+/HER2+ 91.3%, ER+/PR-/HER2+ 88.0%, ER-/PR+/HER2- 82.7%, ER-/PR+/HER2+ 78.8%, ER-/PR-/HER2- 76.2%, ER-/PR-/HER2+ 75.9%) and a linear approximation between the data points. Now using the cell data of the 5-year survival statistics over all invasive tumor cells can be investigated for the basis of a new scoring scheme.

Manual Gating

[0101] The cell population analysis program can provide manual gating, whereby cell populations are identified by defining sequential gating functions of selected cell features. Manual gating typically supports a specific definition of cell populations.

[0102] Regions-of-examples for the different cell populations in the images can be defined from the tissue sections to identify representative cells with their cell features. Those identified cells and their association to certain cell populations can be visualized (e.g. separate plots, color-coded) in the gate definition process. All cells or only those cells from the regions-of-examples can be selected for use in the gate definition process.

[0103] One-, two- or three- dimensional graphical representations of selected cell features are used to identify the regions (e.g. min./max., multiple polygons) of the cells of interest. A sequence (or hierarchy) of intersections and unions of those regions definitions of one to three cell features can be created to define a cell population. The sequential gating function can be incrementally created, at each step only looking at the resulting cells from the previous step.

[0104] A very simple and effective way to create a gating function is to use a marker stain for a certain cell type and to gate on the cell feature that measures the marker stain. Using multiplexing, a marker stain can be used for a certain cell type on one tissue section and thereby identify that cell type on all other tissue sections.

[0105] The cells corresponding to the cell populations can be identified at the same time in the images from the tissue sections.

[0106] Fig. 19 illustrates the gate definition process using representative cells for the different cell populations. The detected cells in the image from a tissue section are illustrated as a grid of black or patterned circles. The graphical representation of a single cell feature is shown as a 1D histogram. Fig. 19A illustrates the defining of two regions-of-examples **40**; **50** for two cell populations of interest and the creation of a new cell annotation for the cell population examples. Fig. 19B illustrates a gate definition step, which defines a minimum and a maximum threshold for the cell feature "X". In addition to a plot that includes all cells, separate plots for each cell population containing only the cells from the corresponding regions-of-examples **40**; **50**, respectively, are shown. The region definition **41** (min. and max.) from the "to be defined" cell population **40** is shown in the other plots as well. Fig. 19C illustrates the creation of a new cell feature for the cell populations and shows the pattern-coded cells **40** from the "to be defined" cell population in the image.

[0107] The gate definition process can include multiple tissue samples. This allows taking the variations between tissue samples into account. The cells from different tissue

samples can be visualized (e.g. separate plots, color-coded) in the gate definition process. The tissue samples' cells and cell features as well as the images of the tissue sections can be conveniently stored and accessed from a database.

[0108] The cell population analysis program allows using a defined gating function for the automated gating of cells in different tissue samples.

Pattern Recognition

[0109] The cell population analysis program can provide cell-based pattern recognition, based on supervised machine learning techniques, for the automatic classification of the cell into the different cell populations. Pattern recognition is typically used for a specific and statistical definition of cell populations. The machine learning technique used determines how important cell features are identified and what kind of calculation model will be used for the cell classification. The cell classification can be expressed using class labels (i.e. a cell can only be part of one cell population) or class probabilities (i.e. a cell can be part of multiple cell populations).

[0110] Regions-of-examples for the different cell populations in the images from the tissue sections can be defined to identify representative cells with their cell features.

[0111] The pattern recognition tool identifies the important cell features and determines the cell classifier based on the provided regions-of-examples.

[0112] Alternatively or in combination to cell-based pattern recognition, the common grid (even when using only a single tissue section) can be used to perform area-based pattern recognition and map the data back and forth between an area-based and a cell-based representation.

[0113] The cells corresponding to the different cell populations can be identified in the images of the tissue section.

[0114] Fig. 20 illustrates the pattern recognition process using representative cells for the different cell populations. Fig. 20A illustrates the defining of two regions-of-examples **60**; **70**, respectively, for two cell classifications of interest and the creation of a new cell annotation **A**, **B**, **NA**, respectively for the cell classification examples. Fig. 20B illustrates the creation of a new cell feature for the cell classification and shows the pattern-coded cells **A**; **B** from the two cell classes in the image.

[0115] The pattern recognition can include multiple tissue samples. This allows taking the variations between tissue samples into account. The tissue samples' cells and cell features

as well as the images from the tissue section can be conveniently stored and accessed from a database.

[0116] The cell population analysis program allows using a defined cell classifier for the automated classification of cells in different tissue samples.

Population Mixture Model Analysis

[0117] The cell population analysis program provides population mixture model analysis, including unsupervised machine learning techniques (e.g. Gaussian mixture model), for the automatic classification of the cell into the different cell populations. Population mixture model analysis is typically used for a representational definition of cell populations. A mixture model is a probabilistic model for representing the presence of cell subpopulations within an overall cell population. Mixture models are used to make statistical inferences about the properties of cell subpopulations given only observations from the overall cell population. Different kinds of distribution models can be used including, a Gaussian model.

Population Comparison

[0118] The cell population analysis program provides population comparisons, based on the comparison of cell population cell features' distribution functions and cumulative distribution functions (CDF). Different cell populations within the same tissue section or across different tissue sections can be compared to define a representational cell population.

Cell Population Features

[0119] Tissue samples are characterized by different cell populations and their cell population features. Once the cell populations are identified, cell population features (e.g. percentage of cells with respect to the total number of cells, statistics of cell features and cell feature histograms) are calculated from the cells and the cell features for each cell population.

Interpretation

[0120] The interpretation of cell populations deals with the comparison of tissue samples and the predication for tissue samples.

Databases

[0121] Looking at the data from tissue samples across a study or an institution or in the context of large populations provides new insights and a new perspective on the data.

[0122] The data set of a tissue sample can include the meta-data (e.g. species, tissue type, sex, age, conditions, drug dosage and clinical outcome), the cell populations with their

cell population features (e.g. percentage of cells with respect to the total number of cells, statistics of cell features and cell feature histograms), but can also include the cells with their cell features and the images of the tissue sections.

[0123] The key is to provide an infrastructure that allows creating, managing and accessing multiple tissue sample databases. An institution could have its own central database and access other public databases. Databases can be on a local computer, on a central server on a network (incl. virtual private networks) and/or be implemented as a cloud application.

[0124] A standardization of the data sets for the tissue samples is important for the utility of the databases.

Data Mining and Visualization

[0125] Data mining techniques, including unsupervised and supervised machine leaning techniques, and visualization techniques, including heatmaps, 1D, 2D and 3D data plots and sortable and searchable data tables, can be used with the databases to automatically or semi-automatically analyze large quantity of data. The goal of data mining and visualization is to extract previously unknown interesting patterns such as groups of data sets (cluster analysis), unusual data sets (anomaly detection) and dependencies (association rule mining), and to be able to provide predications for data sets in form of classification or regression analysis.

[0126] Fig. 21 illustrates a heatmap for tissue samples from a database. The rows in the heatmap represent the different characteristics and the columns the different records. Specific meta-data, cell populations and cell population features are selected in the "Select Items" category. The "Options" category allows limiting the range of the characteristics and to define the way the data should be displayed in the heatmap (e.g. values, color- or intensity-coded). The heatmap is limited to a certain species and tissue type. The "Sort" category allows sorting the data by characteristics. The black boxes illustrate the capability to access the complete information for a tissue sample. In addition to the meta-data and the cell population data in the heatmap, access to the corresponding cell data and the images is provided.

[0127] Additional knowledge or assumptions about the tissue samples can be included in the analysis by simply adding the information to the meta-data.

Decision Support System

[0128] The cell population analysis program can include a decision support system based on the tissue analysis data and related meta-data of a tissue sample under investigation.

[0129] The decision support system can provide a similarity analysis that allows quantifying and visualizing (e.g. heatmaps) the differences in the tissue analysis data (e.g. human – breast tissue - invasive tumor cell population - HER2, ER and PR expressions) and meta-data from a tissue sample under investigation and data sets from a database with known conditions or outcome (e.g. survival after different treatment options). The decision support system can provide access to the slide images, the cells with their cell features, the cell populations with their cell population features and related meta-data, as well as additional case information, including data from other modalities (e.g. radiology data), for the data sets in the database. This allows answering questions such as, how close is the data to that of known conditions or outcome.

[0130] A simple way to implement a decision support system based on a similarity analysis is to expand the heatmap as shown in Fig. 21 with a similarity criterion (e.g. weighted sum of absolute differences) for selected meta and cell population data. The heatmap can now show a tissue sample under investigation in the context of similar tissue samples in the database.

[0131] Rather than providing a decision support system based on the actual data sets in a database, a special data set can be created for a decision support system that is based on a synthesis of clinical findings and example cases. The tissue sample under investigation can now be shown in the context of all the different significant characterizations of a case with possible decisions and to be expected outcome.

[0132] A possible application would be to use the 5-year survival statistics based on the HER2, ER and PR expression levels published in [1] for the synthesis of clinical findings and to select representative and interesting cases as examples. Eight cases could be created based on the data points from the publication (ER+/PR+/HER2 - 96.4%, ER+/PR-/HER2 - 91.9%, ER+/PR+/HER2+ 91.3%, ER+/PR-/HER2+ 88.0%, ER-/PR+/HER2- 82.7%, ER-/PR+/HER2+ 78.8%, ER-/PR-/HER2- 76.2%, ER-/PR-/HER2+ 75.9%).

Automatic Meta-Data Prediction

[0133] The cell population analysis program can provide the capability to predict meta-data based on a tissue sample's cell populations and the cell population features. Supervised machine learning techniques can be used to learn the classifier from tissue sample examples with the existing meta-data.

[0134] A possible application is the prediction of the species and tissue type as the cell and cell population analysis is species and tissue specific. The information about the species

(e.g. human) and tissue type (e.g. breast tissue) is typically provided with the tissue sample. The cell population analysis program can also predict this information automatically by comparing the cell populations of the tissue sample with known species-tissue type cell populations.

In Vitro Diagnostic Multivariate Index Assays (IVDMIA)

[0135] IVDMIAs can be defined and implemented based on the tissue samples' cell population features (e.g. HER2 expression, ER expression, PR expression) from one (e.g. invasive tumor cells) or multiple cell populations (e.g. stroma cells) and related meta-data (e.g. age, sex, conditions).

[0136] Uni-, bi- and multivariate analysis and data mining and visualization can be used to identify common patterns and to determine the important features and their relationship to clinical outcome.

[0137] New scoring and interpretation schemes can be defined using different calculation models, including linear and non-linear functions, template similarities (e.g. normalized sum of absolute differences) and any algorithm (e.g. program code implementation) that calculates a new value from a set of values.

[0138] An approach for defining an IVDMIA involves the use of profile histograms to define the characteristics of cell population features used to determine the IVDMIA. An example of a profile histogram is a dataset that displays the mean value of Y and its root mean square (RMS) for each bin in X. Each Y parameter would be chosen from the cell population features. Each X parameter is arbitrary, and may be a related, a non-obviously related or unrelated dataset of criteria related to the cell-based tissue analysis approach. In this case, the included X and Y values are plotted in the profile histogram to create a unique pattern which represents a specific IVDMA output. This pattern captured in the IVDMIA output is correlated to the meta-data in a fashion that specific IVDMIA output patterns (profile histogram) are predictive of a clinical output related to the meta-data. For example, this profile histogram, or a range of similar profiles, could be used to define a binary (yes/no, positive/negative, etc) output of the IVDMIA which has prognostic or diagnostic value. Thus, the IVDMIA captures the multidimensional outputs of the cell-based tissue analysis, and captures it in a summary format that is amenable to algorithmic determination of an output with direct application for the IVDMIA.

[0139] A possible application would be to implement the calculation of the 5-year survival statistics based on the HER2, ER and PR expression levels as published in [1]. A

simple calculation model could be to use the data points from the publication (ER+/PR+/HER2 - 96.4%, ER+/PR-/HER2 - 91.9%, ER+/PR+/HER2+ 91.3%, ER+/PR-/HER2+ 88.0%, ER-/PR+/HER2- 82.7%, ER-/PR+/HER2+ 78.8%, ER-/PR-/HER2- 76.2%, ER-/PR-/HER2+ 75.9%) and to provide a linear approximation between the data points.

[0140] New scoring and interpretation schemes can be determined automatically, assisted by supervised machine learning techniques and based on selected tissue samples with corresponding meta-data (e.g. clinical outcome).

Calibration

[0141] The cell population analysis program can provide an automatic calibration of the cell population features (e.g. HER2, ER, PR expressions) to adjust for the variations in the tissue and slide preparation process by comparing the detected cell populations and their cell population features to a norm. A norm can be established from tissue sections that have been prepared using controlled procedures varying the different tissue and slide preparation parameters (e.g. varying fixation time) and/or analyzing the data sets from large populations. This kind of calibration can be thought of as using internal tissue controls all the way from a biopsy to the final analysis.

[0142] An approach for defining a calibration output involves the use of profile histograms to define the characteristics of cell population features to determine the calibration output values. Each Y parameter would be chosen from the cell population features. Each X parameter is a criteria related to the cell-based tissue analysis approach. In this case, the included X and Y values are plotted in the profile histogram to create a unique pattern which represents a specific calibration output. If the profile histogram falls within a defined range of the predetermined normal standard, then the assay used to create the sample is considered properly calibrated. This approach can be used to assess all points in the tissue and slide preparation process.

Quality Assessment

[0143] Artifacts from the tissue and slide preparation, digitization, or other factors may confound the tissue analysis and lead to errors. Therefore, quality assessment is a crucial step of the tissue analysis. The quality assessment includes the measurements of inter tissue sample variability and the detection of variations that are not likely biologically motivated. Biologically motivated variations can be defined by analyzing the data sets from the target population.

[0144] An approach for defining a quality output involves the use of profile histograms to define the characteristics of cell population features used to determine the quality output values. Each Y parameter would be chosen from the cell population features. Each X parameter is a criteria related to the cell-based tissue analysis approach. In this case, the included X and Y values are plotted in the profile histogram to create a unique pattern which represents a specific output. If the histogram profile falls within an empirical or intuitively defined range of normalcy, then the sample meets the quality standards. An example of an intuitive definition of the defined range normalcy may be subjective to a pathologist interpretation. An example of an empirical definition of the defined range of normalcy may be inclusion of the sample within one standard deviation of the mean of population of all samples analysed using the same methodology.

[0145] Fig. 22 illustrates quality assessment based on the comparison of cell population feature distributions from different groups. The cell population feature distributions for the same cell population feature "X" are shown for two different groups "A" **80** and "B" **90** in the same 1D data plot. It can easily be seen that the distributions are different in terms of their mean (M) and standard deviation (SD). Additional statistical tests can provide clear acceptance criteria.

[0146] A possible application would be to compare the distribution of protein expressions (e.g. HER2, ER and PR) of a study or institution to the distribution of the same protein expressions in the target population. There could be international, national and regional databases that accumulate and provide the data sets from the target populations.

INDUSTRIAL APPLICABILITY

[0147] The invention relates to methods for cell based tissue analysis used in medical imaging and analysis fields.

CLAIMS

We claim:

1. A method for cell-based tissue analysis, comprising:

preparing one or more tissue sections and respective histology slides from one or more tissue samples using stains, staining processes and controls that are optimized for use with image analysis;

creating calibrated digital images of the histology slides using the controls and stain-specific color calibration targets;

detecting cells within the digital images, with the detected cells:

measuring and calibrating one or more cell features and multiplexing cell features across adjacent tissue sections to assemble cell-based data, and

representing the cell-based data with interactive computer-generated visualizations;

identifying cell populations using the cell-based data, with the identified cell populations:

calculating and calibrating one or more cell population features,

interpreting cell populations, and

representing the cell populations with interactive computer-generated visualizations; and

storing information associated with the tissue samples in one or more databases for comparative analysis.

2. The method of claim 1, wherein said cell-based features comprise one or more from the group consisting of: cell morphology, cell expressions, staining characteristics, cell neighborhood characteristics, and cell region characteristics.

3. The method of claim 1, wherein said cell population features comprise one or more of: percentage of cells with respect to the total number of cells, statistics of cell features, and cell feature histograms.

4. The method of claim 1, wherein said multiplexing comprises the use of image registration techniques to map the cells and the cell features between the different tissue sections.

5. The method of claim 1, wherein said identification of cell populations provides uni- bi- and multivariate analysis.

6. The method of claim 1, wherein said cell populations are identified using sequential gating.

7. The method of claim 1, wherein said cell populations are identified using pattern recognition techniques.

8. The method of claim 1, wherein said cell populations are identified using population mixture model analysis.

9. The method of claim 1, wherein said cell populations are identified using population comparisons based on the comparison of cell populations' cell features' distribution and cumulative distribution functions.

10. The method of claim 1, wherein said interpretation of cell populations provides data mining and visualization.

11. The method of claim 1, wherein said interpretation of cell populations provides decision support systems that are based on a similarity analysis of tissue samples under investigation with data sets in the data bases or with a special data set created from clinical findings and example cases.

12. The method of claim 1, wherein said interpretation of cell populations provides automatic predication of meta-data based on a tissue sample's cell populations and their cell population features using machine learning techniques.

13. The method of claim 1, wherein said interpretation of cell populations allows for the definition of an in vitro diagnostic multivariate index assays (IVDMIA) for tissue samples from one or multiple cell populations and their cell population features and related meta-data.

14. The method of claim 13, wherein said definition of an in vitro diagnostic multivariate index assays (IVDMIA) is defined using machine learning techniques.

15. The method of claim 1, wherein said calibration of the cell features and cell population features is done for the variations in the tissue and slide preparation process by comparing the detected cell populations and their cell population features to a norm that can be established using controlled procedures varying the different tissue and slide preparation parameters or analyzing the data sets from large populations.

16. The method of claim 1, wherein said interpretation of cell populations provides a quality assessment that measures the inter tissue sample variability and detects variations that are not likely biologically motivated.

17. The method of claim 1, wherein said computer-generated visualizations comprise one or more of: interactive linkage between data and display, image overlays, heatmaps, data plots, 3D displays of aligned images, fused images from aligned images and mapped cell data and image overlays of mapped cell data onto an image from a different tissue section.

18. The method of claim 1, wherein said information associated with the tissue samples comprises one or more of: meta-data, cell populations with their cell population features, cells with their cell features, images of the tissue sections, additional case information and data or links to data from other modalities.

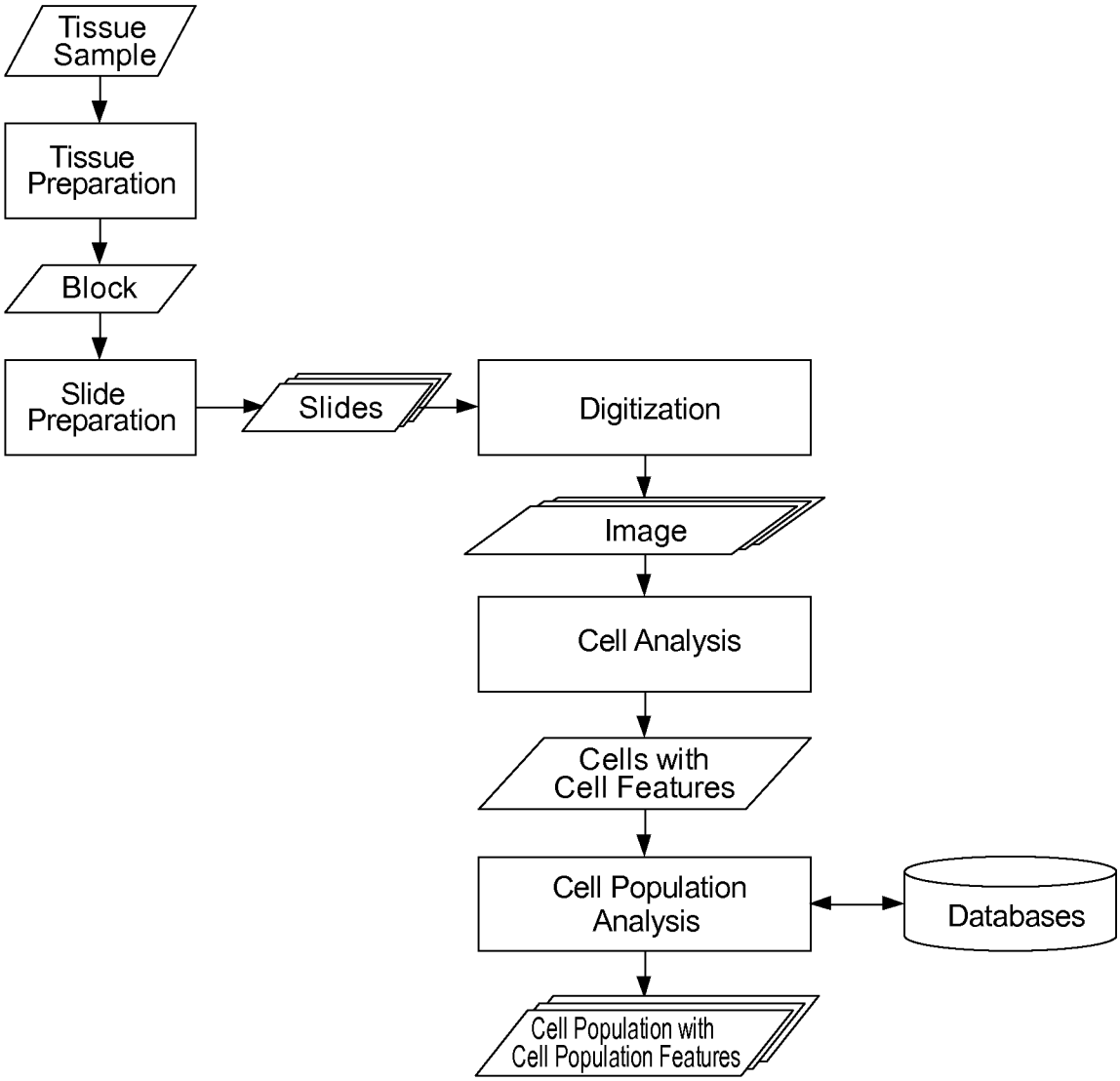
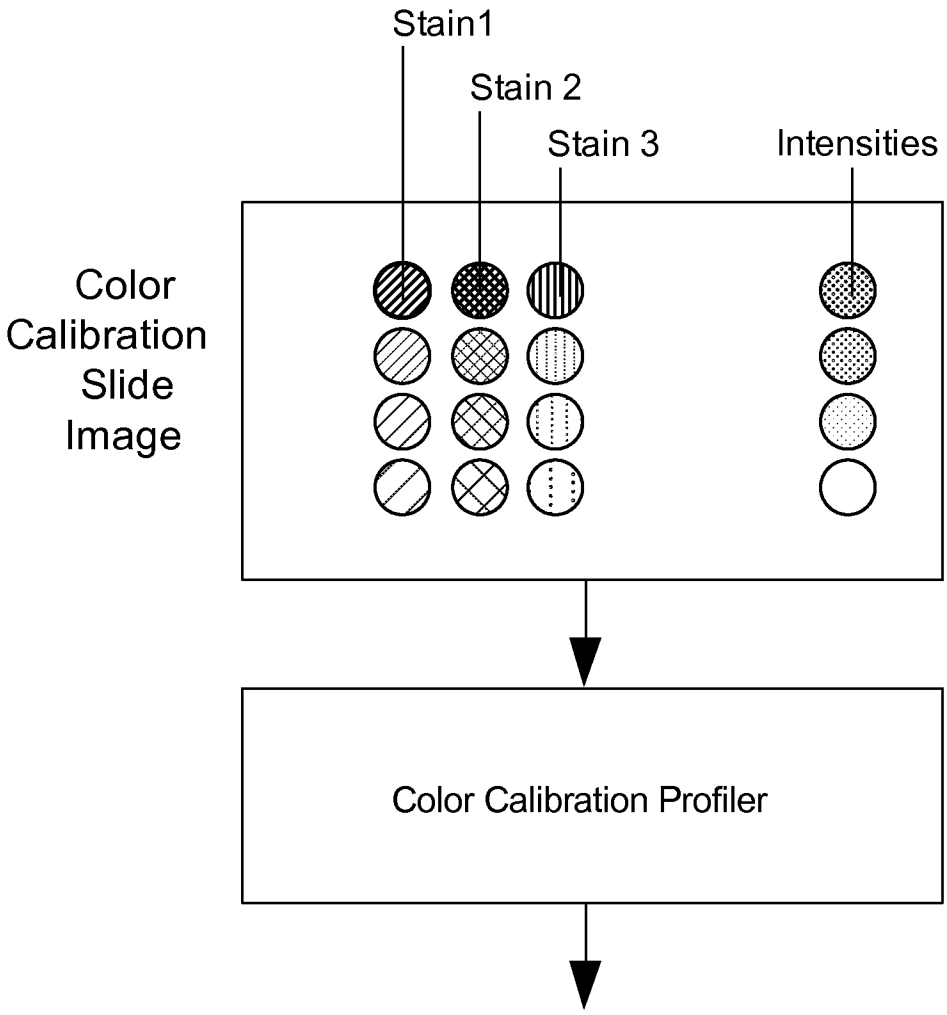

Tissue Sample
Tissue Preparation
Block
Slide Preparation
Slides
Digitization
Image
Cell Analysis
Cells with Cell Features
Cell Population Analysis
Databases
Cell Population with Cell Population Features

Fig. 1



Look-Up Table

Instrument				sRGB		
R	G	B	→	R	G	B
0	0	0		0	0	0
0	0	1		0	1	1
0	0	2		0	1	2
0	0	3		1	1	4
...				...		

Fig. 2

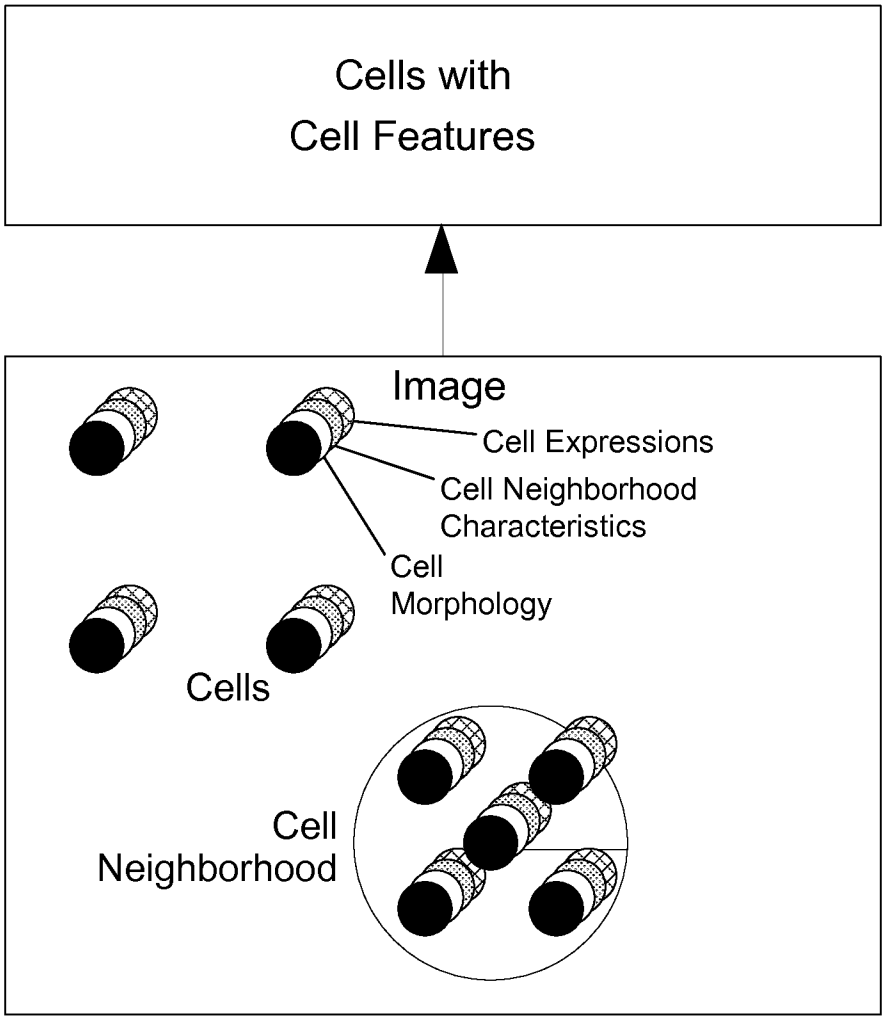

Cells with
Cell Features
Image
Cell Expressions
Cell Neighborhood
Characteristics
Cell
Morphology
Cells
Cell
Neighborhood

Fig. 3

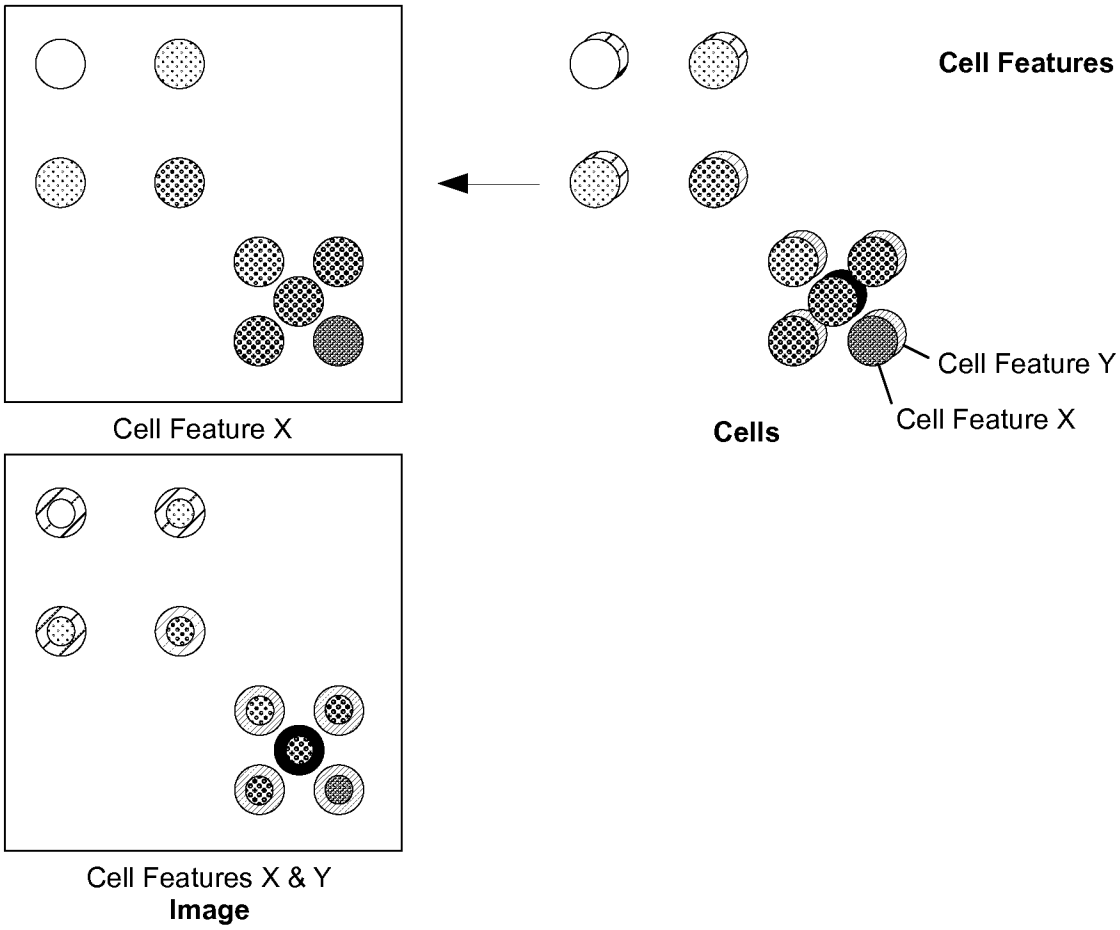

Cell Features
Cell Feature Y
Cell Feature X
Cells
Cell Feature X
Cell Features X & Y
Image

Fig. 4

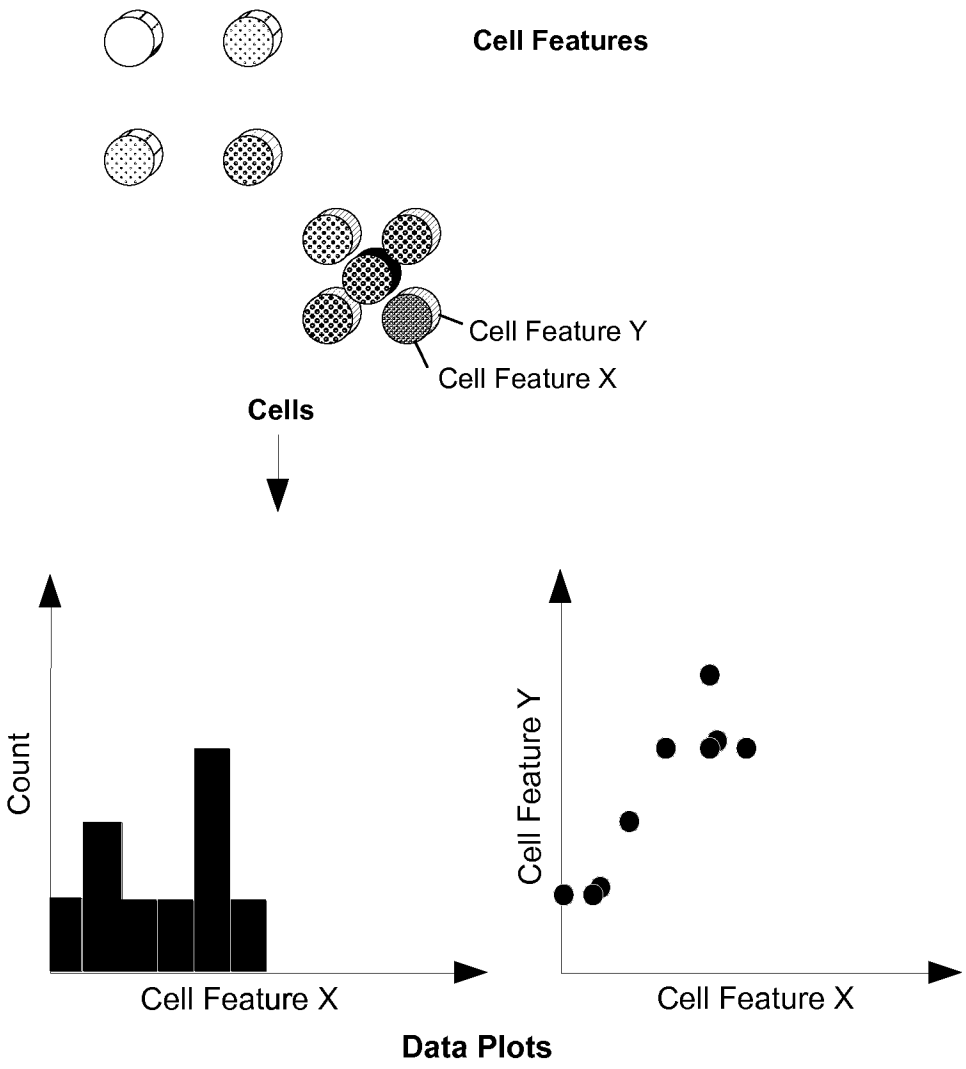
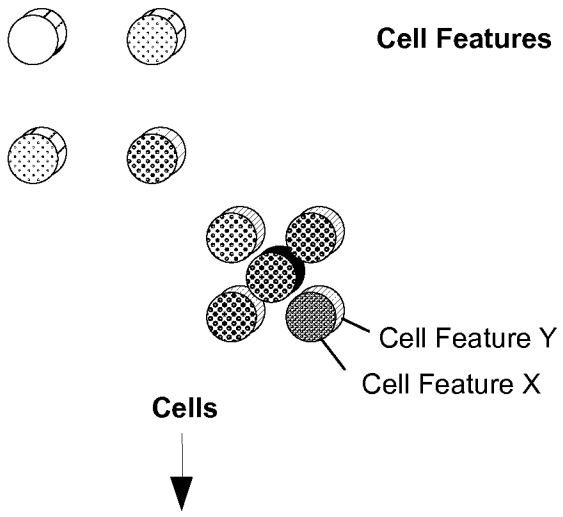

Cell Features
Cell Feature Y
Cell Feature X
Cells
Count
Cell Feature X
Cell Feature Y
Cell Feature X
Data Plots

Fig. 5



Cell Feature X ▼	Cell Feature Y ▽
0%	20%
10%	20%
10%	20%
20%	40%
30%	60%
40%	60%
40%	60%
40%	80%
50%	60%

Data Table

Fig. 6

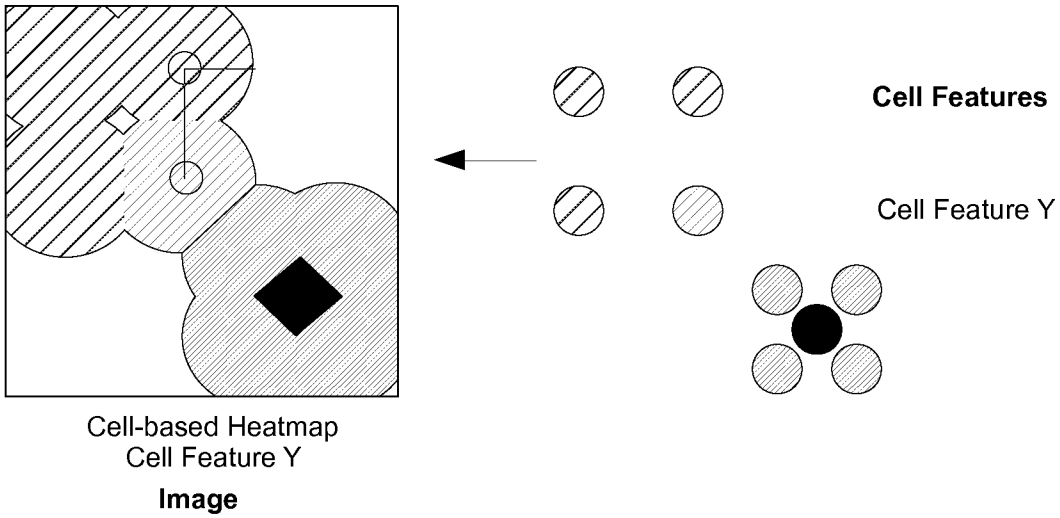

Cell Features
Cell Feature Y
Cell-based Heatmap
Cell Feature Y
Image

Fig. 7

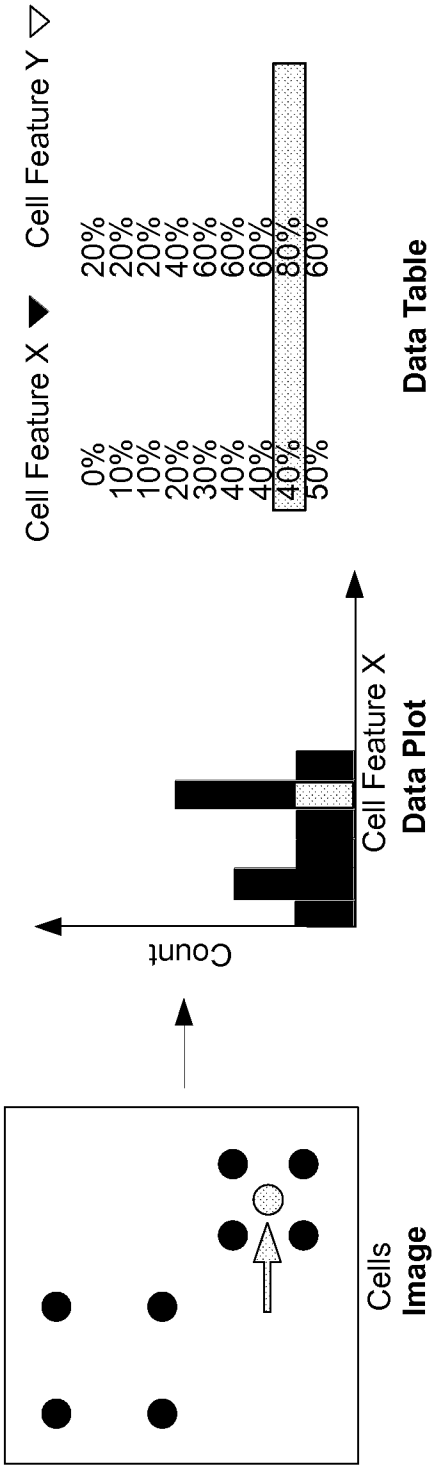

Cell Feature X
Cell Feature Y
0% 20%
10% 20%
10% 20%
20% 40%
30% 60%
40% 60%
40% 60%
40% 80%
50% 60%
Data Table
Count
Cell Feature X
Data Plot
Cells
Image

Fig. 8A

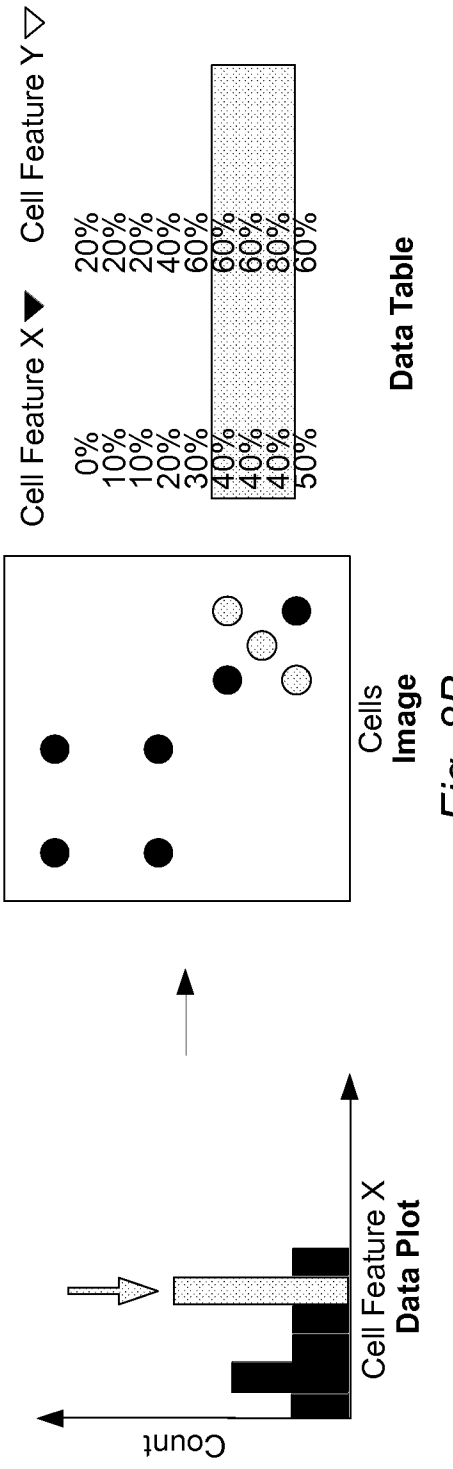

Cell Feature X
Cell Feature Y
0% 20%
10% 20%
10% 20%
20% 40%
30% 60%
40% 60%
40% 60%
40% 80%
50% 60%
Data Table
Count
Cell Feature X
Data Plot
Cells
Image

Fig. 8B

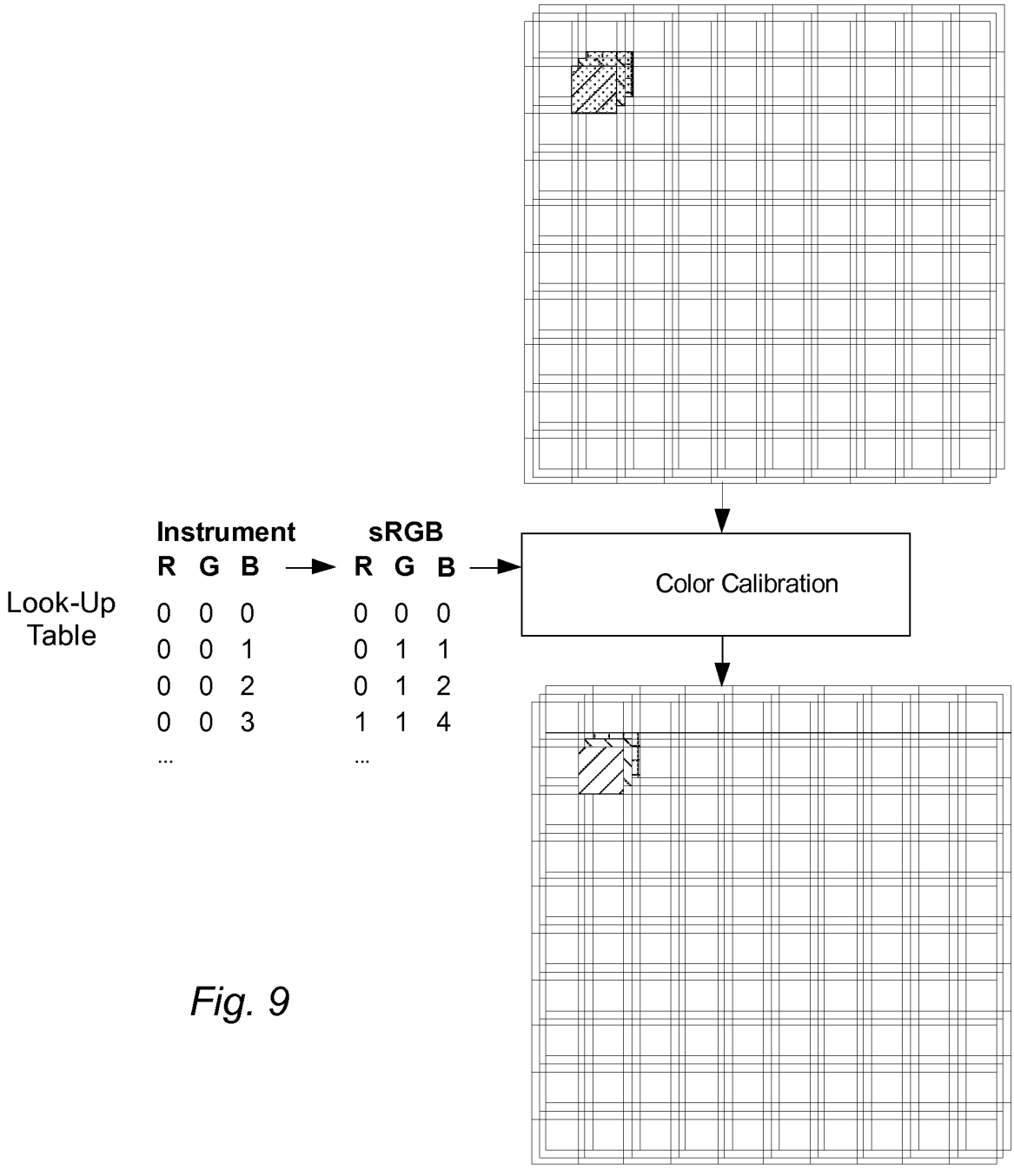

Look-Up Table
Instrument
R G B
sRGB
R G B
0 0 0 0 0 0
0 0 1 0 1 1
0 0 2 0 1 2
0 0 3 1 1 4
... ...
Color Calibration

Fig. 9

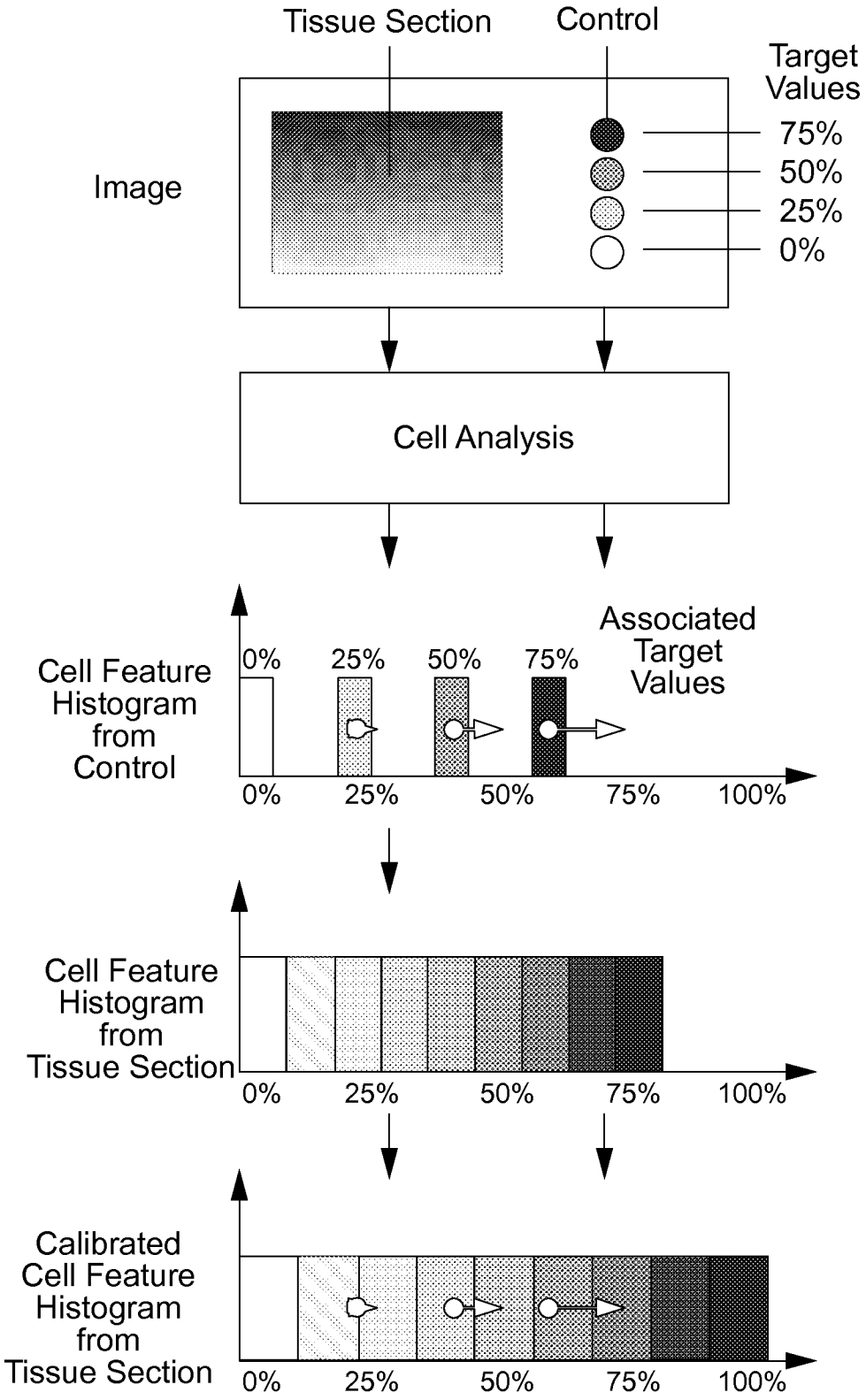

Tissue Section
Control
Target Values
75%
50%
25%
0%
Image
Cell Analysis
Cell Feature Histogram from Control
0%
25%
50%
75%
Associated Target Values
0%
25%
50%
75%
100%
Cell Feature Histogram from Tissue Section
0%
25%
50%
75%
100%
Calibrated Cell Feature Histogram from Tissue Section
0%
25%
50%
75%
100%

Fig. 10

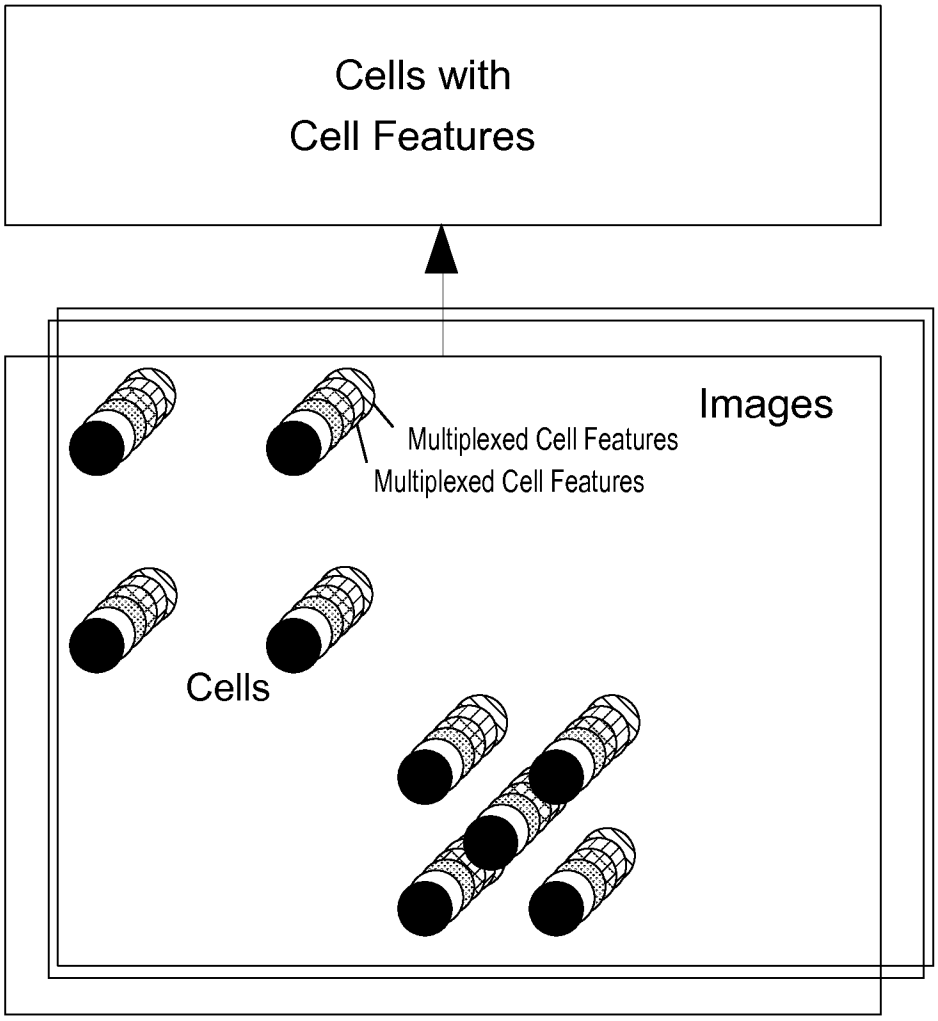
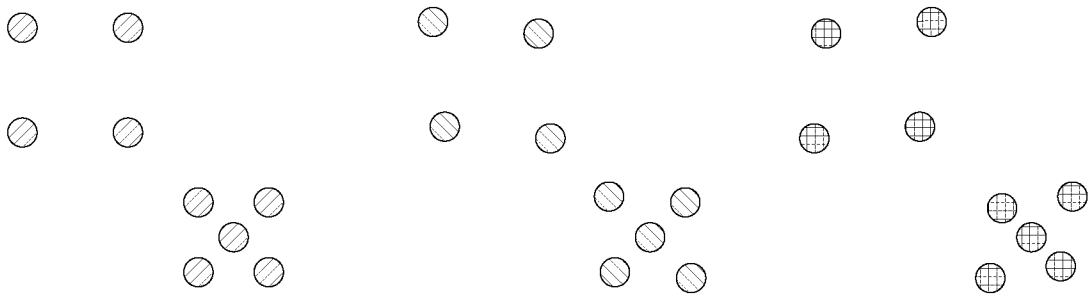

Cells with
Cell Features
Images
Multiplexed Cell Features
Multiplexed Cell Features
Cells

Fig. 11

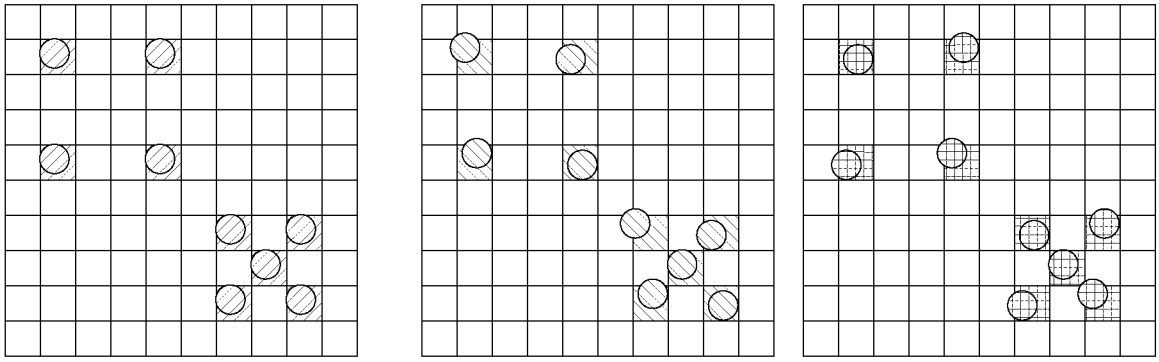


Tissue Section 1
Cells

Tissue Section 2
Cells

Tissue Section 3
Cells

Fig. 12A

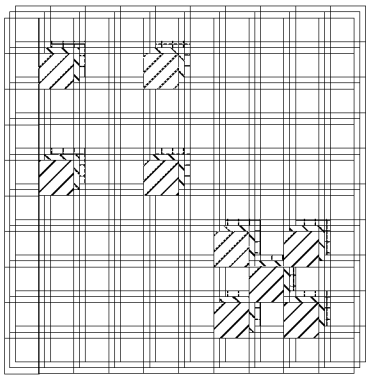


Tissue Section 1

Tissue Section 2
Common Grid

Tissue Section 3

Fig. 12B



Common Grid

Fig. 12C

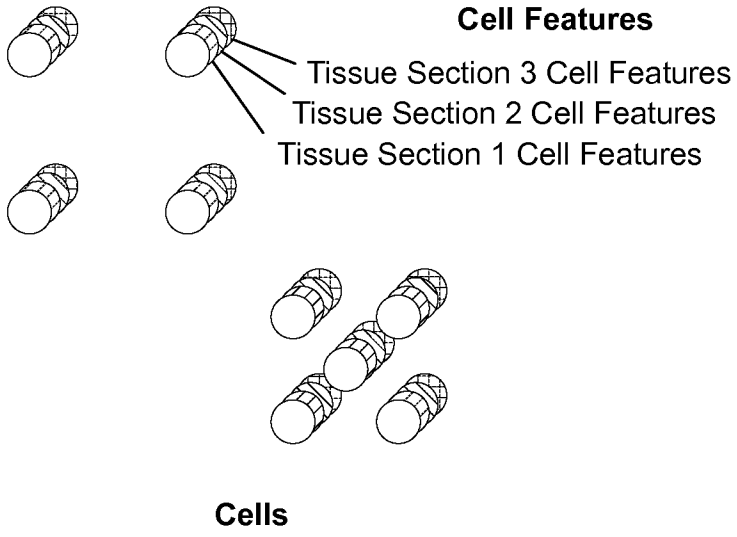

Cell Features
Tissue Section 3 Cell Features
Tissue Section 2 Cell Features
Tissue Section 1 Cell Features
Cells
Common Grid

Fig. 12D

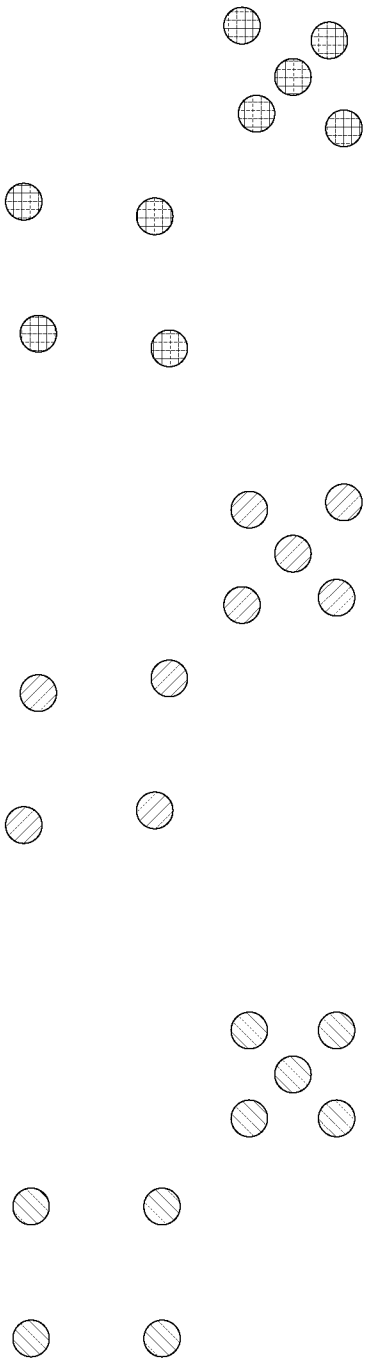


Fig. 13A

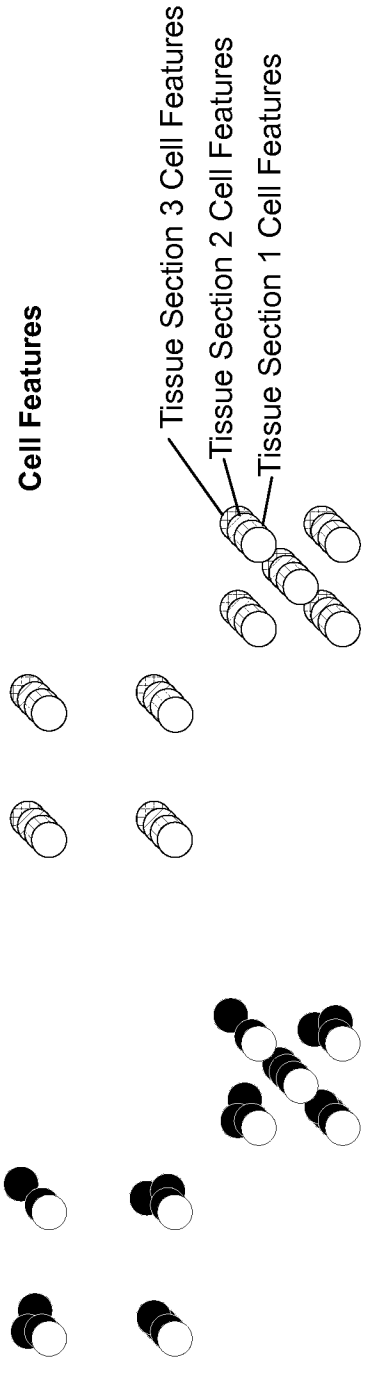


Fig. 13B

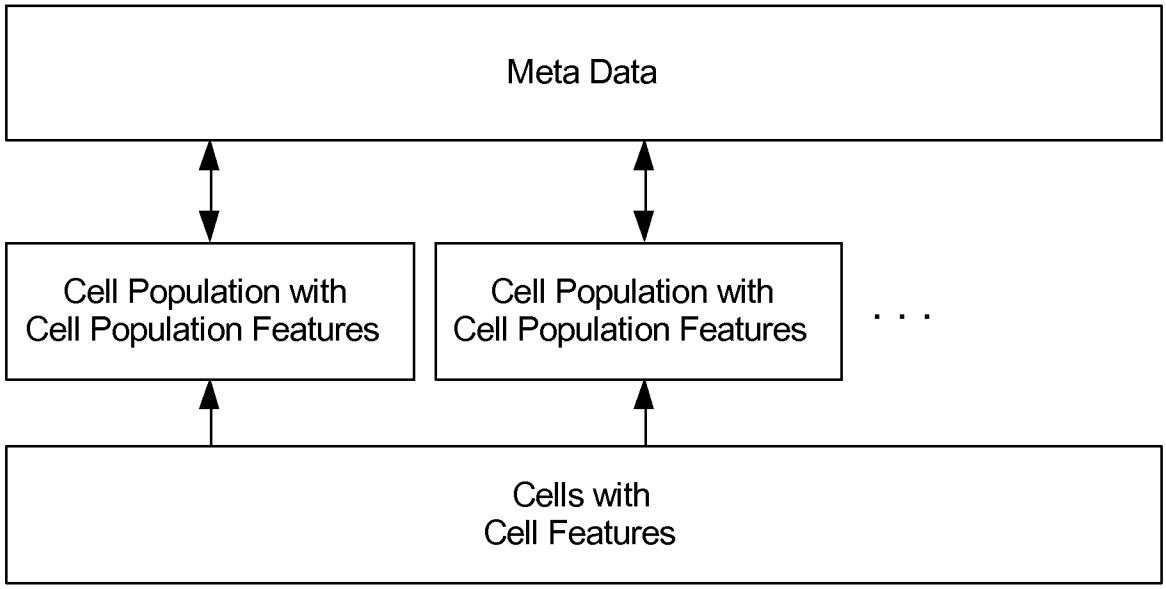

Meta Data
Cell Population with
Cell Population Features
Cell Population with
Cell Population Features
. . .
Cells with
Cell Features

Fig. 14

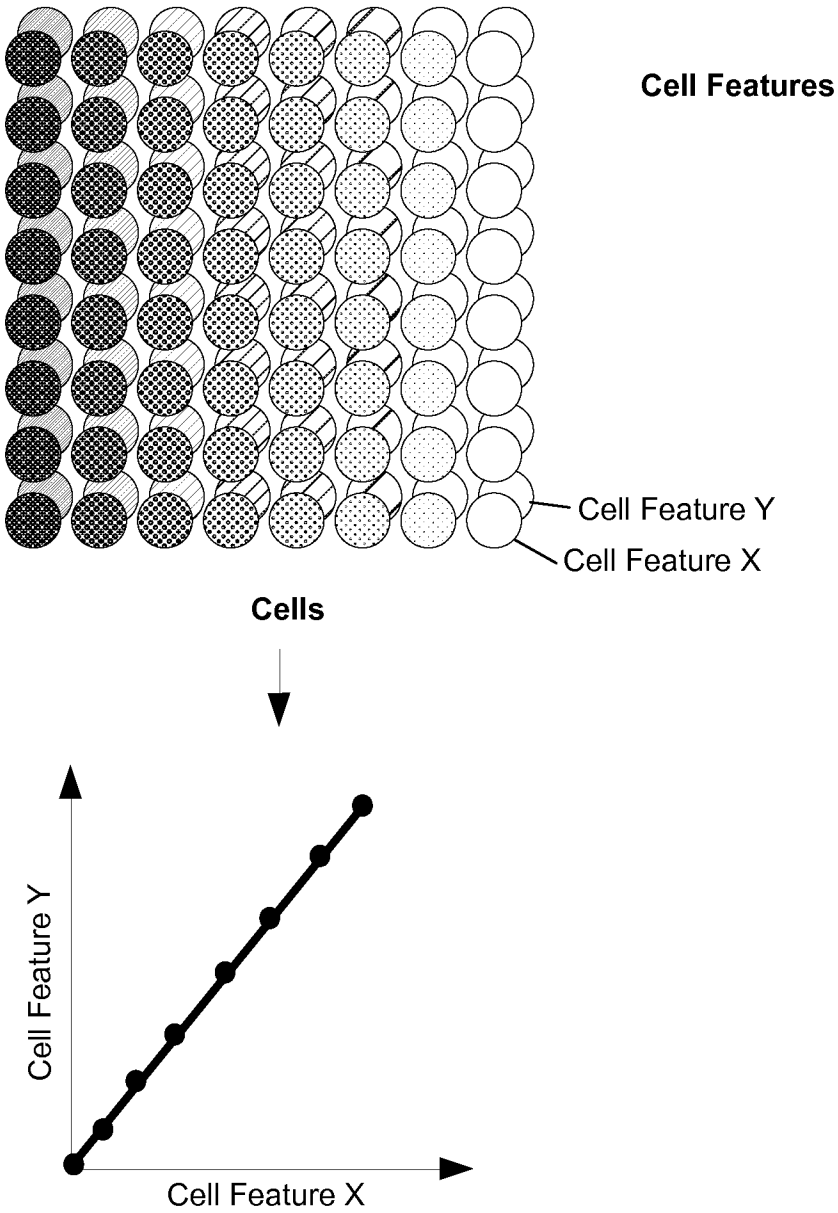

Cell Features
Cell Feature Y
Cell Feature X
Cells
Cell Feature Y
Cell Feature X

Fig. 15

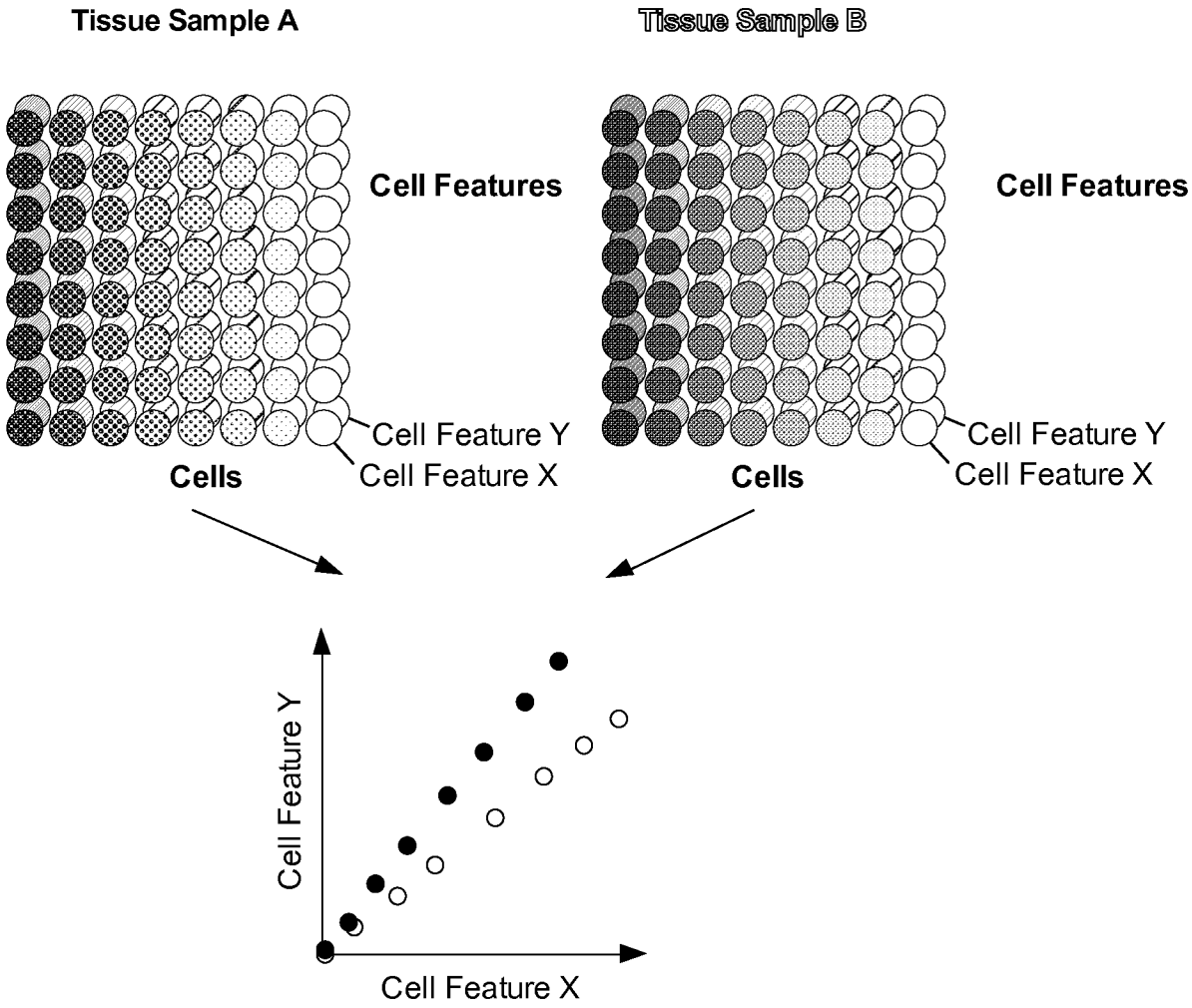

Tissue Sample A
Tissue Sample B
Cell Features
Cell Features
Cell Feature Y
Cell Feature X
Cells
Cells
Cell Feature Y
Cell Feature X
Cell Feature Y
Cell Feature X

Fig. 16

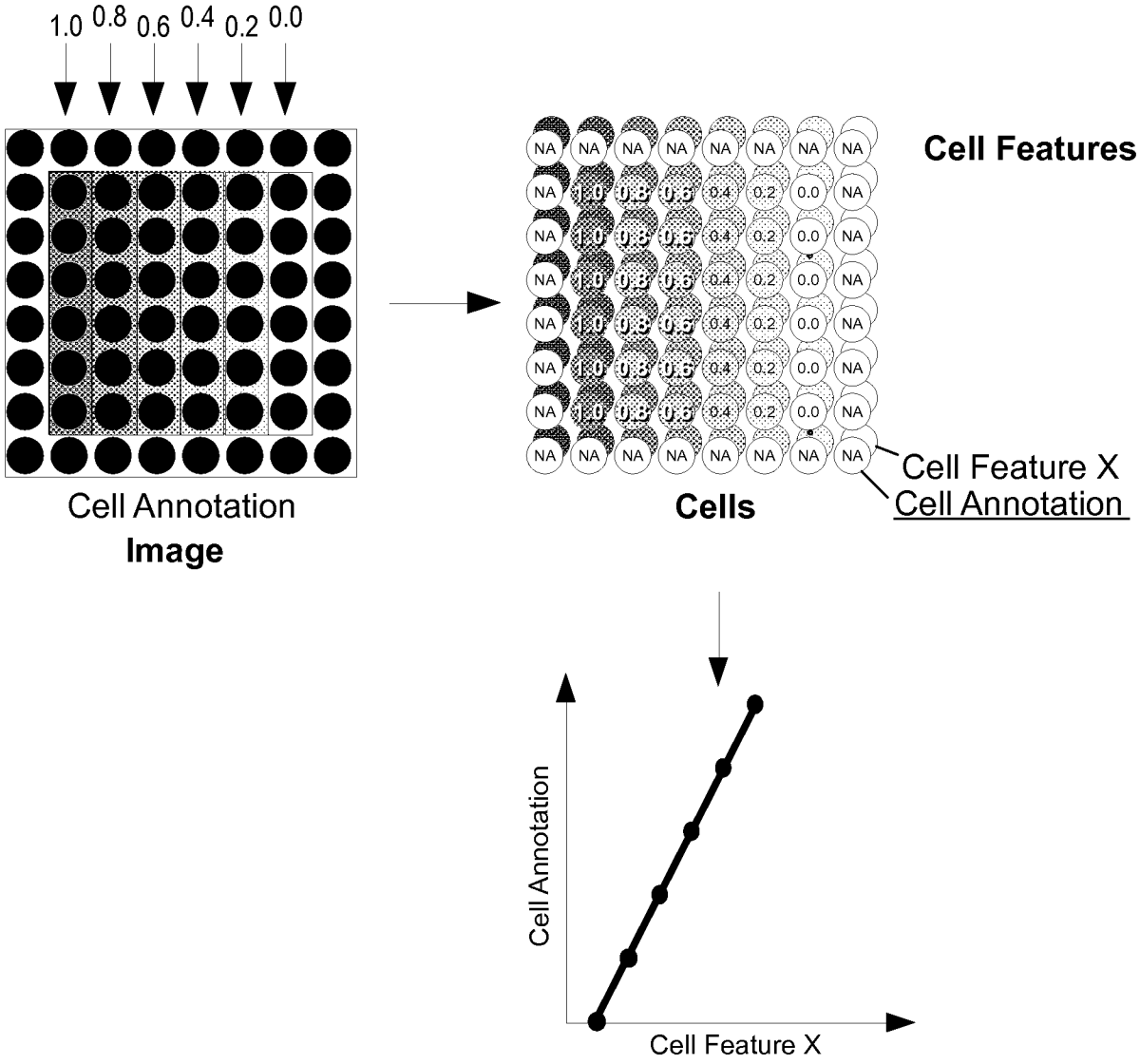

1.0 0.8 0.6 0.4 0.2 0.0
Cell Annotation
Image
Cell Features
Cells
Cell Feature X
Cell Annotation
Cell Annotation
Cell Feature X

Fig. 17

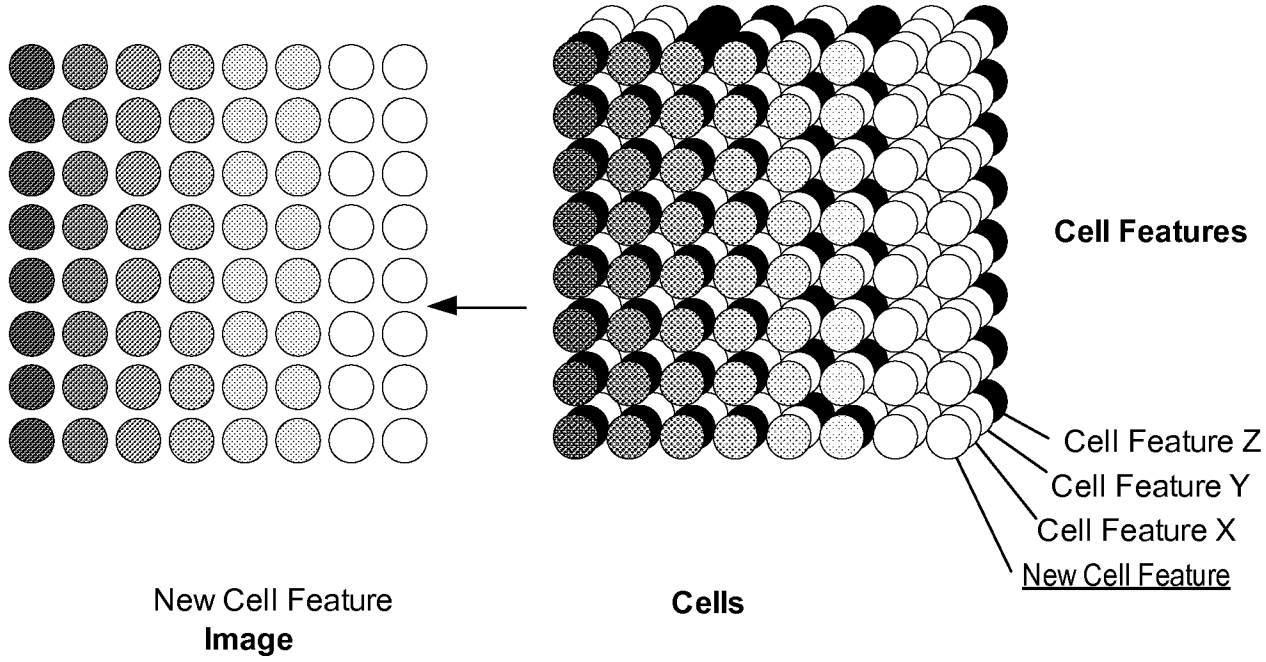

Cell Features
Cell Feature Z
Cell Feature Y
Cell Feature X
New Cell Feature
New Cell Feature
Image
Cells

Fig. 18

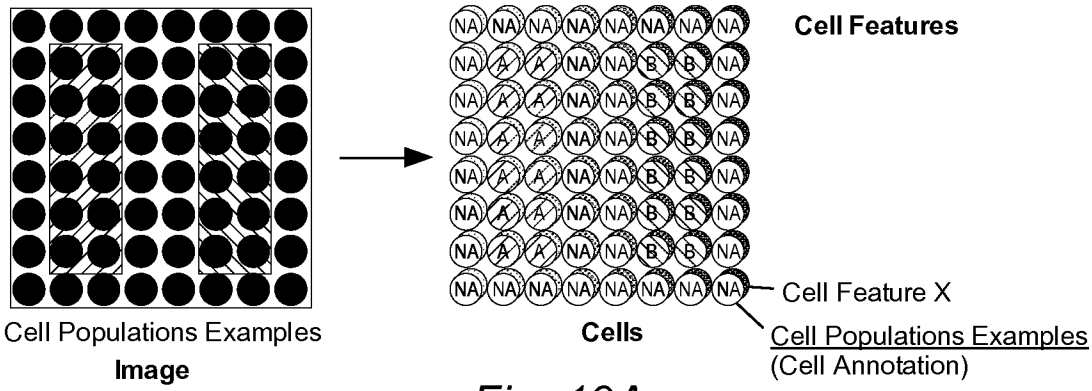


Fig. 19A

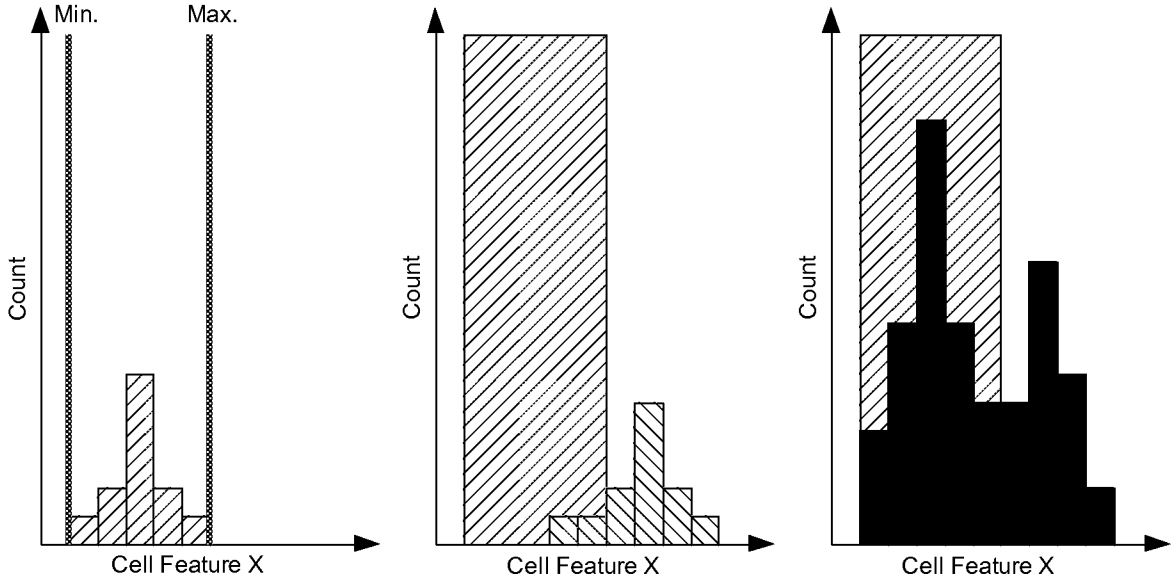


Fig. 19B

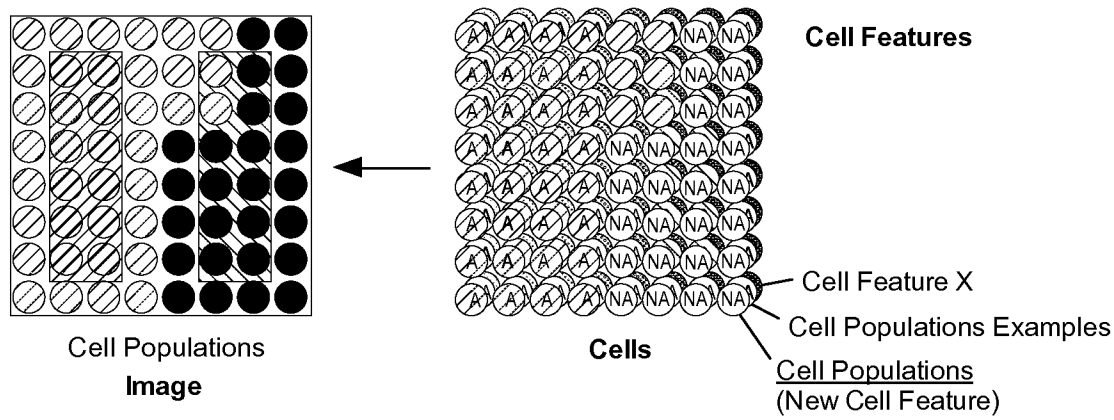


Fig. 19C

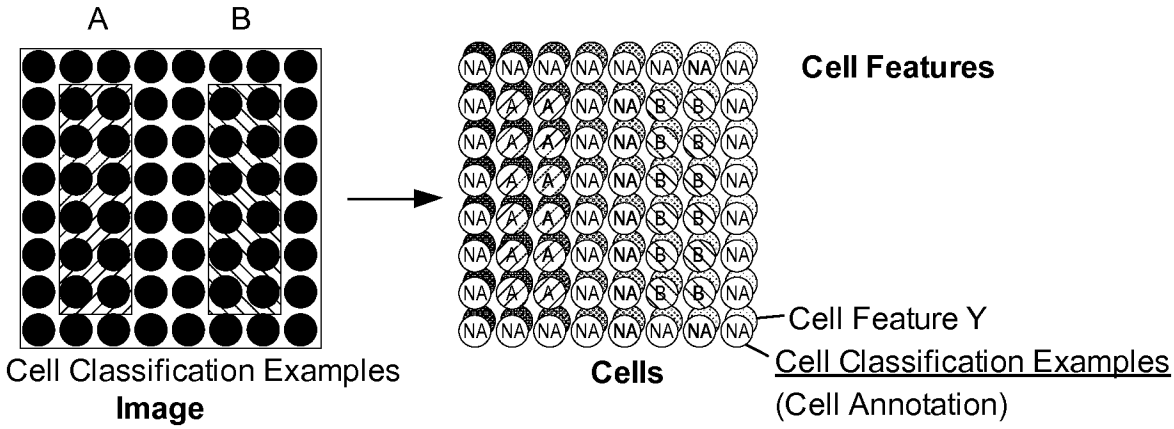


Fig. 20A

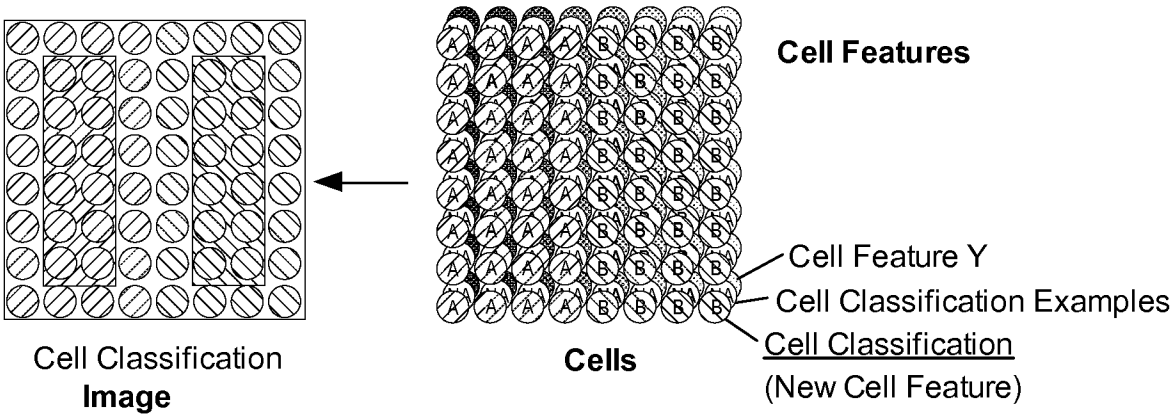


Fig. 20B

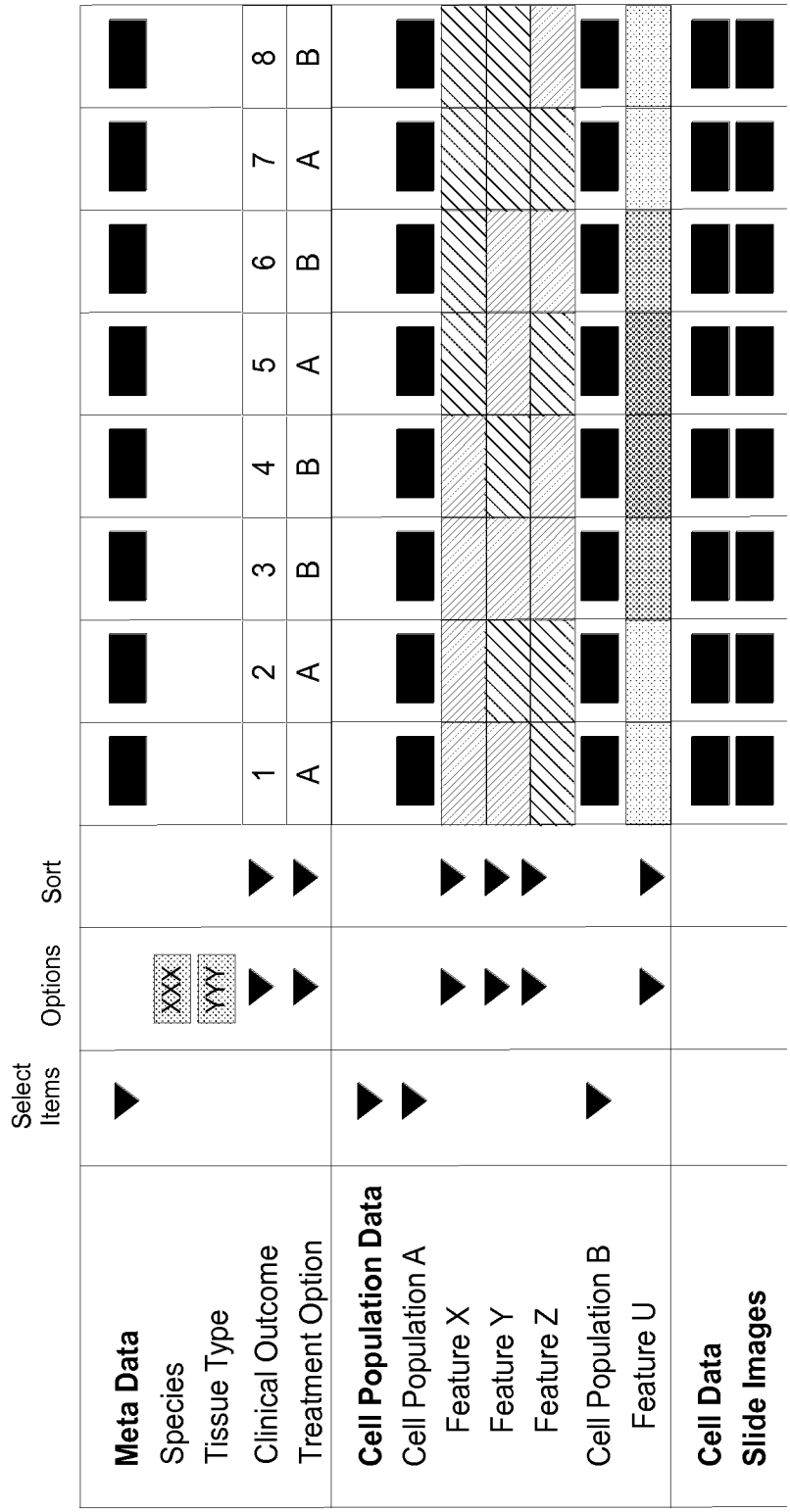

Select Items
Options
Sort
XXX
YYY
Meta Data
Species
Tissue Type
Clinical Outcome
Treatment Option
Cell Population Data
Cell Population A
Feature X
Feature Y
Feature Z
Cell Population B
Feature U
Cell Data
Slide Images
1 2 3 4 5 6 7 8
A A B B A B A B

Fig. 21

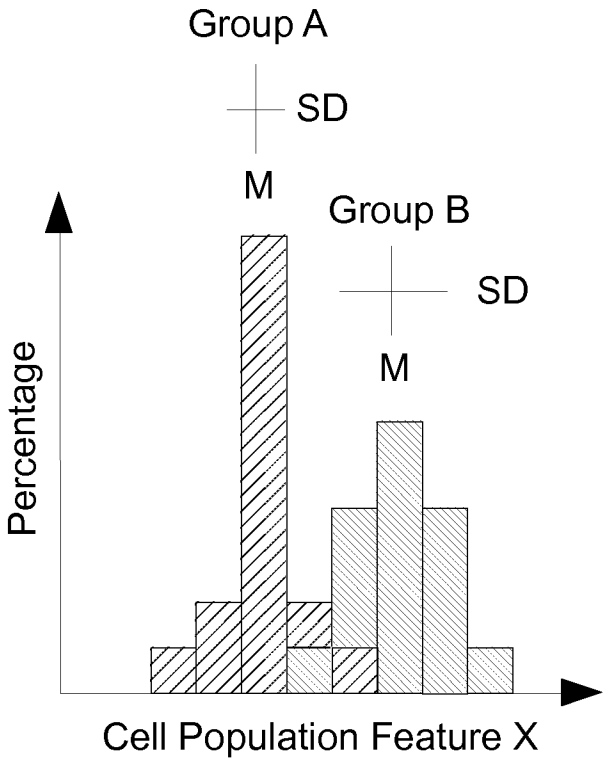

Group A
SD
M
Group B
SD
M
Percentage
Cell Population Feature X

Fig. 22

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US 2014/018322

A. CLASSIFICATION OF SUBJECT MATTER

(see extra sheet)

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G01N 1/28, 1/30, 1/42, G06K 9/68, 9/36, 9/54, 9/612

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CIPO, DWPI, EAPATIS, EMBL,E ESP@CENET, PAJ, PubMED, PatSearch (RUPTO internal), USPTO,WIPO

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	RU 2332223 C2 (HENRY FORD HEALTH SYSTEN et al) 27.08.2008, pp. 23-25	1-18
Y	US 2012/0076390 A1(FLAGSHIP BIO) 29.03.2012, pp.1-7, claims	1-18
A	US 2005/0111758 A1(HOLGER LANGE et al) 26.05.2005, pp.1-3, claims	1-18
A	US 2008/0240558 A1 (STI MEDICAL SYSTEMS, LLC) 02.10.2008, pp.1-6, claims	1-18

☐ Further documents are listed in the continuation of Box C.

☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
06 June 2014 (06.06.2014)	19 June 2014 (19.06.2014)
Name and mailing address of the ISA/ FIPS Russia, 123995, Moscow, G-59, GSP-5, Berezhkovskaya nab., 30-1 Facsimile No. +7 (499) 243-33-37	Authorized officer O.Kolontaevskay Telephone No. 495 531 65 15

INTERNATIONAL SEARCH REPORT
Classification of subject matter

International application No.
PCT/US 2014/018322

G01N 1/28 *(2006.01)*
G01N 1/30 *(2006.01)*
G01N 1/42 *(2006.01)*
G06K 9/68 *(2006.01)*
G06K 9/36 *(2006.01)*
G06K 9/54 *(2006.01)*
G06K 9/62 *(2006.01)*