the diagnostic pathology journal DIAGNOSTIC PATHOLOGY

Kayser et. al. diagnostic pathology 2015, 1:14
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:14

Analysis of texture and objects in microscopic images

Klaus Kayser[1*], Stephan Borkenfeld[2], Amina Djenouni[3], Gian Kayser[4]

*- corresponding author: Institute of Pathology, Charite, Charite Platz 1, D-10118 Berlin, Germany, email: klaus.kayser@charite.de
1- Humbold Unitersity, Berlin, Germany
2- International Academy of Telepathology, Heidelberg, Germany
3- Institute of Pathology, Batna, Algeria
4- Institute of Pathology, University of Freiburg, Freiburg, Germany

Abstract

Background:

Tissue based diagnosis or surgical diagnostic pathology undergoes significant changes and focuses on image content analysis in our days. Herein we describe and discuss new approaches of content image analysis and compare their applications, benefits and constraints.

Theory*:*

Any useful microscopic image contains information that can be evaluated and transferred into a tissue-based diagnosis. A correctly derived diagnosis depends upon the image information and the pathologist's knowledge, i.e. his ability to recognize and transfer the image content information into clinical application. Thus, image information is related to external "disease" information, i.e. interpretation, and "pure" image content information, which the pathologist has to interpret. Application of external image information requires definition and separation of objects from the background, or segmentation procedures. Observer free image information is solely pixel based. It can be analyzed using different approaches, such as entropy measure, construction of image primitives and their spatial distribution, or image similarity operations. Our approach uses entropy calculations dependent upon all possible gray value thresholds in combination with syntactic analysis of pixel based image primitives.

Implementation:

Virtual slides underwent the evaluation of regions of interest (ROI) as described previously. ROIs were interactively controlled and subject for application of developed image analysis procedures. The "classic method" of object recognition and syntactic structure analysis is applied too. Trials were performed on anti Her2_new stained, DAB visualized, and glycohistochemically stained, AP visualized slides. The images were measured at magnifications, which correspond to x20, and x40 objectives.

Results:

The algorithm displayed only weak changes of the evaluated gray level based structural entropy (GL-MST)-entropy in the selected ROIs in contrast to the whole image. In

the diagnostic pathology journal DIAGNOSTIC PATHOLOGY

Kayser et. al. diagnostic pathology 2015, 1:14
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:14

addition, significant differences could be obtained when the measures were associated to the clinical impact (diagnosis, fetal developing stage).

Conclusions: Images textures and pixel primitives can serve for evaluating "pure" image content information. They do not require segmentation and additional external information for measurement. Interpretation of the measures, i.e. external information can be implemented at a later stage. The described algorithm is probably applicable for image analysis of different fields such as histology, forests, traffic, or can serve for objective image quality evaluation.

Key words: tissue-based diagnosis, content image analysis, MST entropy, immunohistochemistry, glycohistochemistry, segmentation.

Virtual Slides: www.diagnosticpathology.eu/vs/2015_1_14/

Introduction

The development and clinical application of virtual slides induced increasing interest in image analysis of histological / microscopic slides [1, 2]. The vendors of slide scanners are forced to develop appropriate image analyzing systems in addition to suitable viewers and data storage systems [3]. Most commercially available virtual microscopy systems include some kinds of image measurements such as cell counting, gray value (intensity) measurements, or calculation of clinically used scores (for example Her2_new) [4].

They try to introduce virtual microscopy which is the diagnostic work with virtual slides rather than with a conventional microscope, in clinical practice. Missing generally accepted standards and appropriate connection to the laboratory information system (LIS) are the main constraints of virtual microscopy at present [5]. However, serious efforts are ongoing to overcome these constraints in the near future [6].

Having solved these contrivers conditions, digital pathology will open a new door, namely the development and implementation of so-called diagnostic assistants [7]. Diagnostic assistants are tools that are comparable to APPs, and support the pathologist in his daily work [8]. In a first step, all of these developments try to directly copy the "conventional", i.e., non – digital diagnostic procedures and to translate them into the digitized workplace [8, 9].

Herein, we describe a different approach: Coming from a basic analysis of a diagnostic procedure we assume that several of the contributing compartments (functions) act independently from each other. They can, therefore, be implemented at different stages of the diagnostic procedure. Whether this "dismounting" will result in additional difficulties or, what we assume, in favorite solutions, cannot be answered at this stage of development. In principle, our proposal can be compared with Grid or APP applications instead of a whole, all issues including "diagnosis machine".

Theory

Visual and acoustic information transfer is the main component of any human communication [10]. Both of them can be digitized, transferred, stored, and retrieved in space and time independently [11]. In medicine, visual information is the main source of disease recognition

Kayser et. al. diagnostic pathology 2015, 1:14
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:14

and classification as well as treatment and quality evaluation of the performed applications [10]. The basic units of our understanding of diseases include a hierarchic arrangement body compartments, that start with the organs, and are followed by cellular basic functional units (vessels, nerves, glomeruli, etc., by cells of different functions (Virchows cellular theory), cellular subunits, such as chromosomes, reticulum, and finally break down to genes, receptors, proteins, ion channels, macromolecules, electric charges [8]. All these different and interacting elements possess a certain "function", i.e. act "as a whole at an individual level". In the 1980th a general theory has been developed, that analyzes and orders the different basic elements reproducible in a hierarchic order. This theory can be expanded to a general theory of diseases and diagnosis [12, 13].

A conventional image analysis searches for objects, for their potential morphologic changes and spatial arrangement [10]. An image can, therefore, be divided into an object space and a background [14]. Assuming the cell as object or basic element in a microscopic image, its information I is a function of objects (f(o)) and their spatial distribution (g(s) structure).I = f(o) & g(s).

The defined objects depend upon the knowledge of the pathologist, i.e., the receiver of the information. The function remains identical even if we define trees or houses, or mice as objects. It also holds true for image sequences such as movies or image comparisons.

We can also go back to the smallest information unit of a digitized image, the pixel. A pixel is characterized by its location and its gray value.

Objects are characterized by a set of pixels that fulfills a spatial (commonly neighborhood) and a gray value condition. The conditions that form a set (object) can only be derived from image information itself if additional (external) features of the image or of the object are known. These features might be precise (upper and lower limits of length of circumference, gray value difference limits within a selected object, etc., or only crude, such as microscopic image of liver biopsy, etc., magnification 10 microns/pixel, or even only a function: please compare the image with a set of images of known origin, that possesses objects which are declared cells, trees, mouse, etc.

Remarks: The reader might recognize that this situation reminds of the well known sentence: You do only see what you already know (Goethe), or of the general problem of recognition processes [15]: How to collect knowledge of events within a closed room (space) whose interpretation cannot be derived from already known information?

The derivative of the function (1.1) induces the common procedure: Do not try to derive information from an image that we do already know; i.e. the common pathologist's answer to digital pathology: Why detect a cell (or nucleus) by a machine when I can see it immediately? Thus, we train the machine to detect a cell (nucleus) and try to compete with the pathologist in speed and reliability. Is this a good approach?

If we go back to the pixel frame, we might avoid this speed and reliability competition between man and machine: We just let the machine work for us, nothing else. Similar to a horse or a car, that is serving us whether it wants or not.

If we consider the term object and distinguish in between meaning which is the (external) information (object cell, nucleus, tree, etc.) and function f(o) which is used to derive objects from the background (pixel conditions of an object), we can then separate q and try to define f(o) by different features such as entropy after a set of transformations, or sets of entropies at

the diagnostic pathology journal DIAGNOSTIC PATHOLOGY

Kayser et. al. diagnostic pathology 2015, 1:14
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:14

different gray value thresholds, or to compose the function q(o), which defines the meaning of an object as follows:

I = q(o) g(s) = q(f(o)) & g(h(s)) = i(f(o) &h(s)); where I is the meaning of the pixel derived features. In other words, we do not have to define objects, meaning etc. when we want to detect image information.

The same consideration holds true if we want to find the "most significant image area; i.e., the image area which contains the "most significant information" for the interested viewer (diagnostic pathologist). We would expect that the region of interest (ROI) displays a nearly homogenous distribution of adequate selected descriptive features in contrast to the whole image that probably displays different values of these features, and a broad variation. Thus a careful selection of pixel related features can be used for proof of correct definition of ROIs [16].

The principal advantages of the described approach include

a) Fast and easy to implement measures that are independent from any meaning or external information, i.e., solely associated with the image content information;
b) Reproducible detection and proof of ROIs;
c) Development of principle or easy to standardize image parameters that are not dependent upon the viewers knowledge, etc.;
d) Application in a broad field of sciences, namely potential all fields that rely on visual information analysis;
e) Potential transfer of the algorithm to a "higher level" of computation, i.e., replacement of fixed measures by functions (flexible boundary parameters).

The constraints of this approach include:

a) No direct visualization of commonly appreciated objects such as cells, nuclei, etc.;
b) No visualization of objects that are unknown to the viewing pathologist;
c) Difficult interpretation of newly detected object associated measures;
d) Demands of strict definition of parameter contribution on clinical decision procedures
e) Potential object – independent classification of diseases with hard-to-foresee impact on treatment.

How to implement the proposed algorithm?

Implementation

Prerequisites

It might be assumed that all commercially available whole slide scanners acquire standardized and reproducible digital images. This seems, however, not to be realized in reality, and numerous virtual slides display with non-negligible aberrations in terms of illumination, color space and glare effects [4, 17, 18]. Thus, prior to applications of automated information detection and extraction the acquired digital images have to be standardized for at least three different parameters:

1) Homogenous illumination or shading. An example is demonstrated in <Figure 1>. Shading correction (vignette) can be performed using the differentiated image or using a standard correctly balanced image [19].

the diagnostic pathology journal DIAGNOSTIC PATHOLOGY

Kayser et. al. diagnostic pathology 2015, 1:14
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:14

2) Adjustment of the gray value range in all three color spaces for the maximum and mean gray value range. This procedure should be performed after the shading correction, and is needed to balance between different laboratory effect such as thickness of the tissue cut, principal staining intensity(age of the dye), etc.
3) Adjustment of gray value distribution in order to avoid any artificial gray value clusters. They can be induced by non homogenous color sensitivity of the scanner chip.

The calculation of the number of potentially useful thresholds in order to separate the object space from the background is a useful step for a) confirming suitable image quality and b) defining mandatory gray value thresholds in relation to the wanted objects [10]. It is less

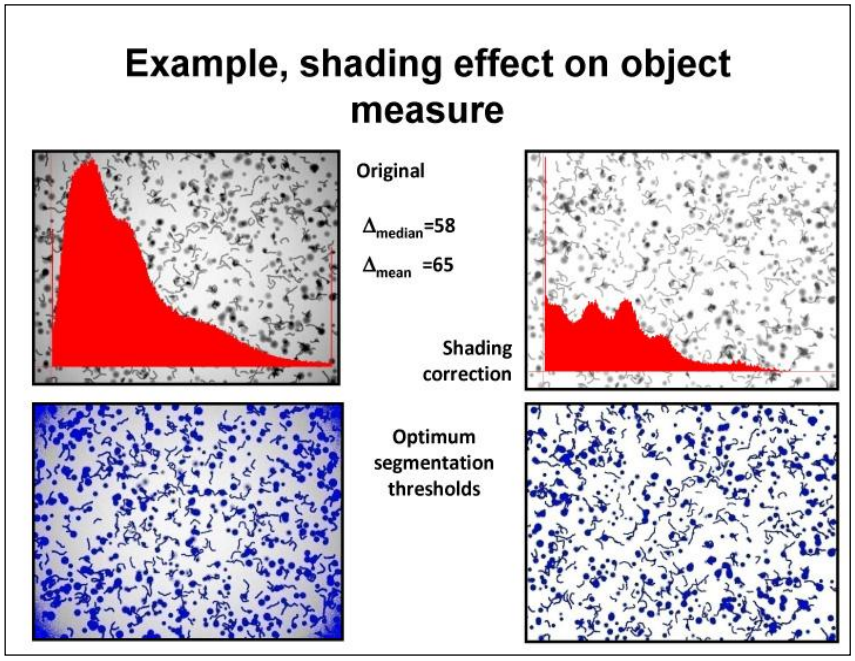

Figure 1: Influence of shading correction on potential segmentation thresholds.

known and less applied in general. The number of statistically useful thresholds (i.e., >95% confidence areas) accounts to three in IHC and conventionally stained images, and to 2 in one-marker fluorescent images, as shown in <Table 1>. An IHC image that contains only 1 – 2 detectable thresholds is probably of poor image quality or of inadequate technical preparation [20].

The next step addresses information distribution within the whole VS, or the detection of ROIs. Several regions might be present in the image, and all of them should serve for additional analysis. Several approaches have been published in the literature [4, 10, 17, 21-24]. The appropriate ROI detection algorithm should be selected to the aim, whether an interactive or an automated diagnosis is addressed to. Interactive diagnosis on ROIs with fixed size and frames is to our experience easier and faster to perform than to evaluate diagnoses on ROIs of variable sizes. Fixed ROI sizes permit, in addition, the calculation of entropy flow between the different neighboring frames; an option, that cannot be calculated on variable areas due to missing area standardization (It might be reminded that entropy calculations depend on the size of the selected space) [10]. Variable ROI sizes are appropriate for automated diagnosis in difficult

the diagnostic pathology journal DIAGNOSTIC PATHOLOGY

Kayser et. al. diagnostic pathology 2015, 1:14
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:14

cases. Based upon structure – associated parameters their boundaries separate diagnosis – relevant areas from diagnosis irrelevant areas more distinctly. In case of image comparison approaches fixed ROI areas are a prerequisite in order to use similarity operations between the source and the test ROIs [10]. The same holds true for pixel based texture approaches. Again, several image transformations can only be performed on square sized images (for example Hough or recursive transformation) [27].

Table 1: *Number of potential segmentation thresholds of IHC (DAB, AP) and fluorescent visualization*

Image quality: number of thresholds & segmented areas

Stain	Number of thresholds	Area covered 1st threshold	Area covered 2nd threshold
H & E, N=1400	3	54%	10 %
DAB, N=1200	3	63 %	12 %
AP, N=400	3	45 %	11 %
DAPI, FITC, TEXAS RED N=120	2	35 %	5 %

Explanations: H & E: hematoxilin – eosin; DAB: diaminobencidine; AP: Alcalic phosphatase; DAPI & Texas Red & FITC: tri-colour fluorescent stain, 5% maximum/minimum significance.

An additional issue of prerequisites arises which color space should be selected for image analysis. Virtual slides are usually displayed in the red-green-blue (RGB) color space, and most of commercially available image analysis systems use the RGB space [25]. Any color space can be transformed in others without artifacts and quality losses [26]. An additional frequently used HSI color space separates the intensity of any color into a separate intensity coordinate, and transforms the RGB colors in hue and saturation coordinates [26]. This color space is practicable if intensities of a specific color should be measured, for example the “brown” intensity of DAB visualized IHC stained glass slides [20]. On the other hand, the RGB color space does not need to be transformed in most of immune fluorescent images [27]. It might also be useful to transfer the three color axes into a gray value image by a principle component transformation, especially, when a one-dye IHC stain has been performed, and simple structures such as membranes or (Her2new) or nuclear counting (MIB-2) are analyzed [8, 11, 16, 20].

The image (ROI) size should be constant for the whole experiment, i.e. for all images that are evaluated and contribute to statistical analysis [8, 11, 16]. The minimum image size depends upon the proposed measurements. For object analysis such as nuclear size, form factor, length of stained membranes, etc. an object size of minimum 1000 pixels and a minimum of 100 objects within the ROI are mandatory if a confidence space >0.95 will be assessed [19]. Object analysis can be performed on images of any form, i.e., rectangular, ellipsoid, circular, triangular, or others

Kayser et. al. diagnostic pathology 2015, 1:14
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:14

in contrast to texture analysis which requires square image sizes for several transformations. Texture analysis needs a minimum ROI size of 256 pixels for reproducible measurement results [10, 27].

Preparation of measurements

Any image analysis should be carefully prepared. The objective includes practical issues that have been mentioned in the previous chapter, and, in addition, theoretical considerations. The principles are explained in<Table 2>.

Any image analysis is a communication procedure between the image (information source); and the beholder (receiver, pathologist). The visual image signals such as colors, their distribution, arrangement, intensities have to be recognized or understood by the viewer. In histopathology the pathologist has to recognize and to classify nuclei, cells, vessels, etc. In image analysis the measurement program “replaces” the pathologist. The outcome of this procedure relies on the pathologist’s knowledge. Therefore, image information can be separated in two terms, namely information that is independent from the receiver, so – called image content information, and information that is acquired by the receiver (external image information) [10, 27]. External information is mandatory to define and consecutively separate objects from the background. This is the principle of object and structure based image analysis [10, 27].

Table 2: Differences between object/structure and texture related image information acquisition

Texture – Object related Diagnosis

Feature	Identification Method	Derived Parameters
Object	Segmentation, Spatial Analysis (Measurement)	Biological classification (cell type), Quantification (size, gray values)
Structure–object derived	Syntactic Structure Analysis	Orders of structure, biological function (vessel, gland, etc)
Texture	Autoregression model Texture primitives	Basic patterns, Entropy gray value calculation

External information can also be used to create “artificial” objects, for example insertion of objects into “empty” image areas. These procedures require additional assumptions, for example a similarity approach. Missed objects should be similar to those of their nearest neighbors, empty spacer significantly larger than the average of “normal” background areas. An example is shown in <Figure 2>.

Textures can be analyzed without any extern knowledge [10, 27]. Its basic principle is the gray value distribution of pixels associated to the pixel position in the image, and its association to

the diagnostic pathology journal DIAGNOSTIC PATHOLOGY

Kayser et. al. diagnostic pathology 2015, 1:14
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:14

neighboring pixels. A pixel (or voxel in a three-dimensional space) can be defined as "object without background space".

Pixel gray values can be subject to thresholds and classification in pixel that posses gray values above and below the threshold, or which exceed the lower and fall below the upper threshold [11, 16]. This is the first step of applying external knowledge to image information. The next

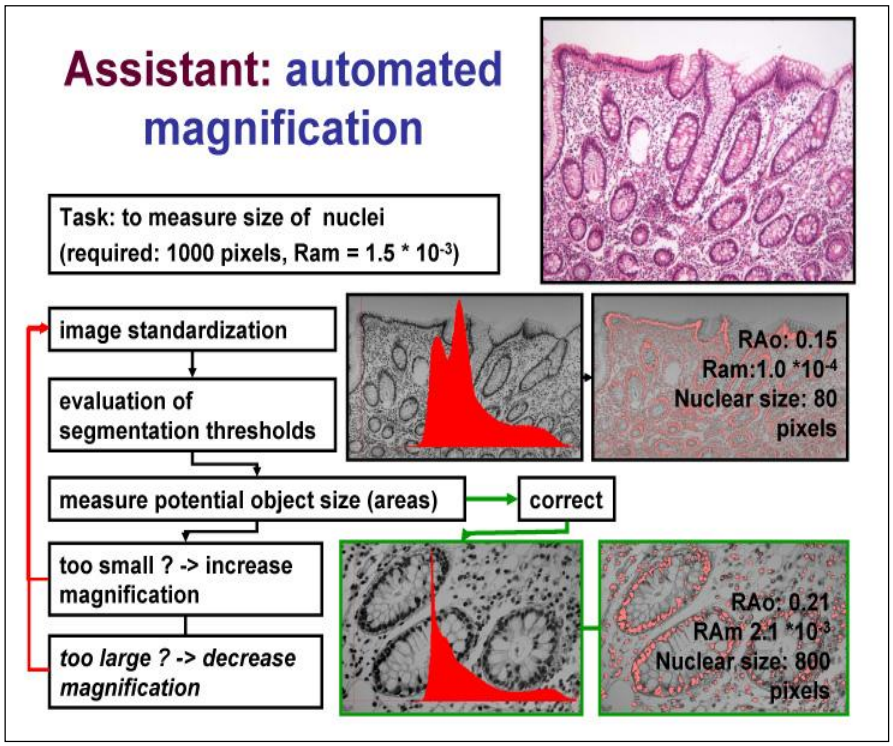

Figure 2: Example of an automated magnification (object size adjustment) diagnosis assistant.

step can be the introduction of neighborhood conditions that create so – called image primitives. Individual point, fibers, circles, plateaus, crosses, or curves are frequently used primitives. They are the "letters" of the image and can be synthesized to "words" and "sentences". "Words" can then be translated into objects, "sentences" into structures. This procedure is called syntactic structure analysis [13]. It has been reported to be a powerful image analysis tool, especially when it is combined with biological meaningful parameters such as Shannon's entropy, structural entropy (MST entropy) or MST entropy flow [28-31]. The principle of syntactic structure analysis for object derived structures is shown in <Figure 3>.Texture analysis can be applied to transformed images in addition to the source image [16]. Well known are Fourier transformations, which are often applied on objects, and interpreted by chaos theory assumptions. Additional transformations include Hadarmad, Hough, and recursive (analogue time series) transformations. For details see [10, 11, 16].

The optimum range of image, object, structure, texture features in relation to the proposed measurement is shown in <Figure 4>.

the diagnostic pathology journal DIAGNOSTIC PATHOLOGY

Kayser et. al. diagnostic pathology 2015, 1:14
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:14

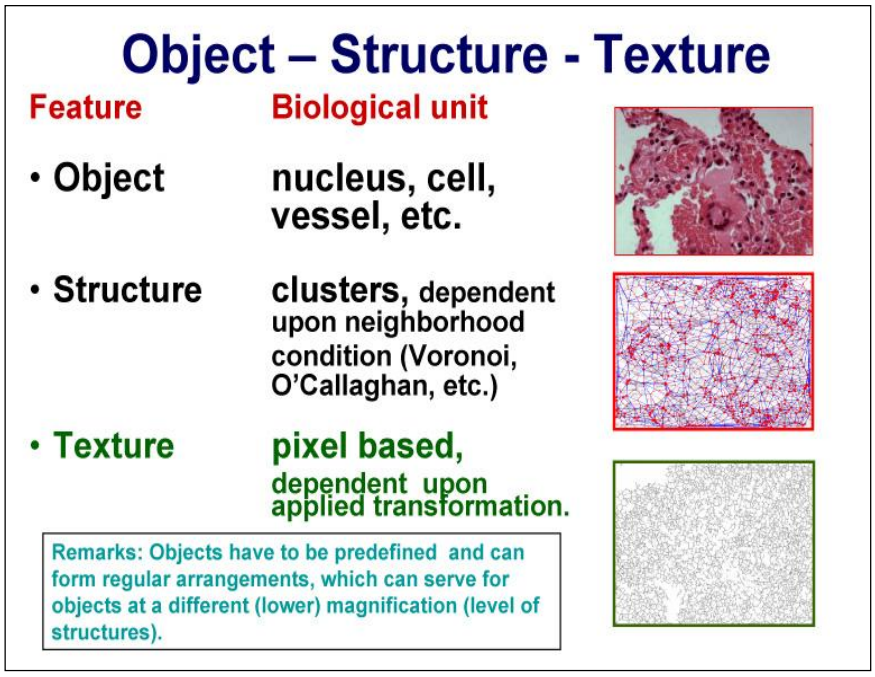

Figure 3: Principle of syntactic structure analysis on object derived structures.

Texture and object related image information

Microscopic images address to objects and their structures in general. Basically, objects include nuclei, cells, nuclear compartment and deposits of extra-cellular material such as mucus, amyloid, asbestos fibers, or invaded living organisms such as fungus, bacteria, or parasites foreign living organisms. Most of them are visible by HE stains, some require specific stains, such as PAS, Silver, Elastica van Gieson's, Ziehl Neelson, or other stains. The more

precise and accurate the laboratory technique is performed, the easier the pathologist can identify and classify the objects and their arrangement. Therefore, digital pathology should provide the pathologist with tools that allow an optimum visual either interactive or automated detection of potential objects. These image quality improving techniques are limited and comprise white balance, color filters, polarization, and usually only 4 – 5 different magnifications (objectives) only when working with a conventional microscope.

These tools can be extended when working with digital images. They include concentration of objects (or their structures only (ROI, and adjusted size of image area to be viewed), several small magnification steps, artificial and feature – adjusted coloring of detected objects and structures, enhancement of object boundaries, for example by image differentiation or other transformations, contemporary display and viewing of images that are stained with different dyes, or have been already diagnosed in the past (recurrent disease, metastasis, etc.). All features that are associated with the different (classes of) objects can be traced down to three different basic measurements, namely of its area, of its circumference, and of the pixel gray value distribution within the object [10, 11, 16]. Therefore, appropriate image standardization should include an adequate object size (magnification) in addition to well known image corrections, such as shading, etc. [10, 11, 16].

the diagnostic pathology journal DIAGNOSTIC PATHOLOGY

Kayser et. al. diagnostic pathology 2015, 1:14
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:14

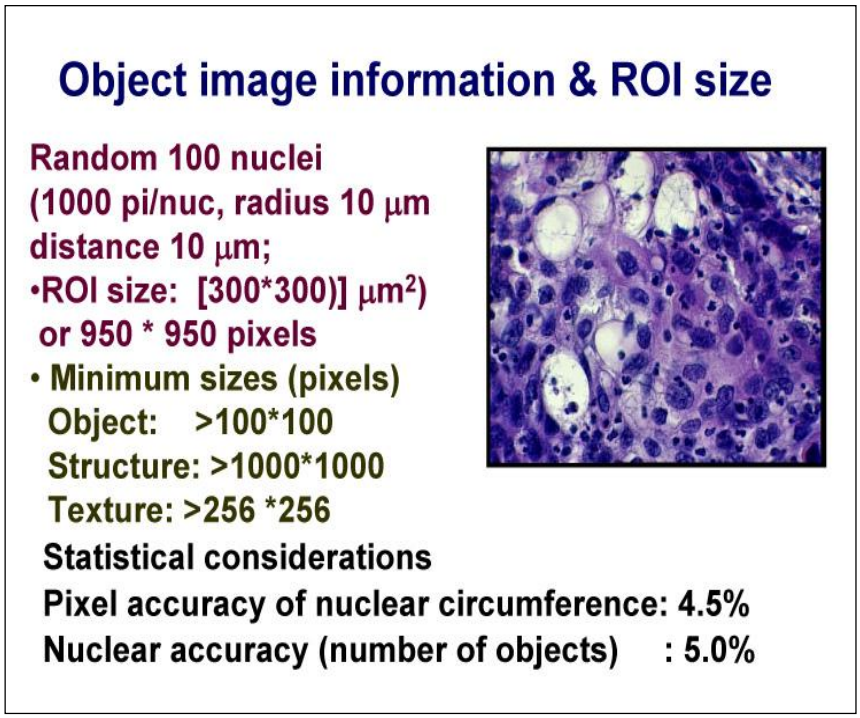

Figure 4: Optimum size of whole image, objects, structures, and textures in relation to the proposed measurement.

Diseases display distinct, however not univocal changes in the tissue of the involved organ. These changes are induced by changes of cells and their compartments, such as chromosomes, genes, receptors, or others. Thus, object – associated image analysis has direct access to functional changes of the underlying "biological units", if chosen correctly [8]. The objects correspond to membranes (IHC), and amplified genes (FISH). Their position and intensity are measured, and serve for additional derived parameters (evaluation of macrostages, microstages, neighborhood). The derived parameters undergo a statistical analysis in association with a proposed interactive grading of membrane receptors and gene products which can easily be judged for its meaning.

Pixel – derived image information

Pixel derived information is an analysis of pixel gray values in relation to the pixel position. It results in a vector (x,y,g) or (x,y,z,g) in case of a three dimensional image (voxel). What kind of features can be derived? They include

a) pixel primitives
b) neighborhood analysis associated to specific grays levels
c) neighborhood analysis of related gray value levels
d) higher order derivatives such as gray value level <> position <> primitive analysis.

The analysis of pixel primitives requires a neighborhood condition, such as Voronoi's or O'Callaghan's tessellation. A potential procedure can be defined as follows: calculate the median of the image's gray value distribution. Define the median as threshold, and use the neighborhood condition to identify: 1) isolated points above and below the threshold {+p,-p}; b) connected fibers {+f(l),-f(l)}; c) crossing points {+cr,-cr}, closed circles {+c(l),-c(l)}, plateau (+p(a),-

the diagnostic pathology journal DIAGNOSTIC PATHOLOGY

Kayser et. al. diagnostic pathology 2015, 1:14
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:14

p(a)). (+) means at and above, (-) below the chosen threshold, (l,a) means the number of pixels within the primitive.

The set of primitives can be used for numerous calculations such as Shannon's entropy (S= Σ[p*ln(p)], p = probability of a primitive within the whole set, weighted by length; or structural entropy:

$E(MST)=p(MST)\ \Sigma[p(MST)*\ln(p(MST))]$; $p(MST)= d[(x,y)*(a,l)]^2/mean(x,y,)*mean(a,l))^2$

The described algorithm can be extended by the introduction of several (segmentation minima, or all) gray values (1,…256), and by neighborhood calculations that include a certain "range" of neighborhood, for example an upper distance which limits the range of a neighborhood, etc. [10, 16, 27].

.

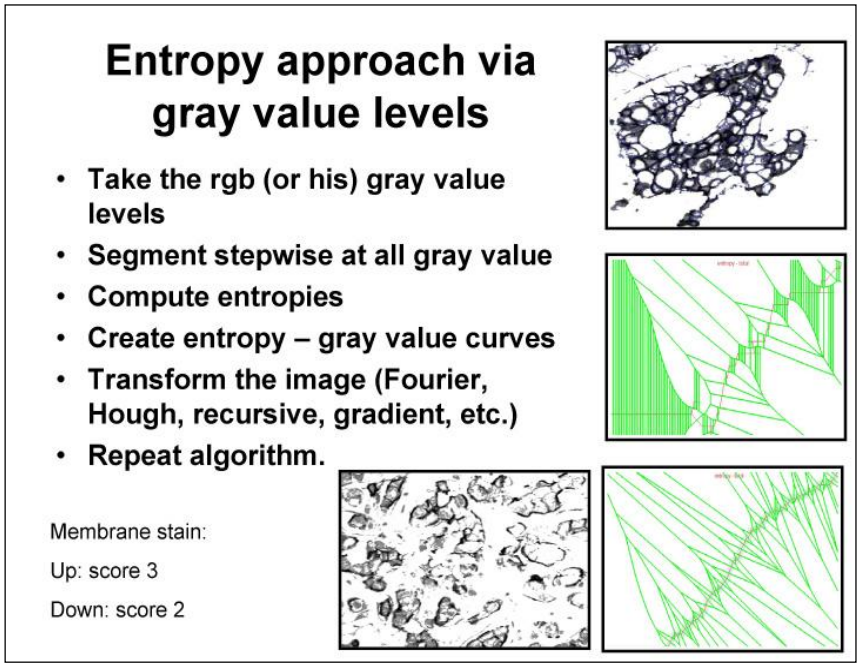

Figure 5: Example of pixel primitive approach on Her2_new (DAB) stained slides (score 2 and score 3) and derived graphs of MST entropies of gray value discrimination.

The principle is explained in <Figure 5> for evaluation of predicitive therapy (Her2_new scoring). Herein, the object related images (membranes) segmented for staining intensity and the pixel derived MST entropies are shown. Both techniques give the same results.

The consecutive statistical calculations can then be applied to calculate the gray value level with the lowest MST entropy, its maximum, its smoothest/steepest increase/decrease if it is assigned to the correspondent gray value level. An example is shown in <Figure 6>.

the diagnostic pathology journal DIAGNOSTIC PATHOLOGY

Kayser et. al. diagnostic pathology 2015, 1:14
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:14

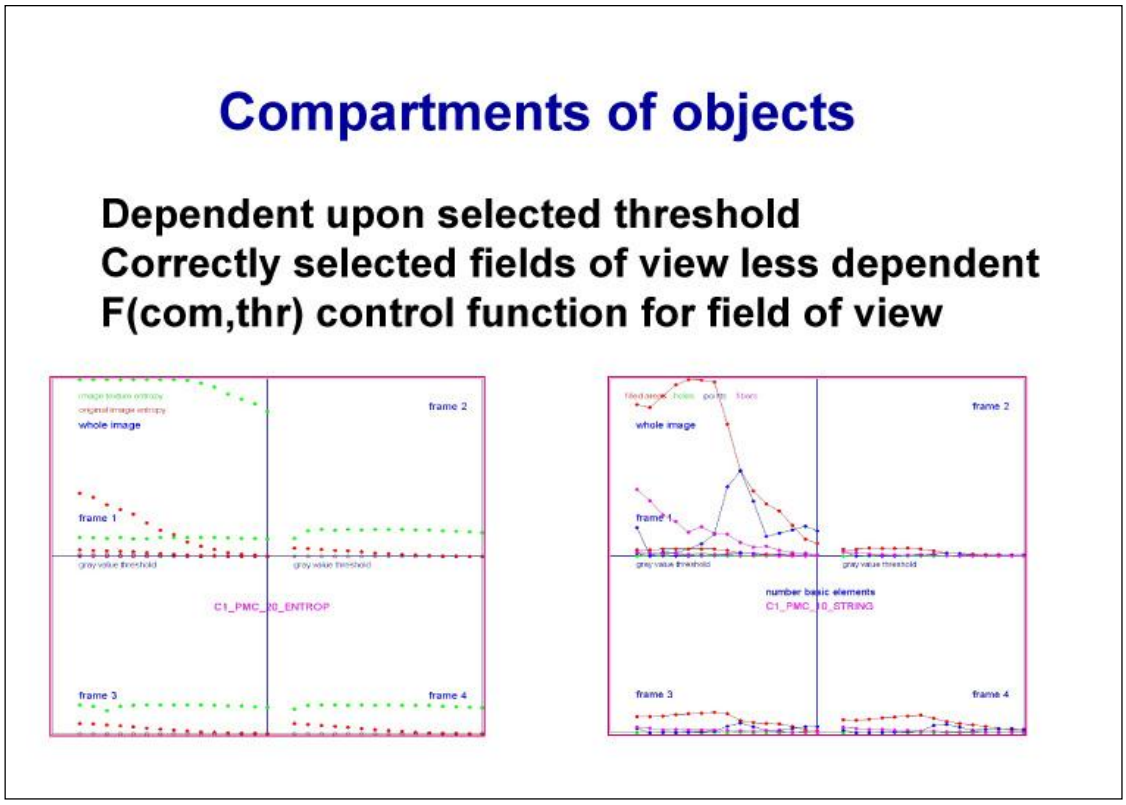

Figure 6: Example of correctly defined ROI. The gray value adjusted entropies displays with negligible variations in comparison to the whole image

Applications

Trials to quantify image objects and structures range back the beginning of microscopy. Natural scientists are not satisfied by "impressions" and try to classify features of objects and their functions, thus to measure. Comparison with a microscopic scale was used by Robert Remak (1815 – 1865) and Rudolf Virchow (1821 – 1902) in the 1850th [32, 33]. Stereology was developed and applied when projections of microscopic images have been available. Measurements on digitized images have been contemporary reported when digital consumer cameras turned up in the market [34]. Measurements focusing on objects displayed with close relationship to diagnosis, especially to cancer cell types, post surgical diseases stages (pTNM), and survival [34]. Structures were less frequently analyzed, despite they revealed a close association with the underlying disease and the clinical outcome too [35]. A close association of parameters derived from texture analysis with the patient's prognosis has been reported too [36, 37].

IHC and FISH promoted digital pathology to a large extent. Exploring cellular feedbacks that control proliferation and apoptosis membrane bound receptors and their associated enzymes have been detected. Their binding capacities and normal/abnormal function can be visualized by IHC, ISH, and similar techniques that display the exact position of the analyzed biological unit [27]. They are important targets of image measurements and often mandatory to optimize cancer treatment, for example in progressive breast or lung cancer [38-40]. Although a semi-quantitative judgment of the pathologist is sufficient for the steering of clinical treatment,

the diagnostic pathology journal DIAGNOSTIC PATHOLOGY

Kayser et. al. diagnostic pathology 2015, 1:14
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:14

quantitative measurements are more reproducible and might, in addition, in 10 – 20% of the cases replace additional expensive molecularbiological investigations [41-43].

Texture analysis, which should not be mistaken for structure analysis is of similar if not even more significant association with the clinical outcome, especially in lung, breast carcinomas, and in the development of metastases [38-40]. It has not been applied in predictive diagnosis to our knowledge.

Lu, Prewitt and Sanfeliu have already demonstrated that the evaluation of image primitives can be of useful clinical application (in differentiating muscular diseases) in the 1980th [44-47] First results indicate that they might serve for understanding thermodynamic properties in fetal development of chicken kidneys.

Discussion and conclusions

Without any doubt digital pathology is on its way to become implemented in tissue-based diagnosis and practice [34]. Several big companies invest in this technology, and certainly they have the power to get their investment reimbursed. One main column of digital pathology is the acquisition of image information, and to transfer this information into clinical practice [10, 11, 16]. The transfer tool is commonly called diagnosis. A diagnosis finally results in a digital clinical decision: For example: to operate yes<> no, to apply herceptin yes<> no, to stop tuberculosis treatment yes<> no, etc. At present, the pathologist reports his/her diagnosis to the clinician, whose final treatment remarkably depends on the pathologist's diagnosis.

Image information influences the pathologist's diagnosis via the pathologist's knowledge: It might be clear and "easy to see"; however, it is useless if the pathologist does not recognize it or cannot state the corresponding diagnosis.

Commonly, investigations on the pathologists' diagnostic algorithms start from the view point of the pathologist. Objects are predefined, and the computer is asked to detect them reproducible and accurate. Our proposed approach differs: The computer is asked to analyze the image texture, combine detectable agglutinations to primitives, compute for potential symmetries and associate the findings with the pathologists knowledge (or underlying disease). The interpretation of the computer information is the final step of the algorithm, and not the first one.

A digitized image such as VS should display the same objects, structures, etc as the original image viewed from the glass slide under a microscope. However, it is different, as all virtual worlds differ from the real environment. Is this difference a disadvantage for the diagnostic procedure?

The pathologist usually views different areas of a glass slide at different magnifications. This procedure can be translated in an algorithm that possesses two functions, a) sampling b) information acquisition [10, 11, 19]. Both are applied separately in VS when searching for ROI and segmentation of objects, etc. A similar separation of a "combined function" into two separated functions can be observed in the flight of birds compared to the flight of an airplane.

Diagnostics on virtual slides has been reported to last longer than using a conventional transmission light microscope [1, 7, 10]. This is mainly due to the long lasting sampling time and the non optimum screen size and image display [10, 14]. Sampling or the calculation of ROIs can be performed prior to and independent from the diagnosis evaluation. This separation results

Kayser et. al. diagnostic pathology 2015, 1:14
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:14

in a remarkable reduction of the "digital pathologist time". It is in most cases shorter than working with a conventional microscope [10, 14].

Crude judgment of objects and structures does not necessary require high image quality. It does not matter to judge the size of a nucleus at low magnification; however reproducible measurement of nuclear sizes mandatory need optimum image quality and adjusted image (object) pixel numbers [4, 6, 9, 14, 22, 26, 48]. Standardized image quality, potential corrections, and failure notifications are prerequisites for any image information acquisition.

It is too early to give a final statement about the feasibility of object-structure image information acquisition in comparison to image primitives or texture analysis. Some issues should be mentioned here:

1. "Classic" object information algorithms require appropriate and reproducible segmentation algorithms. There are several approaches to provide nearly errorless object detection [10]. The most frequently applied is the Otsu algorithm, which is dependent upon the gray value distribution [49-51]. Investigations that investigate in different objects (and their features) and the most applicable segmentation algorithm in histopathology are still missing to our knowledge.
2. "Classic" object information acquisition requires prior to its application the introduction of external knowledge, for example the definition of "the object". Therefore, for any different approach (or image content) another algorithm has to be developed.
3. External image information can be used for a more detailed and reproducible information acquisition, for example to more precisely segment an object.
4. "Classic information acquisition" induces images that are sued in the pathologist's work, and can be easily judged for correctness and completeness.
5. "Pixel derived" information can be transferred either into the "object world" by synthesis of image primitives to objects and structures, or to the final aim, the diagnosis directly by passing the 2object world".
6. "Pixel derived information" algorithms can work on images which are of less image quality. They can be derived from images that underwent several locally dependent and independent image transformations.
7. Both "pixel derived" and "object derived" information acquisition algorithms can be combined and be used for diagnosis decision algorithms contemporary [49-51].

In perspective, both algorithms can be implemented in serial and hierarchic assessment. Texture evaluation and measurement are suitable to work on images that are not classified and of unknown origin at the beginning, and that are suitable for further evaluation by object information driven algorithms. Thinking about these proposals, a new door will open into a world that permits insight into information "outside of that what we do see", in expansion of Goethe's statement: you do only see what you do know" [15].

Acknowledgement

Kayser et. al. diagnostic pathology 2015, 1:14
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:14

The financial support of the Verein zur Förderung des biologisch technologischen Fortschritts in der Medizin e.V. is gratefully acknowledged.

References

1. Racoceanu, D. and P. Belhomme, *Breakthrough technologies in digital pathology.* Comput Med Imaging Graph. **42**: p. 1.
2. Wright, A.I., H.I. Grabsch, and D.E. Treanor, *RandomSpot: A web-based tool for systematic random sampling of virtual slides.* J Pathol Inform. **6**: p. 8.
3. Lundstrom, C., et al., *Summary of 2(nd) Nordic symposium on digital pathology.* J Pathol Inform. **6**: p. 5.
4. Al-Janabi, S., et al., *Whole slide images for primary diagnostics of urinary system pathology: a feasibility study.* J Renal Inj Prev. **3**(4): p. 91-6.
5. Fine, J.L., *21(st) century workflow: A proposal.* J Pathol Inform. **5**: p. 44.
6. Greenbaum, A., et al., *Wide-field computational imaging of pathology slides using lens-free on-chip microscopy.* Sci Transl Med. **6**(267): p. 267ra175.
7. Kumar, A., et al., *Automated analysis of immunohistochemistry images identifies candidate location biomarkers for cancers.* Proc Natl Acad Sci U S A. **111**(51): p. 18249-54.
8. Amselgruber, W.M., et al., *Expression of galactosyltransferase in prostatic tumors.* Nutrition, 1995. **11**(5 Suppl): p. 638-42.
9. Kobara, H., et al., *Analysis of the amount of tissue sample necessary for mitotic count and Ki-67 index in gastrointestinal stromal tumor sampling.* Oncol Rep. **33**(1): p. 215-22.
10. Kayser, K., B. Molnar, and R.S. Weinstein, *Virtual Microscopy Fundamentals - Applications - Perspectives of Electronic Tissue - based Diagnosis*. 2006, Berlin: VSV Interdisciplinary Medical Publishing.
11. Kayser, K., et al., *Digitized pathology: theory and experiences in automated tissue-based virtual diagnosis.* Rom J Morphol Embryol, 2006. **47**(1): p. 21-8.
12. Kayser, K. and H. Hoffgen, *Pattern recognition in histopathology by orders of textures.* Med Inform (Lond), 1984. **9**(1): p. 55-9.
13. Kayser, K. and W. Schlegel, *Pattern recognition in histo-pathology: basic considerations.* Methods Inf Med, 1982. **21**(1): p. 15-22.
14. Kayser, G., et al., *Towards an automated morphological classification of histological images of common lung carcinomas.* Elec J Pathol Histol, 2002. **8**: p. 022-03.
15. von Goethe, J.W., *Zur Farbenlehre*. 1810, Tübingen: Cotta.

Kayser et. al. diagnostic pathology 2015, 1:14
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:14

16. Kayser, K., et al., AI (artificial intelligence) in histopathology--from image analysis to automated diagnosis. Folia Histochem Cytobiol, 2009. 47(3): p. 355-61.

17. Corredor, G., E. Romero, and M. Iregui, *An adaptable navigation strategy for Virtual Microscopy from mobile platforms.* J Biomed Inform.

18. Moles Lopez, X., et al., *An automated blur detection method for histological whole slide imaging.* PLoS One. **8**(12): p. e82710.

19. Kayser, K., et al., *Theory of sampling and its application in tissue based diagnosis.* Diagn Pathol, 2009. **4**: p. 6.

20. Gray, A., et al., *Quantification of histochemical stains using whole slide imaging: development of a method and demonstration of its usefulness in laboratory quality control.* J Clin Pathol. **68**(3): p. 192-9.

21. Khalbuss, W.E., J. Cuda, and I.C. Cucoranu, *Screening and dotting virtual slides: A new challenge for cytotechnologists.* Cytojournal. **10**: p. 22.

22. Keller, K.K., et al., *Improving efficiency in stereology: a study applying the proportionator and the autodisector on virtual slides.* J Microsc. **251**(1): p. 68-76.

23. Garcia-Rojo, M., L. Goncalves, and B. Blobel, The COST Action IC0604 "Telepathology Network in Europe" (EURO-TELEPATH). Stud Health Technol Inform. 179: p. 3-12.

24. Oger, M., et al., *Automated region of interest retrieval and classification using spectral analysis.* Diagn Pathol, 2008. **3 Suppl 1**: p. S17.

25. Li, X. and K. Plataniotis, *A Complete Color Normalization Approach To Histo-pathology Images Using Color Cues Computed From Saturation-Weighted Statistics.* IEEE Trans Biomed Eng.

26. Voss, K. and H. Süsse, *Praktische Bildverarbeitung*. 1991: München, Wien: Carl Hanser Verlag.

27. Kayser, K. and H.J. Gabius, *The application of thermodynamic principles to histochemical and morphometric tissue research: principles and practical outline with focus on the glycosciences.* Cell Tissue Res, 1999. **296**(3): p. 443-55.

28. Kayser, K., et al., *Glyco- and immunohistochemical refinement of the differential diagnosis between mesothelioma and metastatic carcinoma and survival analysis of patients.* J Pathol, 2001. **193**(2): p. 175-80.

29. Kayser, K., et al., *Application of computer-assisted morphometry to the analysis of prenatal development of human lung.* Anat Histol Embryol, 1997. **26**(2): p. 135-9.

30. Kayser, K., et al., *Carcinoid tumors of the lung: immuno- and ligandohistochemistry, analysis of integrated optical density, syntactic structure analysis, clinical data, and prognosis of patients treated surgically.* J Surg Oncol, 1996. **63**(2): p. 99-106.

Kayser et. al. diagnostic pathology 2015, 1:14
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:14

31. Kayser, K., et al., *Association of prognosis in surgically treated lung cancer patients with cytometric, histometric and ligand histochemical properties: with an emphasis on structural entropy.* Anal Quant Cytol Histol, 1998. **20**(4): p. 313-20.

32. Virchow, R., *Cellular Pathologie.* Virchows. Arch. Path. Anat., 1855. **8**: p. 3-39.

33. Remak, R., *Ein Beitrag zur Entwicklungsgeschichte der krebshaften Geschwuelste.* Dtsch. Klin., 1854. **6**: p. 170-175.

34. Park, S., et al., *The history of pathology informatics: A global perspective.* J Pathol Inform. **4**: p. 7.

35. Kayser, K., et al., *New developments in digital pathology: from telepathology to virtual pathology laboratory.* Stud Health Technol Inform, 2004. **105**: p. 61-9.

36. Szoke, T., et al., *Prognostic significance of endogenous adhesion/growth-regulatory lectins in lung cancer.* Oncology, 2005. **69**(2): p. 167-74.

37. Szoke, T., et al., *Prognostic significance of microvascularization in cases of operated lung cancer.* Eur J Cardiothorac Surg, 2005. **27**(6): p. 1106-11.

38. Zink, S., et al., *Survival, disease-free interval, and associated tumor features in patients with colon/rectal carcinomas and their resected intra-pulmonary metastases.* Eur J Cardiothorac Surg, 2001. **19**(6): p. 908-13.

39. Werle, B., et al., Cathepsin B in infiltrated lymph nodes is of prognostic significance for patients with nonsmall cell lung carcinoma. Cancer, 2000. 89(11): p. 2282-91.

40. Kayser, K., et al., *Pulmonary metastases of breast carcinomas: ligandohistochemical, nuclear, and structural analysis of primary and metastatic tumors with emphasis on period of occurrence of metastases and survival.* J Surg Oncol, 1998. **69**(3): p. 137-46.

41. Schmitt, F., *HER2+ breast cancer: how to evaluate?* Adv Ther, 2009. **26 Suppl 1**: p. S1-8.

42. Bilous, M., et al., *Predicting the HER2 status of breast cancer from basic histopathology data: an analysis of 1500 breast cancers as part of the HER2000 International Study.* Breast, 2003. **12**(2): p. 92-8.

43. Aitken, S.J., et al., *Quantitative analysis of changes in ER, PR and HER2 expression in primary breast cancer and paired nodal metastases.* Ann Oncol. **21**(6): p. 1254-61.

44. Sanfeliu, A., K.S. Fu, and J.M.S. Prewitt. *An application ofa distance measure between graphs to the analysis of muscle tissue patterns.* in *Workshop on Structural and Syntactic Pattern recognition*. 1981. Saratoga Springs, New Yorck.

45. Prewitt, J.M.S. and S.C. Wu. *An application of pattern recognition to epithelial tissues*. in *Computer Applications in Medical Care*. 1978: IEEE Computer Society.

46. Lu, S.Y. and K.S. Fu, *A syntactic approach to texture analysis.* Computer Garphics Image Processing, 1978. **7**: p. 303 - 330.

Kayser et. al. diagnostic pathology 2015, 1:14
ISSN 2364-4893
DOI: http://dx.doi.org/10.17629/www.diagnosticpathology.eu-2015-1:14

47. Lu, S.Y., *A tree-to-tree distance and its application to cluster analysis.* IEEE Trans Pattern Analysis and Maschine Intelligence, 1979. **PAMI-1**: p. 219-234.

48. Helps, S.C., et al., *Automatic nonsubjective estimation of antigen content visualized by immunohistochemistry using color deconvolution.* Appl Immunohistochem Mol Morphol. **20**(1): p. 82-90.

49. Kayser, K., et al., *Texture- and object-related automated information analysis in histological still images of various organs.* Anal Quant Cytol Histol, 2008. **30**(6): p. 323-35.

50. Walkowski, S. and J. Szymas, *Histopathologic patterns of nervous system tumors based on computer vision methods and whole slide imaging (WSI).* Anal Cell Pathol (Amst). **35**(2): p. 117-22.

51. Wang, F., et al., *A data model and database for high-resolution pathology analytical image informatics.* J Pathol Inform. **2**: p. 32.