 Cytometry Part A 69A:588–591 (2006)

Extracting Quantitative Information from Tissue—An Industrial Perspective

Peter Van Osta*
MAIA SCIENTIFIC, Cipalstraat 3, B-2440 Geel, Belgium

Received 12 January 2005; Revision Received 15 August 2005; Accepted 18 November 2005

The focus of this article is to provide an overview of the current technologies for the pharmaceutical and biotech industry. Disease processes express themselves in the functional and structural disturbance of cellular systems. Cells and their metabolites constitute the building blocks of tissues and entire organisms. Studying the spatial and temporal phenotype of disease processes in tissues at the cellular level reveals a multitude of information about the progress and status of a disease. Detailed exploration of tissues by slide-based cytometry is an important source of information about disease processes. Technological and analytical advances allow us to shed a new light on tissues and to come to a better understanding of the complexity of disease processes. Dealing with complex multidimensional datasets from tissue samples requires an advanced approach to image processing and data management. The increase in computing power and the continuing research into imaging algorithms allow us to improve the exploration of the data content of tissues. © 2006 International Society for Analytical Cytology

Key terms: image-based cytometry; microscopy; image analysis; quantification; tissue

Studying a disease process is possible at several levels of biological integration, ranging from individual genes up to the entire organism. Each level of biological integration will reveal part of the entire dynamical and structural process underlying a disease. An organism, however, exists by the grace of its cells and, as such, all processes in health and diseases have to express themselves through the structure and function of cells and the tissues they contain.

Tissues and the microscopic exploration of biological process at the level of tissue samples are already a long and well-established source of information. Studying cells in the context of a tissue allows for getting a detailed view on how cells change in relation to the tissue they have built and maintain.

In recent years we have witnessed an increase of our understanding of biological process at the cellular and tissue level. Advances in microscopical techniques and the advent of digital microscopy allow for a shift from a qualitative to a quantitative and objective understanding of cellular disease processes (1). This article aims at providing an entry into quantitative content extraction from tissues (cells), for the pharmaceutical and biotech industry.

IMAGING TECHNOLOGY

To study biological process in cells and tissues, we need technologies to detect structures of interest (staining, contrast techniques) and imaging techniques to capture an image of the sample in order to detect and quantify the biological process. Staining techniques allow us to distinguish structures and molecules of interest from their surroundings. Slide-based cytometry allows for (re-)staining multiple markers and exploring complex patterns of molecular phenotypes (2). Tissue samples can be studied individually or organized in arrays to allow for the simultaneous study of multiple structural and functional expression patterns. Analyzing a large number of tissues for candidate gene expression is now greatly facilitated by using tissue microarray (TMA) technology (3–5). Digital microscopy allows for correlating the patterns of gene expression to their spatial and temporal expression patterns in tissues (6).

Laser scanning and wide-field microscopes allow for studying molecular localization of proteins and their dynamics in cells and tissues in great detail (7–9). Confocal and multiphoton microscopy allow for a detailed exploration of cells in 3D and beyond (10). Multiphoton microscopy allows for exploring the deep structure of tissues (11). Modern auto focus algorithms allow for fast and robust autofocusing (12).

The resolving power of optical microscopy beyond the diffraction barrier is a new and interesting development, which will lead to the so-called super-resolving fluorescence microscopy (13). New microscopy techniques,

*Correspondence to: Peter Van Osta, De Zwaantjes 7, 2390 Malle, Belgium.
E-mail: pvosta@cs.com
Published online 5 May 2006 in Wiley InterScience (www.interscience.wiley.com).
DOI: 10.1002/cyto.a.20286

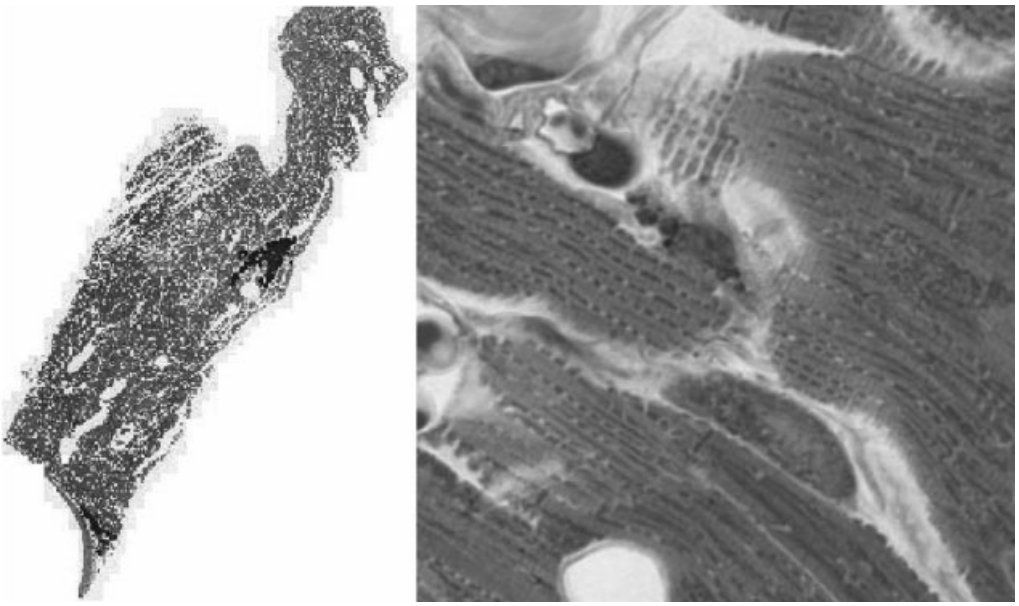

FIG. 1. Overview on the left and one individual image on the right. Toluidin blue stained rabbit heart tissue, Epon 2 μ semithin slice taken at ×63 (1.4 N.A.) consisting of 1,300 individual 512 × 512 pixel tiled images, taken with an automatic tissue-edge detection system.

such as standing wave microscopy, 4Pi confocal microscopy, I^5M, and structured illumination, are breaking the diffraction barrier and allow for improving the resolving power of optical microscopy (14,15). We are now heading toward fluorescence nanoscopy, which will improve spatial resolution far below the physical diffraction limits of 150 nm in the focal plane (XY) and 500 nm along the optical axis (Z) (16,17).

The high-order and the complex structural and functional organization of cell and tissues require capturing the multiparametric molecular morphology of its constituting components, such as the nucleus and other organelles (18,19).

The analogue image created by the microscope is converted by appropriate sampling into a digital representation, which is accessible to image detection and quantification techniques (20–22). The digitizing device either captures a broad part of the (visible) light, such as with a single-CCD camera or a multispectral image by using an array of detectors, a 3CCD color camera, or more channels for spectral imaging (23,24).

CREATING A DIGITAL REPRESENTATION

Let us take a more overall look at the process, to create a digital representation of cells and more particular tissues (Fig. 1). A tissue sample represents a continuum of biological information, which we convert into a digitized presentation for quantification. In order to quantify the physical properties of space and time of a tissue sample, we must be able to create an appropriate digital representation of these physical properties *in-silico*. This digital representation is then accessible to algorithms for content extraction. The content or objects of interest are then to be presented to a quantification engine that associates physical meaningful properties or features to the extracted objects (25). Finally, these features are analyzed to find structural and functional clusters, trends, periodicities, associations, and correlations.

Each imaging device enables us to create an image of a sample at a range on one side limited by its inner resolution (XYZ, spectral, temporal) and on the other side by its outer resolution (XYZ, spectral). Each spot captured by the imaging device represents a finite aperture of the imaging instrument in space (XYZ), spectrum (λ, wavelength) and time (t). The profile and extent of this single spot varies extensively, depending on the physical properties and capacity of the imaging device as such. Every image represents a cuboid extraction of the original sample put on or in the imaging device and represents five dimensions (XYZ, spectral, time). A traditional 2D image represents an optical depth defined by the optical characteristics of the microscope (N.A. of the objective) and, as such, constitutes the single slice version of a 3D image at its z-axis. A single 3D image exists for a stack of 2D slices, taken with ideally an equally spaced sampling in all planes, which is in reality not the case. The optical characteristics, such as the point-spread-function of the optics, distort the isospatial model of the sample. Each inner image element or voxel represents a sample point, while the outer extent of the 3D image represents a physical volume within the sample.

The use of electromagnetic waves and their interaction with the sample allow for the spectral and spatial resolution of the organization of the sample. The interaction and diffraction of the electromagnetic waves by the sample allow us to create an image of its spatial details and at a certain spectral distribution. Optical microscopy is only one way of using electromagnetic energy to interact with a specimen, to explore its inner structure.

The numerical representation of the created images in the memory of the computer is only a representation of the spatial, temporal, and spectral layout. The physical dimensions of each data point are only important when we want to back-propagate onto the physical extent of the sample afterward. In general, the size of the structures we explore by using digital microscopy, range from micrometers to millimeters in space (XYZ), the nanometer range in optical wavelength (λ), and milliseconds to several hours in time (t).

Each point represented by a number meaning intensity, spectrally distributed over the electromagnetic range captured by the optics and digitizing equipment. In conventional digital microscopy, the spectral characteristics of each physical point are represented by just one or three numbers (RGB), as such representing a wide spectral range, but there is no physical limitation on the spectral sampling. Software should be capable to deal with a wide range of data types and process them *in-silico* (26). A device attached to the software core should inform the system about its capacity to explore the spatial, spectral, and temporal range of a sample.

DEALING WITH IMAGE DATA

In order to be able to store and retrieve the image data and to analyze them, we need algorithms and software libraries capable, with a wide range of image content layouts. The data for a computer are only a pointer to a memory address, with additional information about the size of the unit data (byte, integer, float, etc.) and the layout of the pattern of the data matrix (intensity, RGB, etc.). The core of a system for large-scale image processing only needs to deal with the size, type, and layout of the data, in order to transfer them to the appropriate software component, for object detection, quantification, and data analysis.

Most software libraries provide special solutions for a specific range of image data and do not allow for an easy exchange of image content, unless the data are continuously reformatted.

Several solutions for image data storage and retrieval are available. The image cytometry standard file format (ICS, v. 1 and 2), allows for a flexible approach to storing and retrieving multidimensional image data (27). The open microscopy environment has the goal to provide a flexible data model, a relational database, and an XML-encoded file standard to allow for the exchange of data (28).

As the high dimensional datasets created by a digital imaging device can become extremely large and surpass the capacity of a single computer, solutions for a flexible increase of the computing capacity are necessary. The necessary increase of computing power requires both a solution at the level of computation as well as an increase in the processing capacity (29,30). For the exchange of digital data in a distributed computing environment (DCE), several solutions are available, such as the DCE, common object request broker architecture, .NET, simple object access protocol, etc. (31–34). Depending on the scale and the range of images generated, these tools should allow us to manage the dataflow throughout a scaleable image processing system.

OBJECT DETECTION

The extraction of content from the digital image of the tissue slice requires the application of some form of object detection. The digital sampling requires paying attention to the quantitative presentation of analogue shapes into their digital counterpart (35,36). Algorithms based on the principle of geometric diffusion allow for the detection of objects, which is related to that of the human visual system (37–39). The same principles can be applied to the detection of color, by using a color model, which takes into account the spatial organization of colors in an image (40). Improvements and new developments of digital imaging algorithms will enable us to increase the potential for quantitative analysis of tissue samples (41,42).

DATA QUANTIFICATION

The next step in an automated image analysis system for tissue analysis is the quantification of the features of the detected objects or of entire images. Both object-related quantification and global-image quantification allows for analyzing either specific or global characteristics of the tissue-image. The size and dimensionality of the created high-dimensional feature space necessitates the development of new feature-extraction approaches.

DATA PATTERN ANALYSIS

Finding biologically relevant patterns in the tissue data is the final step to come to an understanding of the *in vivo* process. Statistical or nonstatistical methods or both are being used to find clusters, trends, periodicities, associations, and correlations between and within samples. From data quantification to image, understanding is still a challenging step for which several approaches are being explored, depending on the underlying sample and experiment (43–45). Both statistical and nonstatistical approaches to data analysis are being used (46). Data mining of high dimensional data sets is still a challenging endeavor (47,48).

CONCLUSION

In recent years, we have seen an enormous increase in the usage and potential of digital microscopy techniques for analysis tissue samples. Developments in (digital) microscopy, image processing, and analysis allow us to shed a new light on the biological process in tissue. Feature extraction and data analysis methods are being developed to handle the increase in large multidimensional datasets. However, the management of the data flow and the analysis of large volumes of images and digital data will require a dramatic increase in the capacity of our digital imaging systems.

For the technology to be applicable for industrial R&D, it is not enough to automate the imaging process as such, but the entire chain of image capturing and content extraction up to knowledge extraction needs to be automated and integrated into the R&D process.

LITERATURE CITED

1. Furness PN. The use of digital images in pathology. J Pathol 1997;183: 253–263.
2. Gerstner AOH, Trumpfheller C, Racz P, Osmancik P, Tenner-Racz K, Tarnok A. Quantitative histology by multicolor slide-based cytometry. Cytometry A 2004;59:210–219.
3. Kononen J, Bubendorf L, Kallioniemi A, Barlund M, Schraml P, Leighton S, Torhorst J, Mihatsch MJ, Sauter G, Kallioniemi OP. Tissue microarrays for high-throughput molecular profiling of tumor specimens. Nat Med 1998;4:844–847.
4. Simon R, Sauter G. Tissue microarrays for miniaturized high-throughput molecular profiling of tumors. Exp Hematol 2002;30:1365–1372.
5. Braunschweig T, Chung JY, Hewitt SM. Perspectives in tissue microarrays. Comb Chem High Throughput Screen 2004;7:575–585.
6. Kriete A, Anderson MK, Love B, Freund J, Caffrey JJ, Young MB, Sendera TJ, Magnuson SR, Braughler JM. Combined histomorphometric and gene-expression profiling applied to toxicology. Genome Biol 2003;4:R32.
7. Denk W, Strickler JH, Webb WW. Two-photon laser scanning fuorescence microscopy. Science 1990;248:73–76.
8. Andrews PD, Harper IS, Swedlow JR. To 5D and beyond: Quantitative fluorescence microscopy in the postgenomic era. Traffic 2002;3:29–36.
9. Tarnok A, Gerstner A. Clinical applications of laser scanning cytometry. Cytometry 2002;50:133–143.
10. Peti-Peterdi J, Bell PD. Confocal and two-photon microscopy. Methods Mol Med 2003;86:129–138.
11. Centonze VC, White JG. Multiphoton excitation provides optical sections from deeper within scattering specimens than confocal imaging. Biophys J 1998;75:2015–2024.
12. Geusebroek JM, Cornelissen F, Smeulders AWM, Geerts H. Robust autofocusing in microscopy. Cytometry 2000;39:1–9.
13. Iketaki Y, Watanabe T, Ishiuchi S, Sakai M, Omatsu T, Yamamoto K, Fujii M. Predicted spatial resolution of super-resolving fluorescence microscopy using two-color fluorescence dip spectroscopy. Appl Spectrosc 2003;57:1312–1316.
14. Gustafsson MG. Extended resolution fluorescence microscopy. Curr Opin Struct Biol 1999;9:627–634.
15. Egner A, Verrier S, Goroshkov A, Söling H-D, Hell SW. 4Pi-microscopy of the golgi apparatus in live mammalian cells. J Struct Biol 2004;147: 70–76.
16. Hell SW. Toward fluorescence nanoscopy. Nat Biotechnol 2003;21: 1347–1355.
17. Hell SW, Dyba M, Jakobs S. Concepts for nanoscale resolution in fluorescence microscopy. Curr Opin Neurobiol 2004;14:599–609.
18. Ecker RC, de Martin R, Steiner GE, Schmid JA. Application of spectral imaging microscopy in cytomics and fluorescence energy transfer (FRET) analysis. Cytometry A 2004;59:179–182.
19. Best CJ, Emmert-Buck MR. Molecular profiling of tissue samples using laser capture microdissection. Expert Rev Mol Diagn 2001;1: 53–60.
20. Young IT. Quantitative microscopy. IEEE Eng Med Biol Mag 1996;15: 59–54.
21. Young IT. Sampling density and quantitative microscopy. Anal Quant Cytol Histol 1988;10:269–275.
22. Price JH, Goodacre A, Hahn K, Hodgson L, Hunter EA, Krajewski S, Murphy RF, Rabinovich A, Reed JC, Heynen S. Advances in molecular labeling, high throughput imaging and machine intelligence portend powerful functional cellular biochemistry tools. J Cell Biochem (Suppl.) 2003;39:194–210.
23. Hiraoka Y, Sedat J, Agard D. The use of a charge-coupled device for quantitative optical microscopy of biological structures. Science 1987;238:36–41.
24. Haraguchi T, Shimi T, Koujin T, Hashiguchi N, Hiraoka Y. Spectral imaging fluorescence microscopy. Genes Cells 2002;7:881–887.
25. Van Osta P, Bart Vanherck B, Kris Ver Donck K, Luc Bols L, Geysen J. The M^5 framework for exploring the cytome. Plenary Session and Discussion: Human Cytome Project, Philadelphia, PA, USA, April 4–7, 2004. Focus On Microscopy (F.O.M.) Abstract p.66.
26. Van Osta P. Automated tiled multi-mode image acquisition and processing applied to pharmaceutical research. 13th ProRISC workshop on Circuits, Systems and Signal Processing, Veldhoven, The Netherlands, November 28–29, 2002.
27. Dean P, Mascio L, Ow D, Sudar D, Mullikin J. Proposed standard for image cytometry data files. Cytometry 1990;11:561–569.
28. Swedlow JR, Goldberg I, Brauner E, Sorger PK. Informatics and quantitative analysis in biological imaging. Science 2003;300:100–102.
29. Seinstra FJ, Koelma D, Geusebroek JM. A software architecture for user transparent parallel image processing. Parallel Comput 2002;28: 967–993.
30. Carrington WA, Lisin D. Cluster computing for digital microscopy. Microsc Res Tech 2004;64:204–213.
31. DCE. Available at www.opengroup.org/dce/
32. CORBA. Available at www.corba.org/
33. NET. Available at www.microsoft.com/net/
34. SOAP. Available at www.w3.org/TR/soap/
35. Dorst L, Smeulders AWM. Length estimators for digitized contours. Comput Vis Graph Image Process 1987;40:311–333.
36. Smeulders AWM, Dorst L. Measurement issues in morphometry. Anal Quant Cytol Histol 1985;7:242–249.
37. Koenderink JJ. The structure of images. Biol Cybern 1984;50:363–370.
38. ter Haar Romeny BM, editor. Geometry-Driven Diffusion in Computer Vision. Dordrecht: Kluwer; 1994.
39. Van Osta P, Geusebroek JM, Ver Donck K, Bols L, Geysen J, ter Haar Romeny BM. The principles of scale space applied to structure and colour in light microscopy. Proc Royal Microsc Soc 2002;37:161–166.
40. Geusebroek JM, van den Boomgaard R, Smeulders AWM, Gevers T. Color constancy from physical principles. Pattern Recogn Lett 2003;24: 1653–1662.
41. Nattkemper TW. Automatic segmentation of digital micrographs: A survey. Medinfo 2004;2004:847–854.
42. Wurflinger T, Stockhausen J, Meyer-Ebrecht D, Bocking A. Robust automatic coregistration, segmentation, and classification of cell nuclei in multimodal cytopathological microscopic images. Comput Med Imaging Graph 2004;28:87–98.
43. Huang K, Murphy RF. From quantitative microscopy to automated image understanding. J Biomed Optics 2004;9:893–912.
44. Liu X, Minin V, Huang Y, Seligson DB, Horvath S, Statistical methods for analyzing tissue microarray data. J Biopharm Stat 2004;14:671–685.
45. McCullough B, Ying X, Monticello T, Bonnefoi M. Digital microscopy imaging and new approaches in toxicologic pathology. Toxicol Pathol 2004;32 (Suppl. 2):49–58.
46. Valet GK, Hoeffkes HG. Data pattern analysis for the individualised pretherapeutic identification of high-risk diffuse large B-cell lymphoma (DLBCL) patients by cytomics. Cytometry A 2004;59:232–236.
47. Wegman EJ. Visual data mining. Stat Med 2003;22:1383–1397.
48. Smyth P. Data mining: Data analysis on a grand scale? Stat Methods Med Res 2000;9:309–327.