ARTIFICIAL INTELLIGENCE IN BIOLOGY

Jan Urban, Petr Císař, Aliaksandr Pautsina, Jindřich Soukup, Antonín Bárta

Institute of Complex Systems, South Bohemian Research Center of Aquaculture and Biodiversity of Hydrocenoses, Faculty of Fisheries and Protection of Waters, University of South Bohemia, Zámek 136, Nové Hrady 37333, Czech Republic.

Abstract

The cybernetics was defined as a description of control and communication in living organisms and machines, by Norbert Wiener in 1948. Unfortunately, the part of living organisms is often underestimated. Recently, the initiative „A New Biology for the 21st Century“ of the US National Research Council of the National Academies, announced a goal of re-integration of the many sub-discipline of biology, and the integration into biology of physicists, chemists, computer scientists, engineers, and mathematicians to permit deeper understanding of biological systems. The similar aspect is expected to be a part of the next European framework Horizon 2020. Contemporary situation has to deal with two complementary issues: 1) The system theory and the artificial intelligence already produced plenty of theories, methods, and algorithms for processing and analysis of the digital (sampled and quantizied) signals, including images, to perform generous amount of possible results for given tasks. Various methods were gradually conditioned properly in specific or general way. 2) On the other hand, biology (biochemistry, biophysics, systems biology) is able to generate troubling problems, which are mathematically analogous to the problems already solved in the other scientific fields. Thus, the interdisciplinary collaboration has a possibility to increase an impact of the joint solution.

Introduction

Institute of complex systems is a part of the Faculty of Fisheries and Protection of Waters, University of South Bohemia. The institute consist of laboratories of Tissue cultures; Laser, Microscopy, Condensed phase and Material Engineering; Macromolecular Structure and Dynamics; and Applied systems biology. In the year 2014 will be created new laboratory for signal processing and analysis. The laboratory developed many software solutions in Matlab environment to help biologists with analysis of microscopical images, and other signals.

FishGui

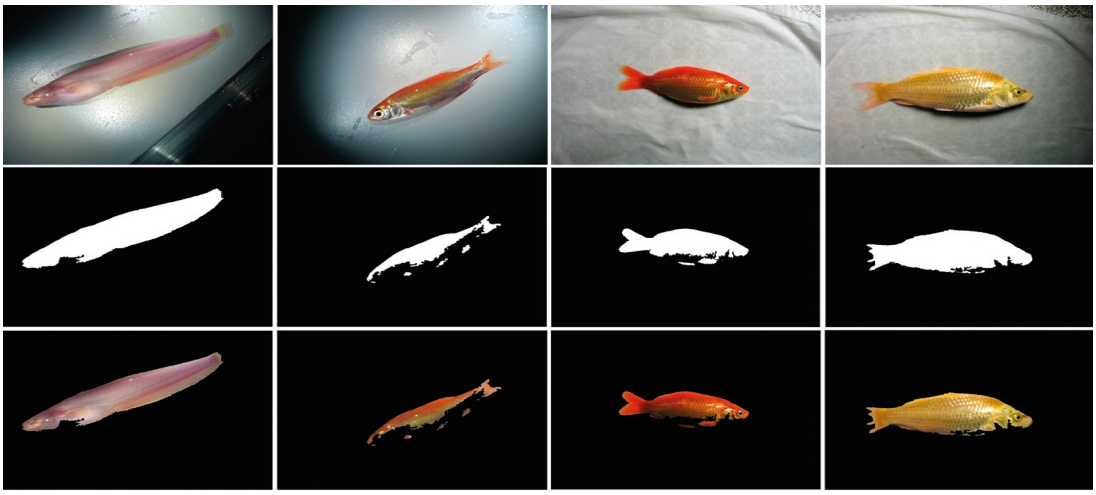

Figure 1: Segmentation of ornamental fish in FishGui aplication.

The colour hue and saturation of the fish skin are important factors in fish breeding, reflecting the momentary physiological status of the fish. Algorithms evaluating fish colour were created in the MATLAB environment and compiled in the stand-alone application, Expertomica Fishgui. Conversion from the Red-Green-Blue (RGB) colour space to chromatic colours was carried out to reduce the colour space. The threshold of skin chromatic colour was set on basis of the chromatic histograms using Otsu automatic bimodal segmentation. The average colour was calculated through all pixels of the selected area in the original image and through all images of the group. Colour conversion to the Hue-Saturation-Value (HSV) colour space was also carried out. Relative deviation between the saturation of samples in the control and the experimental groups provided a means of judgement. The dominant wavelength was also determined. [1,2,3,4]

Photosystems

Algorithm, which helps biologists to examine structure of very small objects, is described. These small objects are many times on the display limit of the best existent microscopes and the quality of the raw images is insufficient. In this case is advantageous to use a digital postprocessing of the images. The digital postprocessing is based on processes of huge amount of the low quality images which depict the same small object and in the specialized software can be estimated one picture of the object in much better quality. [5]

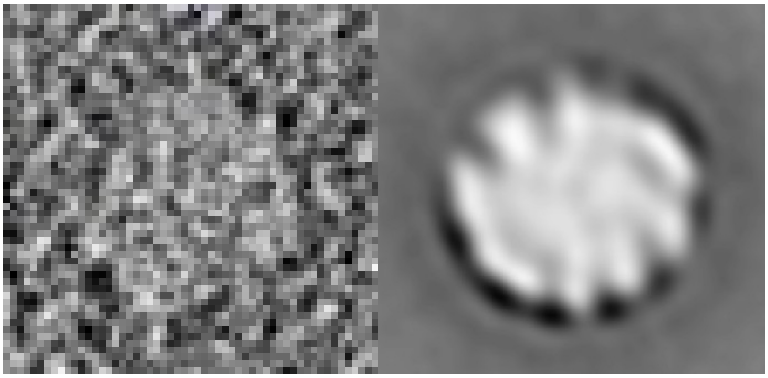

Figure 2: Left: photosystem I in raw source image. Right: postprocessed photosystem I from 2 467 particles.

Cells segmentation

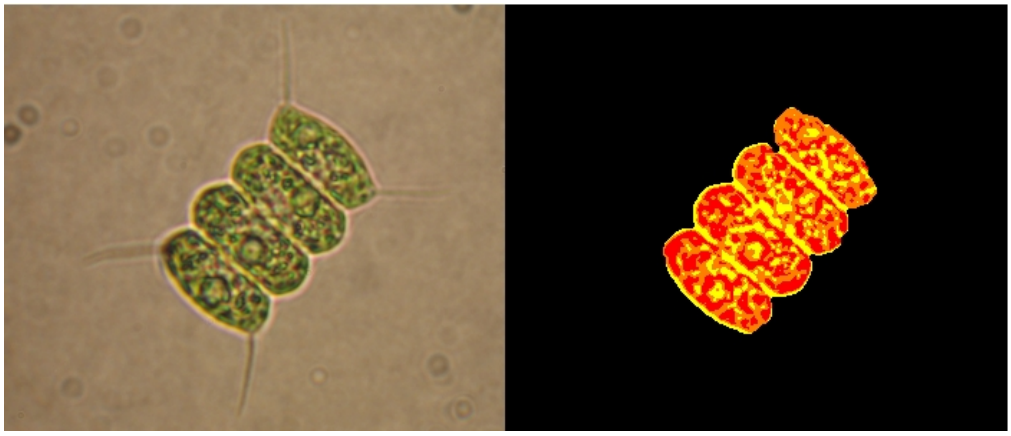

Figure 3: Illustration of image processing, inner structures of cells on input image (left) are classified into three classes (right).

Applying image analysis algorithms in biological field: all the modern measurement equipments produce a large amount data in digitally form including microscopes. Developing of automatic or semi-automatic tools e.g. in Matlab for various tasks could be helpful. The algorithm set described in this paper analyze a large sequence of microscope images of scenedesmus growth and evaluate quantitative features. Within developing this set, a novel segmentation algorithm which is based on artificial neural networks was proposed. This algorithm highly outperform the classical image processing approaches. [6, 7, 8,9]

HeLa Cells

The possibilities for segmentation of Helacyton gartleri cell line from their background as well as from each other in digital images from phase contrast time lapse microphotography are tested, combined and improved. Proper segmentation is one of the main issue of image evaulation and steps order differs from task to task, depending on input images. For reaching information about cell sphericity, several approaches are applied, including filtration, details emphasizing, and segmentation techniques. [10,11]

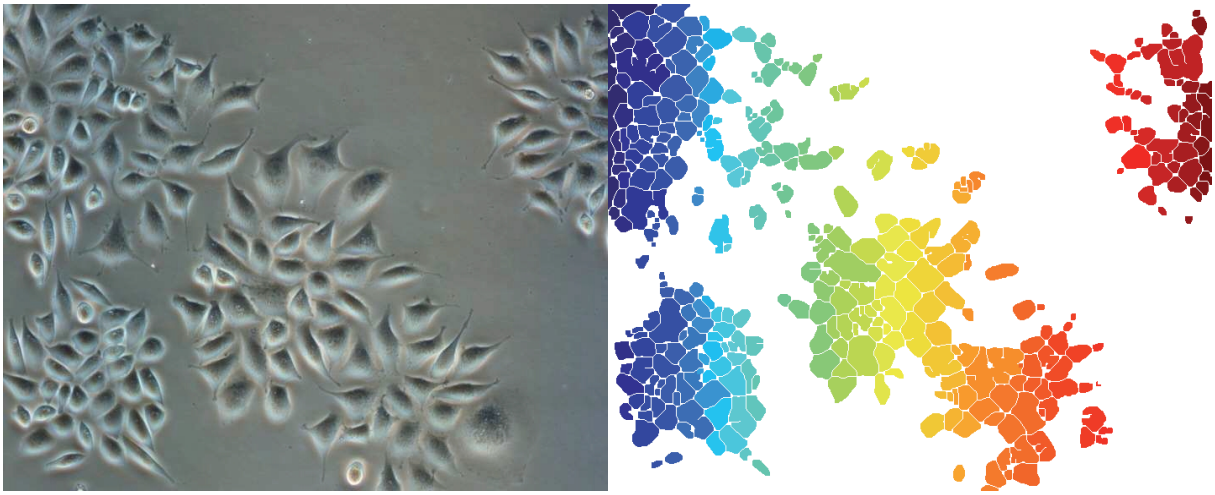

Figure 4: Automatic segmentation of HeLa cells in phase contrast microphotography.

Updopsi2dge

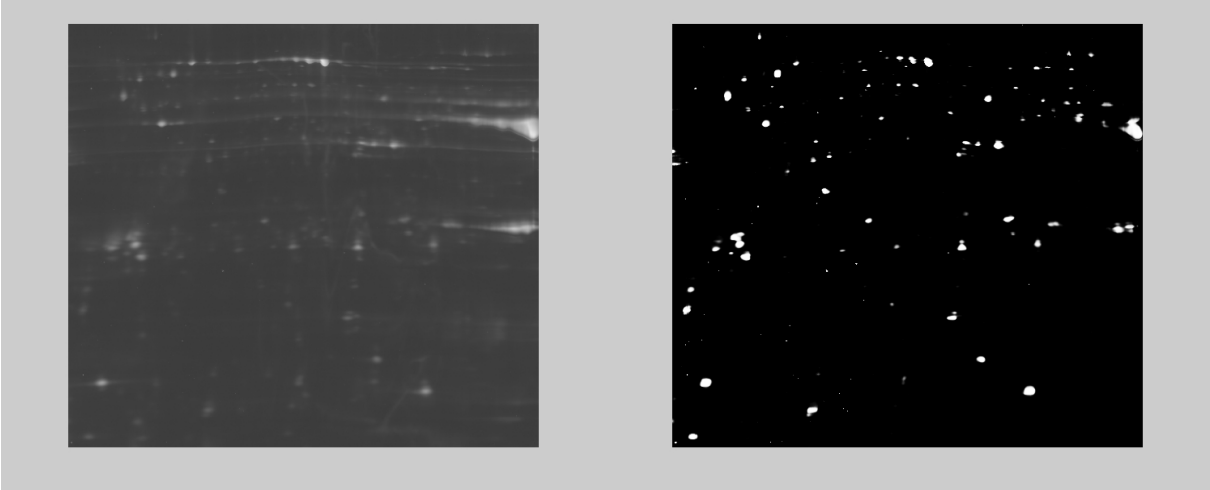

Figure 5: Filtration of 2D electrophoresis gel images based on probabilistic approach: Raw image on the left, filtration result on the right. Segmentation performed using Expertomica Updopsi2dge toolbox for Matlab.

The toolbox was originally developed for unsupervised probabilistic detection of protein spots in two-dimensional gel electrophoresis images as a toolbox for Matlab.2DE gel images are processed for background thresholding as automatic evaluation of each pixel probability to decide if it is a meaningful signal. The method is based on evaluation of noise signal behavior. Presence of useful signal can be detected as a violation in the noise behavior. Background in 2DE gel should be considered as a special case of noise. The fitted distribution depends only on the type of electrophoresis separation. [12]

LC-MS

The application was developed for analysis of datasets with high resolution from Liquid chromatography – mass spectrometry. It works with several ascii and xml datafiles. Native Matlab and specified cdf files are also supported. Application performs precison and resolution estimation, mass peaks centroidisation, unsupervised noise and mobile phase removal, iterative peak deconvolution and confidence factors evaluation. No information is lost during the processing. Program enables graphical user interface with basic plotings (Total Ion Current chromatograms and Mass spectra) of raw, processed and resulted data. [13-22]

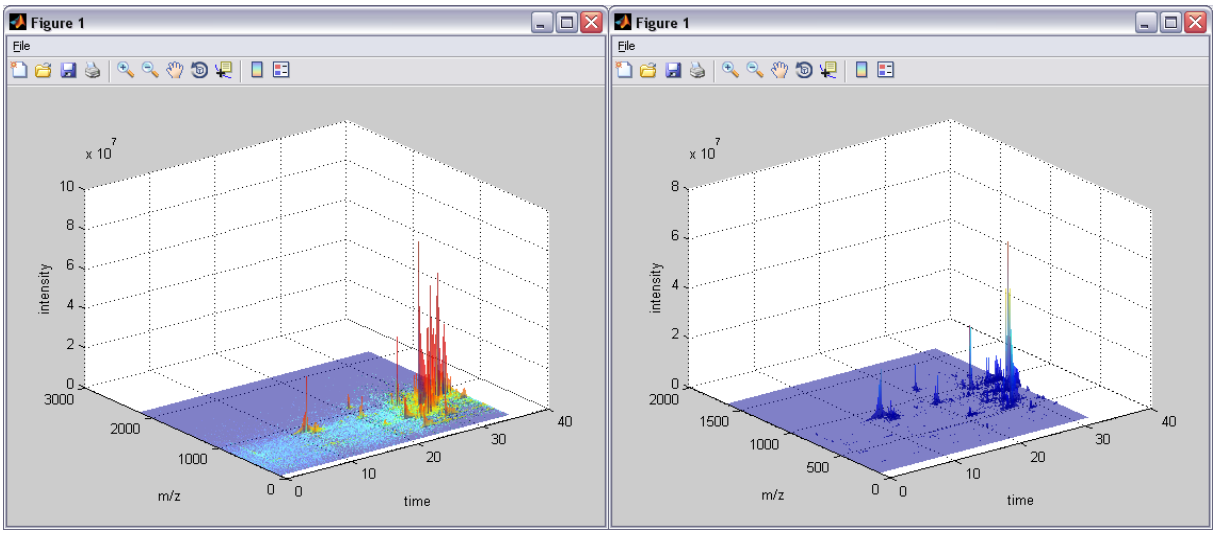

Figure 6: Filtration of LC-MS signal using probalisitc approach.

We present the structure of stochastic systems theory for formal description of the liquid chromatography - mass spectrometry (LC-MS) measurements. The states of the system are assigned by the chromatographic column elution in so-called retention time. As system variables are usually considered levels of analytes (metabolites, proteins, lipids) in extracts of cell compartments measured by MS detector. Thus, constituents are chemical individui detectable in the system, and components are sets of constituents which change in behavior dependently, e.g isotopologues, fragments, or adducts. In practical terms it means to determine measurable phenomenological variables in time change of the state. The examining of the behavior as state trajectory leads to probability distribution functions for deconvolution between the constituents. Therefore, related system based approach is introduced for the description of LC-MS measurement data. The abstract model is constructed according to the paradigms of the system theory. Therefore, denitions of attributes and their sets of variables are consistent and explicit for each mapped Cartesian products in the state space as well as for every data processing or analysis step.

Point information Gain

Novel method of image preprocessing based on Shannon's entropy was developed a specially for microscopy images captured in phase-contrast mode. But it can be used in many others applications. Entropy filtration is strictly depended on size of the neigbourhood and produce noisy or

blurry outputs. On the other hand, entropy thresholding is not sensitive for local changes. Our approach, the entropy contribution allows to deal with local changes without blurring and partial denoising. Used equation is very simple, but computational time for each pixel is time consuming. Therefore, we propose to using parallelisation on graphics cards. The algorithm should be used for other images with objects on more or less simple background. [23-34]

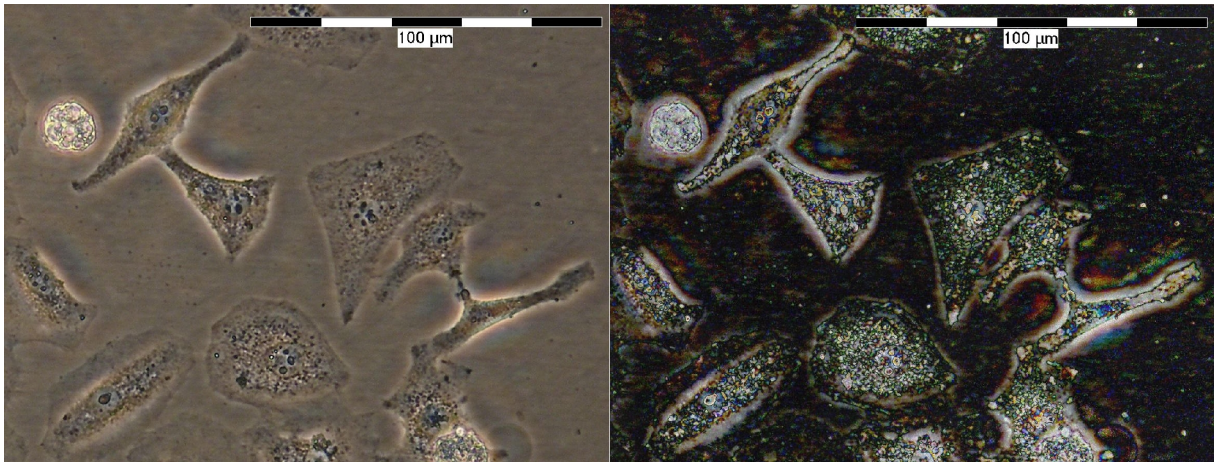

Figure 7: Detail of Point Information Gain of HeLa cells image. Left: original image. Right: PIG.

VioPin

Viopin is Graphical User Interface and Matlab compiled software application. The application was developed for analysis of image sets from fluorescent microscopy with red/pink autofluorescence of the cells, and blue/violet fluorescence of the nuclei. The application processes common image file formats (jpeg, png, bmp, tiff). Application performs noise filtration, cells to background segmentation, and nuclei to cells segmentation. Ratio of the cells to nuclei is computed individually for each image, as well as average ratio and standard deviation for the whole set. Program enables graphical user interface with basic functions as image view, zoom, and pan. Automatic export of the results is available. [35]

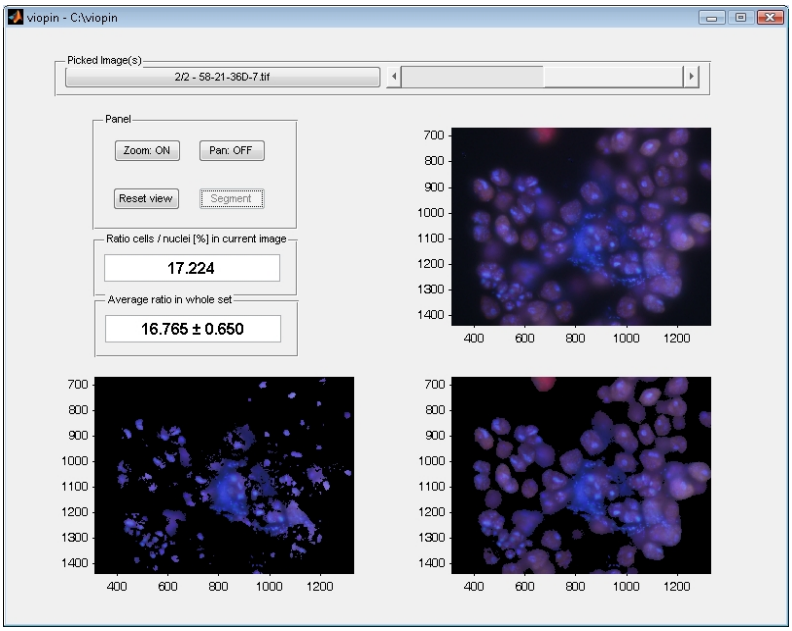

Figure 8: Main function of the software provides segmentation of all images in the directory. The results are then shown in remaining two figures, lower right for the segmentation cells / background, and lower left for the segmentation nuclei / cells.

Superresolution

Resolution in confocal fluorescence microscopy depends on the whole optical system: detector (chip), lenses (magnification, NA), dynamic range, Signal to Noise Ratio (SNR), gain, Point Spread Function (PSF). Theoretical limit for lenses at 550 nanometers: R = 0.20 μm. PSF is the impulse response of optical system. Impulse response is the system response to the Dirac unit input. In practice, it is not physically feasible to produce Dirac input, therefore it is evaluated as the first derivation of the transient response. Therefore, the integral of PSF is the transient response of the Microscope. PSF is in spatial domain, while Optical Transfer Function (OTF) is in frequency domain. OTF consist of Modulation TF and Phase TF (ISO 9334 standard). Diffraction pattern of a point light source has inner light circle,known as Airy diameter (1 Airy Unit). Pinhole on confocal fluorescence microscopy suppress the light from outside the focal plane. Open one brings brighter but blurred image, while closed pinhole produce darker, suppress diffused light, sharper. Subtractive imaging is done in frequency domain: (Fourier transformation):

$$Isub = Iclose - \gamma * Iopen \text{ [36].}$$

Intensity of open to be half of close. Subtraction is done in frequency domain.

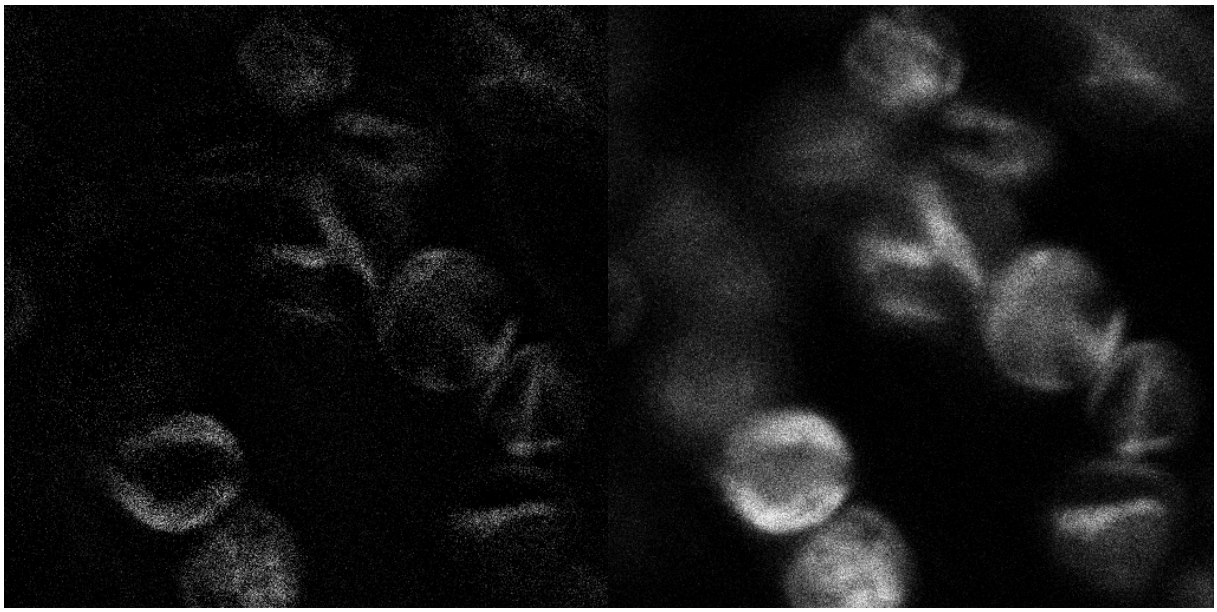

Figure 9: Subtractive imaging on the left, one Airy unit opened pinhole on the right. The blured parts, and contributions outside the focal plane were removed in subtractive image (left).

Conclusions

Biological data are troubling in interpretation. However, they are mathematically analogous to the problems already solved in the different fields. The processing and analysis of digital images, processing of signals for classification offers pre-implemented method for task designed solutions. Currently, the expansion of the collaboration between data producers and application developers is on the rise [37].

Acknowledgement

This work was supported and co-financed by the project Postdok JU CZ.1.07/2.3.00/30.0006; by the South Bohemian Research Center of Aquaculture and Biodiversity of Hydrocenoses CENAKVA CZ.1.05/2.1.00/01.0024; and by the South Bohemia University grant Selected complexity phenomena in condensed phase:experiment, development and theory GA JU 134/2013/Z (in part).

The authors thank to Zdeno Gardian, Vítězslav Březina, Iva Zaťková, Antonín Kouba, Dalibor Štys, Milada Víťová, Irene Lichtscheidl, Pavel Hrouzek, Pavel Souček, Tomáš Náhlík, Jan Vaněk, Jiří Soukup, Tomáš Levitner, Jiří Kopecký, Petr Kohout, Harald Martens, Štěpán Papáček, Nils Kristian Afseth.

References

[1] Urban, J., Zaťková, I., Vaněk, J. Comparison of the Fish Skin Colour, Technical Computing Prague, 2006, CZ.

[2] Zaťková I., Sergejevová M., Urban J., Vachta R., Štys D., Masojídek J., Carotenoid-enriched microalgal biomass as feed supplement for freshwater ornamentals: albinic form of wels catfish (Silurus glanis), Aquaculture Nutrition, doi: 10.1111/j.1365-2095.

[3] Urban J, Štys D, Sergejevová M, and Masojídek J, Expertomica Fishgui: Comparison of fish skin colour, Journal of Applied Ichthyology, Vol. 29 (1), p 172–180, doi: 10.1111/jai.12011

[4] Urban J., Sergejevová M., Slepičková I., Kouba A., Štys D., Masojídek J., Hodnocení změn vybarvení okrasných ryb, 2012. Edice Metodik (Technologická řada), FROV JU, Vodňany, č. 131, ISBN: 978-80-87437-55-1.

[5] Vaněk, J., Urban, J., Gardian, Z. Automatic Detection of Photosystems II in Electron Microscope Photographs, Technical Computing Prague, 2006, CZ.

[6] Urban J., Vaněk J., Štys D. Automatic analysis of Nostoc sp. growth, Materials structure in chemistry, biology, physics and technology (abstract), Nove Hrady, CZ, 2007.

[7] Vaněk, K., Urban, J., Březina, V. Study of Scenedesmus Alga Growth Via Image Analysis, Technical Computing Prague, 2007, CZ.

[8] Vaněk J., Urban J., Štys D., RIV/60076658:12640/09:00010083 - Algae and other microorganisms growth monitoring toolbox based on neural networks.

[9] Urban J., Březina V., Štys D., RIV/60076658:12640/09:00010085 - Expertomica Caamio.

[10] Urban J, Automatic image segmentation of HeLa cells in phase contrast microphotography, Lap LAMBERT Academic Publishing, 2012. ISBN-13: 978-3846585320.

[11] Urban, J., Thörnberg, B., Březina V., Štys, D. Evaluation of HeLa Cells Growth, Technical Computing Prague, 2007, CZ.

[12] Urban J., Štys D., RIV/60076658:12640/12:43882587, Expertomica Updopsi2dge toolbox, 2011.

[13] Urban J., Vaněk J., Soukup J., Štys D. Expertomica metabolite profiling: getting more information from LC-MS using the stochastic systems approach. Bioinformatics. 2009 Oct 15;25(20):2764-7.Vol. 25 no. 20 2009, pages 2764–2767 doi:10.1093/bioinformatics/b

[14] Urban J., Vaněk J., Štys D., Mass spectrometry: system based analysis, Computer Applications in Biotechnology, Volume 11 | Part 1, 10.3182/20100707-3-BE-2012.0040, ISBN: 978-3-902661-70-8.

[15] Hrouzek P., Tomek P., Lukešová A., Urban J., Voloshko L., Pushparaj B., Ventura S., Lukavský J., Štys D., Kopecký J.,Cytotoxicity and secondary metabolites production in terrestrial Nostoc strains, originating from different climatic/geographic regions and habitats: Is their cytotoxicity environmentally dependent?, Environmental Toxicology, doi: 10.1002/tox.20561

[16] Urban J, Hrouzek P, Štys D. Time-alignment in high performance liquidchromatography - mass spectrometry based on blankmeasurement, Lap LAMBERT Academic Publishing, 2012. ISBN-13: 978-3-659-13020-5

[17] Urban J, Vaněk J, Štys D. Current State of HPLC-MS Data Processing and Analysis in Proteomics and Metabolomics. Current Proteomics, Vol. 09, No. 2., pp.80-93 (14)

[18] Urban J, Vaněk J, Štys D. Unsupervised Adaptive Filter for Baseline Thresholding and Elimination in Liquid Chromatography - Mass Spectrometry via Approximation of the Standard Deviation of Baseline Distribution in Retention Time Domain. Acta Chromatographica.

[19] Urban J, AfsethNK, Štys D, Mass assignment, centroiding, and resolution – fundamental definitions and confusions in mass spectrometry, TrAC Trends in Analytical Chemistry

[20] Urban J, Vaněk J, Štys D. Unsupervised Adaptive Filter for Baseline Thresholding and Elimination in Liquid Chromatography - Mass Spectrometry via Approximation of the Standard Deviation of Baseline Distribution in Retention Time Domain. Acta Chromatographica. 257-273, Vol.25, N. 2, 2013 doi: 10.1556/AChrom.25.2013.2.4

[21] Urban J., Hrouzek P., Štys D., Martens H., Estimation of ion competition via correlated responsivity offset in linear ion trap mass spectrometry analysis: theory and practical use in the analysis of cyanobacterial hepatotoxin microcystin-LR in extracts of food additives. BioMed Research International (formerly titled Journal of Biomedicine and Biotechnology). Doi: 10.1155/2013/414631

[22] Urban J, Vaněk J, Štys D. Systems Theory in Liquid Chromatography - Mass Spectrometry, Lap LAMBERT Academic Publishing, 2012. ISBN-13: 978-3-659-29816-5

[23] Levitner T., Urban J., Vaněk J., Štys D. Sledování biologických interakcí z pohledu dostupných dat, Modelování a měření interakcí v mechanice, Nečtiny, CZ, 2009

[24] Urban, J., Vaněk J., Štys, D. Using Information Entropy for Camera Settings, Technical Computing Prague, 2008, CZ.

[25] Urban J., Vaněk J., Štys D. Entropy fluxes in cell culture and its parallel implementation on GPU, SIMONA 2009 (SImulace, MOdelování a Nejrůznější Aplikace), Liberec, CZ, ISBN 978-80-7372-543-3, p170-177.

[26] Urban J., Vaněk J., Štys D. Preprocessing of microscopy images via Shannon's entropy, In Proc. Of Pattern Recognition and Information Processing, pp.183-187, Minsk, Belarus, ISBN 978-985-476-704-8, 2009.

[27] Urban J., Vaněk J., Štys D. Preprocessing of microscopy images via Shannon's entropy on GPU, Computer Graphics, Computer Vision and Mathematics, ISBN 978-80-86943-90-9, p48-51.

[28] Štys D., Urban J., Vaněk J, Císař P,. Analysis of biological time-lapse microscopic experiment from the point of view of the information theory, 2010, Micron, S0968-4328(10)00026-0, doi:10.1016/j.micron.2010.01.012

[29] Štys D., Urban J., Levitner T. Decomposition of cellular system via causal relations, Materials Structure inChemistry, Biology, Physics and Technology, vol. 18, no. 1, 2011, ISSN 1211 – 5894.

[30] Štys D., Urban J., Císař P., Náhlík T., Pautsina A.,Vaněk J., The state trajectory of cell using Renyi entropy coeficients, Procedia Computer Science, Volume 7 (2011), Pages Pages 212-213. ISSN: 1877-0509.

[31] Štys, D ; Vaněk, J; Náhlík, T; Urban, J; Císař, P., The cell monolayer trajectory from the system state point of view ; 2011, Molecular Biosystems, 7/10, pp:2824-2833, DOI: 10.1039/c1mb05083d

[32] Náhlík T., Urban J., Císař P., Vaněk J., Štys D., Entropy based approximation to cell monolayer development, ISSN: 1680-0737, IFMBE Proceedings, 2012, Volume 37, Part 1, Part 6, 563-566, DOI: 10.1007/978-3-642-23508-5_146

[33] Urban J, Štys D, Could New Paradigms of System Theory Help with the Understanding of Molecular Interactions Inside a Cell? International Journal of Virology and Molecular Biology, Vol. 1 No. 3, 2012, pp. 23-27. doi: 10.5923/j.ijvmb.20120103.01.

[34] Štys D, Urban J., Náhlík T., Vaněk J., Císař P., Technické a teoretické aspekty mikroskopie živé buňky, Bulletin VÚRH Vodňany 48, 2012/3, 27–47, ISSN 0007-389X.

[35] Urban J., Software Expertomica VioPin.

[36] Rainer Heintzmanna, Vassilios Sarafisb, Paul Munroec, John Nailond, Quentin S. Hanleya Thomas M. Jovina. Resolution enhancement by subtraction of confocal signals taken at different pinhole sizes. 293–300, Micron 34 (2003).

[37] http://www.frov.jcu.cz/en/sks-frov-ju/skola-komkplexnich-systemu

Contact: urbanj@frov.jcu.cz