

## Chapter 4.6

# Image Processing and Machine Learning Techniques for the Segmentation of cDNA Microarray Images

**Nikolaos Giannakeas**

*University of Ioannina, Greece*

**Dimitrios I. Fotiadis**

*University of Ioannina, Greece*

### **ABSTRACT**

Microarray technology allows the comprehensive measurement of the expression level of many genes simultaneously on a common substrate. Typical applications of microarrays include the quantification of expression profiles of a system under different experimental conditions, or expression profile comparisons of two systems for one or more conditions. Microarray image analysis is a crucial step in the analysis of microarray data. In this chapter an extensive overview of the segmentation of the microarray image is presented.

DOI: 10.4018/978-1-60960-818-7.ch4.6

Methods already presented in the literature are classified into two main categories: methods which are based on image processing techniques and those which are based on Machine learning techniques. A novel classification-based application for the segmentation is also presented to demonstrate efficiency.

### **INTRODUCTION**

Several types of microarrays have been developed to address different biological processes: (i) cDNA microarrays (Eisen, 1999) are used for the monitoring of the gene expression levels to

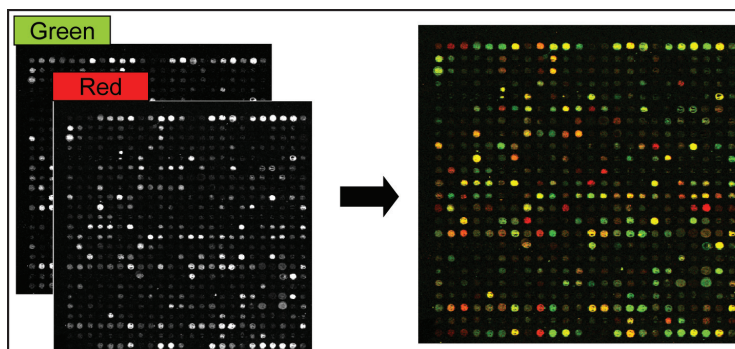
study the effects of certain treatments, diseases, and developmental stages on gene expression. As a result, microarray gene expression profiling can be used to identify disease genes by comparing gene expression in diseased and normal cells. (ii) Comparative genomic hybridization application assesses genome content in different cells or closely related organisms (Pollack et al., 1999). (iii) SNP detection arrays identify single nucleotide polymorphism among alleles within or between populations (Moran & Whitney, 2004). (iv) Finally, Chromatin immunoprecipitation (chIP) technologies determine protein binding site occupancy throughout the genome, employing ChIP-on-chip technology (Buck & Lieb, 2004).

The experiment of cDNA microarrays typically starts by taking two biological tissues and extracting their mRNA. The mRNA samples are reverse transcribed into complementary DNA (cDNA) and labelled with fluorescent dyes resulting in a fluorescence-tagged cDNA. The most common dyes for tagging cDNA are the red fluorescent dye Cy5 (emission from 630-660 nm) and the green-fluorescent dye Cy3 (emission from 510-550 nm). Next, the tagged cDNA copy, called the sample probe, is hybridized on a slide containing a grid or array of single-stranded cDNAs called probes. Probes are usually known genes of interest which were printed on a glass microscope slide by a robotic arrayer. According to the hybridization principles, a sample probe will only hybridize

with its complementary probe. The probe-sample hybridization process on a microarray typically occurs after several hours. All unhybridized sample probes are then washed off and the microarray is scanned twice, at different wavelengths corresponding to the different dyes used in the assay. The digital image scanner records the intensity level at each grid location producing two greyscale images. The intensity level is correlated with the absolute amount of RNA in the original sample, and thus, the expression level of the gene associated with this RNA.

Automated quantification of gene expression levels is realized analyzing the microarray images. Microarray images contain several blocks (or subgrids) which consist of a number of spots, placed in rows and columns (Figure 1). The level of intensity of each spot represents the amount of sample which is hybridized with the corresponding gene. The processing of microarray images (Schena et al., 1995) includes three stages: initially, spots and blocks are preliminarily located from the images (gridding). Second, using the available gridding information, each microarray spot is individually segmented into foreground and background. Finally, intensity extraction, calculates the foreground fluorescence intensity, which represents each gene expression level, and the background intensities. Ideally, the image analysis would be a rather trivial process, if all the spots had circular shape, similar size,

*Figure 1. A Block of a Typical Microarray Image*



11 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

[www.igi-global.com/chapter/image-processing-machine-learning-techniques/56176](http://www.igi-global.com/chapter/image-processing-machine-learning-techniques/56176)

## Related Content

---

### Prerequisites of Autonomous Software Operations in Home UbiHealth Described Formally with Denotational Mathematics

John Sarivougioukas, Aristides Vagelatos and Isaac Lagaris (2016). *International Journal of Software Science and Computational Intelligence* (pp. 16-31).

[www.irma-international.org/article/prerequisites-of-autonomous-software-operations-in-home-ubihealth-described-formally-with-denotational-mathematics/161710](http://www.irma-international.org/article/prerequisites-of-autonomous-software-operations-in-home-ubihealth-described-formally-with-denotational-mathematics/161710)

### A General Knowledge Representation Model for the Acquisition of Skills and Concepts

Carlos Ramirez and Benjamin Valdes (2012). *Breakthroughs in Software Science and Computational Intelligence* (pp. 28-45).

[www.irma-international.org/chapter/general-knowledge-representation-model-acquisition/64601](http://www.irma-international.org/chapter/general-knowledge-representation-model-acquisition/64601)

### Search of Web Service based on Weighted Association Rule

Lei Wang, Lingyu Xu, Yunlan Xue, Gaowei Zhang and Xiangfeng Luo (2015). *International Journal of Software Science and Computational Intelligence* (pp. 46-57).

[www.irma-international.org/article/search-of-web-service-based-on-weighted-association-rule/155158](http://www.irma-international.org/article/search-of-web-service-based-on-weighted-association-rule/155158)

### Improved SfM-Based Indoor Localization with Occlusion Removal

Yushi Li, George Baciu, Yu Han and Chenhui Li (2018). *International Journal of Software Science and Computational Intelligence* (pp. 24-40).

[www.irma-international.org/article/improved-sfm-based-indoor-localization-with-occlusion-removal/207743](http://www.irma-international.org/article/improved-sfm-based-indoor-localization-with-occlusion-removal/207743)

### Learning Manifolds: Design Analysis for Medical Applications

Diana Mateus, Christian Wachinger, Selen Atasoy, Loren Schwarz and Nassir Navab (2012). *Machine Learning in Computer-Aided Diagnosis: Medical Imaging Intelligence and Analysis* (pp. 374-402).

[www.irma-international.org/chapter/learning-manifolds-design-analysis-medical/62239](http://www.irma-international.org/chapter/learning-manifolds-design-analysis-medical/62239)