Doctorate Thesis:

"DNA Microarray image processing based on advanced pattern recognition techniques"

Emmanouil I. Athanasiadis

Medical Image and Signal Processing (*MED.I.S.P.*) Laboratory, Department of Medical Instruments Technology, Technological Educational Institute of Athens, Ag. Spyridonos Street, Aigaleo, 122 10, Athens, Greece.
Email: mathan@teiath.gr Tel: 0030-210-5385375

In the present thesis, a novel gridding technique, as well as, two new segmentation methods applied to complementary *DNA* (*cDNA*) microarray images are proposed.

More specifically, a new gridding method based on continuous wavelet transform (*CWT*) was performed. Line profiles of x and y axis were calculated, resulting to 2 different signals. These signals were independently processed by means of CWT at 15 different levels, using daubechies 4 mother wavelet. A summation, point by point, was performed on the processed signals, in order to suppress noise and enhance spot's differences. Additionally, a wavelet based hard thresholding filter was applied to each signal for the task of alleviating the noise of the signals. 10 real microarray images were used in order to visually assess the performance of our gridding method. Each microarray image contained 4 sub-arrays, each sub-array 40x40 spots, thus, 6400 spots totally. According to our results, the accuracy of our algorithm was 98% in all 10 images and in all spots. Additionally, processing time was less than 3 sec on a 1024×1024×16 microarray image, rendering the method a promising technique for an efficient and fully automatic gridding processing.

Following the gridding process, the Gaussian Mixture Model (GMM) and the Fuzzy GMM algorithms were applied to each cell, with the purpose of discriminating foreground from background. In addition, Markov Random Field (*MRF*), as well as, a proposed wavelet based MRF model (*SMRF*) were implemented. The segmentation abilities of all the algorithms were evaluated by means of the segmentation matching factor (*SMF*), the Coefficient of Determination (r^2), and the concordance correlation (p_c). Indirect accuracy performances were also tested on the experimental images by means of the Mean Absolute Error (*MAE*) and the Coefficient of Variation (*CV*). In the latter case, SPOT and SCANALYZE software results were also tested. In the former case, SMRF attained the best *SMF*, r^2, and p_c (92.66%, 0.923, and 0.88, respectively) scores, whereas, in the latter case scored *MAE* and *CV*, of 497 and 0.88, respectively. Our results support the performance superiority of the SMRF algorithm in segmenting *cDNA* images.