A Fully Automatic Approach for Enhancement of Microarray Images

Nagaraja J and Manjunath S S
Dayananda Sagar College of Engineering, Bangalore 560078, India
nagaraj.dsce@gmail.com, mnj_ss2002@yahoo.co.in

Dr.Pradeep
Rajarajeshwari College of Engineering, Bangalore 560074, India
pradeepbs78@yahoo.com

***Abstract*—DNA microarray technology has promised a very accelerating research inclination in recent years. There are numerous applications of this technology, including clinical diagnosis and treatment, drug design and discovery, tumor detection, and in the environmental health research. Enhancement is the major pre-processing step in microarray image analysis. Microarray images when corrupted with noise may drastically affect the subsequent stages of image analysis and finally affects gene expression profile. This paper presents an approach that achieves an automated way for applying mathematical morphology for the enhancement of microarray images. White and black top-hat transform is performed to denoise the image. A threshold is estimated which is dependent on image characteristics to remove artifacts present in the image. Experiments on Stanford, TBDB and UNC database illustrate robustness of the proposed approach in the presence of noise, artifacts and weakly expressed spots. Experimental results and analysis illustrates the performance of the proposed method with the contemporary methods discussed in the literature.**

***Index Terms*—microarray, morphology and white & black top-hat transform.**

I. Introduction

DNA microarray technology [1] has a large impact in many application areas, such as diagnosis of human diseases and treatments (determination of risk factors, monitoring disease stage and treatment progress, etc.), agricultural development (plant biotechnology), and quantification of genetically modified organisms, drug discovery, and design. In cDNA microarrays, a set of genetic DNA probes (from several hundreds to some thousands) are spotted on a slide. Two populations of mRNA, tagged with fluorescent dyes, are then hybridized with the slide spots, and finally the slide is read with a scanner. The outlined process produces two images, one for each mRNA population, each of which varies in intensity according to the level of hybridization represented as the quantity of fluorescent dye contained in each spot.

Microarray image processing consists of the following sequence of three stages 1. Gridding, separation of spots by assignment of image coordinates to the spots [2]. 2. Segmentation, separation between the foreground and background pixels and 3. Intensity extraction, computation of the average foreground and background intensities for each spot of the array [3]. Microarray image may contain different sources of errors. Such as electronic noise, dust on slide, photon noise and other sources causes high level of noise which may propagate through higher image analysis leading to difficulty in identifying the genes that each type of cells is expressing to draw accurate biological conclusions. Spot recognition is complicated task as microarray image gets corrupted by noise sources during image acquisition also bright artifacts may be detected incorrectly as spots of microarray image. Hence it is very much essential to remove the noise present in the image .The image enhancement is necessary to improve the interpretability of information in images to provide better input for the higher image processing applications. Low quality images are thus to be enhanced by appropriate methods to interpret the accurate expression levels.

Image Enhancement improves the image quality by refining the image with respect to structural content, statistical content, edges, textures and presence of noise. It can be further used for accurate measurement of gene expression profiling.

The organization of rest of the paper is as follows: Section II describes the literature survey carried out in the areas of microarray image enhancement. Section III presents morphological approach which makes use of white and black top-hat transform to enhance microarray images. Section IV highlights the results of extensive experimentation conducted on some benchmark images. Finally conclusion is discussed.

II. Related Work

The literature survey carried out has revealed that a fair amount of research has been put in the areas of

Manuscript received December 21, 2012; revised April 18, 2013.

doi: 10.12720/joace.1.4.285-289

microarray image enhancement. X. H. Wang, Robert S. H. Istepanian and Yong Hua Song [4] have proposed a new approach based on wavelet theory to provide a denoising approach for eliminating noise source and ensure better gene expression. Method of denoising applies stationary wavelet transform to pre-process the microarray images for removing the random noises. Rastislav Lukac and Bogdan Smolka [5] have proposed novel method of noise reduction, which is capable of attenuating both impulse and Gauassian noise, while preserving and even denoising the sharpness of the image edges. R. Lukac, et.al [6] have proposed vector fuzzy filtering framework to denoise cDNA microarray images. This method adaptively determines weights in the filtering structure and provides different filter structures. Noise removal using smoothening of coefficients of highest sub bands in wavelet domain is described by Mario Mastriani and Alberto E. Giraldez [7]. Denoising switching scheme based on the impulse detection mechanism using peer group concept is discussed by N. Plataniotis et.al [8]. A two-stage approach for noise removal that processes the additive and the multiplicative noise component, which decomposes the signal by a multiresolution transform, is described by Hara Stefanou, Thanasis Margaritis, Dimitris Kafetzopoulos, Konstantinos Marias and Panagiotis Tsakalides [9]. Guifang Shao, Hong Mi, Qifeng Zhou and Linkai Luo [10] have proposed a new algorithm for noise reduction which included two parts: edge noise reduction and highly fluorescence noise reduction. Ali Zifan, Mohammad Hassan Moradi and Shahriar Gharibzadeh [11] have proposed an approach using of decimated and undecimated multiwavelet transforms. Denoising of microarray images using the standard maximum a posteriori and linear minimum mean squared error estimation criteria is discussed by Tamanna Howlader et.al [12]. J. K. Meher et.al [13] have proposed novel pre-processing techniques such as optimized spatial resolution (OSR) and spatial domain filtering (SDF) for reduction of noise from microarray data and reduction of error during quantification process for estimating the microarray spots accurately to determine expression level of genes. Weng Guirong has proposed a novel filtering method to denoise microarray images using edge enhancing diffusion method [14].Factorial analysis on simulated microarray images to study the effects and interaction of noise types at different noise levels is discussed by yogananda Balagurunathan et.al [15]. Chiatra Gopalappa et.al [16] have proposed a novel methodology for identification and scanning noise from microarray images using a dual tree complex wavelet transform. A two phase scheme for removing impulse noise from microarray images by preserving the feature of interest is discussed by Ram murugesan et.al [17].Arunakumari Kakumani et.al [18] have proposed a method to denoise microarray images using independent component analysis. Enhancement approach which uses principles of fuzzy logic in conjunction with data adaptive filter to enhance noisy microarray images is presented by Rastislav Lukac et.al [19]. Wang li-qiang et.al [20] presents a novel method to reduce impulse noise by employing the switching scheme which uses differences between the standard deviation of the pixels within the filter window and the current pixel of concern. Nader Suffarian et.al [21] have proposed an approach which is implemented as conditional sub-block bi-histogram equalization (CSBE) which has the ability to improve the gridding results in DNA microarray analysis.

Most of the methods proposed by researchers have either considered high SNR (signal-to-noise ratio) images or various assumptions on factors such as type of threshloding used, parametric assumptions and decomposition levels, which in turn leads to misclassification of foreground pixels from the background pixels in the segmentation process and finally affects gene expression profile. Also some of the methods have discussed only impulse, Gaussian noise and fluorescent noise. A method has to be proposed which works with low SNR images and estimate other types of noises so has to accurately denoise the image. This is very essential at the pre-processing stage because in the microarray image analysis each stage affects subsequent stage, so that an accurate biological conclusion can be drawn. Denoising of microarray image is a challenging task in the pre-processing step of microarray image analysis. So, techniques without the above mentioned constraints and which depends exclusively on the image characteristics is in demand. Fig. 1 shows a sub grid of microarray image.

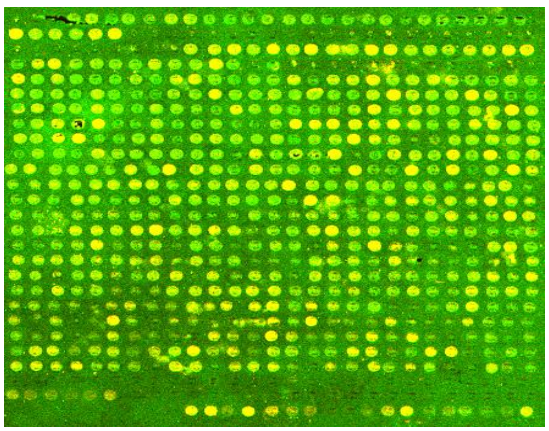

Figure 1. Sub grid of microarray image (Database: TBDB, ID: 422471)

III. Proposed Approach

A. Denoising using Mathematcial Morphology

In this section an automated way to achieve enhancement of microarray images using mathematical morphology is presented. The noisy microarray image I(x,y) of x rows and y columns is preprocessed to get grayscale image P(x,y). Fig. 2 shows dataflow diagram of the proposed approach.

White top-hat transform (WTT) is performed on P(x,y). WTT is the difference between P(x,y) and its opening by

structuring element (S).Morphological opening is erosion followed by dilation. The resultant image will be WTT(x,y).

$$\left(WTT(x,y) = (P(x,y) - (P(x,y) \circ S))\right) \quad (1)$$

$$\left(P(x,y) \circ S = ((P(x,y) \Theta S) \oplus S)\right) \quad (2)$$

Black top-hat transform (BTT) is performed on P(x,y). BTT is the difference between closing by structuring element (S) and P(x,y).Morphological closing is dilation followed by erosion. The resultant image will be BTT(x,y).

$$\left(BTT(x,y) = ((P(x,y) \bullet S) - P(x,y))\right) \quad (3)$$

$$\left(P(x,y) \bullet S = ((P(x,y) \oplus S) \Theta S)\right) \quad (4)$$

where $\oplus$ and Θ denotes dilation and erosion respectively. $\circ$ and $\bullet$ denotes opening and closing respectively.

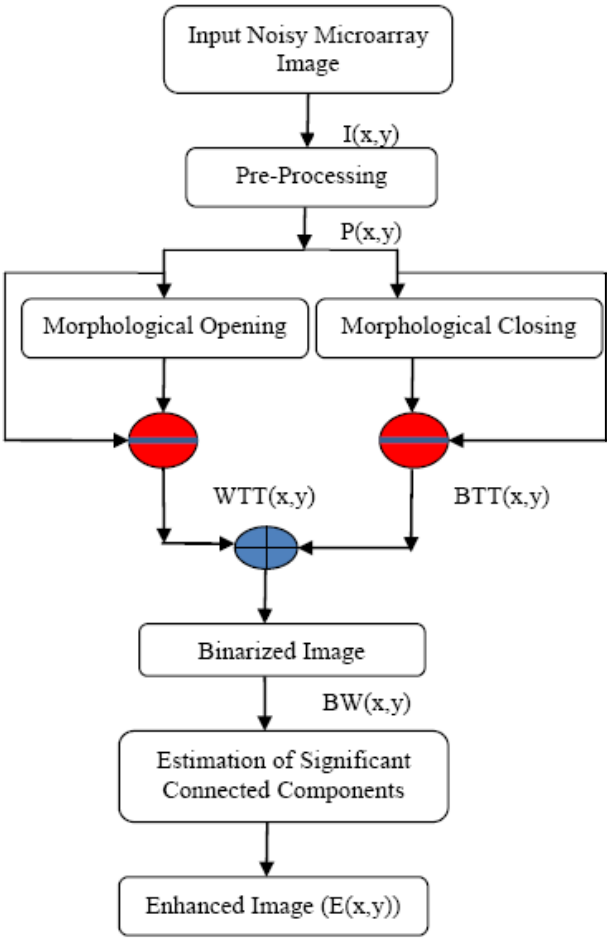

Figure 2. dataflow diagram of proposed approach

Dilation is an operation that grows or thickens: objects in a gray scale image. The specific manner and extent of thickening is controlled by a shape referred through structuring element. Erosion shrinks or thins objects in a gray scale image. The manner and extent of shrinking is controlled by a structuring element. Structuring element is still the key factor of morphology operations. Applying structuring elements with different radius leads to diverse results of analyzing and processing of geometric characteristic. Therefore, structuring element determines the effect and performance of morphological transformation. Structuring element used for dilation and erosion process is shown in Fig. 3.

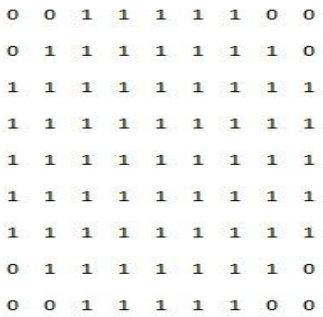

Figure 3. Disk structuring element with radius-5

B. Threshold

To remove artifacts present in the subgrid after morphological denoising, a threshold is estimated which depends on image characteristics. This threshold used estimate significant connected components to draw accurate biological conclusions.

Threshold value is calculated using the equation 5.

$$\left(C_T = \frac{\text{Total Number of pixels}}{\text{Total number of connected Components}}\right) \quad (5)$$

If the number of pixels in a component is less than threshold value (C_T) in each segment, then remove the spot (insignificant spot) by setting intensity zero to all pixels in that component.

The results of the proposed filtering process in removing the insignificant spots using the threshold values are reported in Table I.

TABLE I. ESTIMATED THRESHOLD VALUES OF PROPOSED APPROACH

Noisy Image ID (Database)	Total Number of Connected Components in Noisy Image	Threshold Value (CT)	Total Number of Conncted Components in Enhanced Image
39119 (TBDB)	820	67	796
40031 (Stanfrod)	2005	47	1376
44004 (TBDB)	777	48	690
17931 (UNC)	721	13	660

IV. RESULTS ANS PERFORMANCE ANALYSIS

In this section, the performance of the proposed approach is evaluated on real noisy microarray images drawn from SMD (Stanford microarray database), UNC (University of North California microarray database) and TBDB database. The images are available for free download from https://genome.unc.edu, http://www.tbdb.org/cgi-in/data/clickable.pl.html. Fig.4

shows noisy microarray image and in Fig.5 Enhanced image using proposed approach is shown.

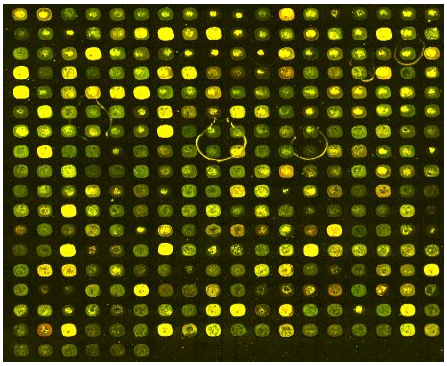

Figure 4. Noisy subgrid, Image ID: 39119, Database:TBDB

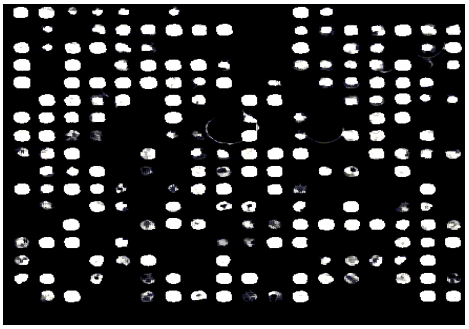

Figure 5. Enhanced subgrid, Image ID: 39119, Database:TBDB

Fig.6. shows one subgrid of noisy microarray image. As discussed in section III, Morphological dilation, erosion and threshold are used to perform filtering. Fig.7. shows enhanced image from this observation, it infers that, most of the contaminated (insignificant, noisy) pixels are removed.

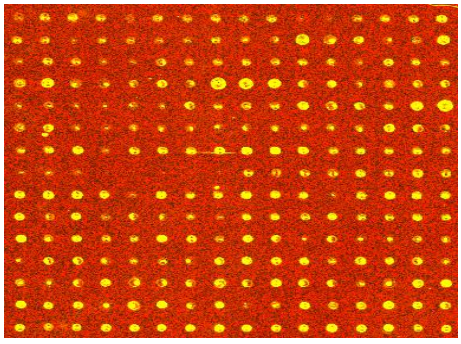

Figure 6. Noisy subgrid, Image ID:35964, Database:TBDB

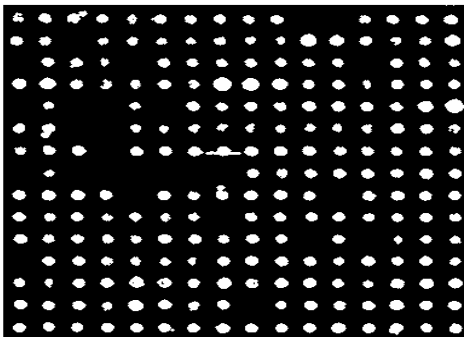

Figure 7. Enhanced subgrid, Image ID:35964, Database:TBDB

To quantify both the degree of filtering as well as the improvements due to enhancement algorithms, various performance measures are used. Such as mean squared error and peak signal to noise ratio. Higher the peak signal to noise ratio value higher is the quality of the image and lower the mean squared value higher is the image quality. Here we have compared the performance of different filters.

Performance analysis is shown in Fig. 8 and 9.Table II and III illustrates comparative results of the proposed method with existing filters.

TABLE II. .NUMERICAL VALUES OF PEAK SIGNAL TO NOISE RATIO FOR DENOISNG METHODS

IMAGE ID	PEAK SIGNAL TO NOISE RATIO			
	WIENER FILTER	LOW PASS FILTER	MEDIAN FILTER	PROPOSED APPROACH
34133	1.152	0.92	1.0742	1.2472
22950	0.4842	0.4594	0.3994	2.7730
34130	0.5442	0.6598	0.3788	2.5538
35964	0.4718	0.5918	0.4390	0.9336

TABLE III. NUMERICAL VALUES OF MEAN SQUARE ERROR DENOISNG METHODS

IMAGE ID	MEAN SQURE ERROR			
	WIENER FILTER	LOW PASS FILTER	MEDIAN FILTER	PROPOSED APPROACH
34133	1.023	0.987	1.231	0.521
22950	0.965	1.675	1.876	0.434
34130	1.425	0.875	1.743	0.673
35964	1.236	0.754	1.421	0.572

Figure 8. Comparison chart of PSNR of different denoising methods for microarray images

Figure 9. Comparison chart of MSE of different denoising methods for microarray images

V. Conclusion

In this work automatic approach for enhancement of microarray image is presented. The noise removal is performed through white and black top- hat transform which are implemented using morphological dilation and erosion .To the morphological image threshold is used to eliminate insignificant spots. From the experimental results it has been observed that most of the contaminated pixels have been removed from the image. The entire process is robust, in the presence of noise, artifacts and weakly expressed spots. The proposed work can be used at pre-processing phase in microarray image analysis before using it in any of the stages of microarray image analysis, which then results in accurate gene expression profiling.

References

[1] Y. F. Leung and D. Cavalieri, "Fundamentals of cDNA microarray data analysis," *TRENDS in Genetics* , vol. 19, no. 11, pp. 649-659, 2003.

[2] P. Bajcsy, "Gridline: Automatic grid alignment in DNA microarray scans," *IEEE Trans. Image Process*, vol. 13, no. 1, pp. 15–25, 2004.

[3] M. Steinfath, W. Wruck, H. Seidel, H. Lehrach, U. Radelof, and J. O'Brien, "Automated image analysis for array hybridization experiments," *Bioinformatics*, vol. 17, no. 7, pp. 634–641, Jul. 1, 2001

[4] X. H. Wang, R. S. H. Istepanian, and Y. H. Song, "Microarray image enhancement by denoising using stationary wavelet transforms," *IEEE Transactions Nanobiosciene*, vol. 2, no. 4, pp. 184–189, 2003.

[5] R. Lukac and B. Smolka, "Application of the adaptive center weighted vector median framework for the enhancement of cDNA microarray images," *International Journal of Applied Mathematics and Computer Science*, vol. 13, no. 3, pp. 369–383, 2003.

[6] R. Lukaca, K. N. Plataniotis, B. Smolka, and A. N. Venetsanopoulos, "cDNA microarray image processing using fuzzy vector filtering frame work," *Fuzzy Sets and Systems*, ELSEVIER, pp. 17-35, 2005.

[7] M. Mastriani and A. E. Giraldez, "Microarray denoising via smoothing of coefficients in wavelet domain," *International Journal of Biological, Biomedical and Medical Sciences*, pp. 7-14, 2005.

[8] B. Smolka, R. Lukac, and K. N. Plataniotis ,"Fast noise reduction in cDNA microarray images," in *Proc. IEEE, 23rd Biennial Symposium on Communications*, 2006, pp.348-351.

[9] H. Stefanou, T. Margaritis, D. Kafetzopoulos, K. Marias, and P. Tsakalides, "Microarray image denoising using a two- stage multiresolution technique," in *Proc. IEEE International Conference on Bioinformatics and Biomedicine*, 2007, pp. 383-389.

[10] G. F. Shao, H. Mi, Q. F. Zhou, and L. K. Luo, "Noise estimation and reduction in microarray images," in *Proc. IEEE, World Congress on Computer Science and Information Engineering*, 2009, pp. 564-568.

[11] A. Zifan , M. Hassan Moradi, and S. Gharibzadeh, "Microarray image nhacment using decimated and undecimated wavelet transforms," *SIViP*, pp. 177-185, 2010.

[12] T. Howlader and Y. P. Chaubey, "Noise reduction of cDNA microarray images using complex wavelets," *IEEE Transactions on Image Processing*, vol. 19, pp. 1953-1967, 2010.

[13] J. K. Meher, P. K. Meher, and G. N. Dash, "Preprocessing of microarray by integrated OSR and SDF approach for effective denoising and quantification," *IPCSIT*, vol. 4, pp. 158- 163, 2011.

[14] W. guirong, "cDNA Microarray image processing using morphological operator and edge Enhancing diffusion," in *Proc. 3rd International Conference on Bioinformatics and Biomedical Engineering*, 2009, pp. 1- 4.

[15] Y. Balagurunathan, N. Wang, E. R. Dougherty, D. Nguyen, and Y. Chen,"Noise factor analysis for cDNA microarrays," *Journal of Biomedical Optics*, vol. 9, no. 4, pp. 663- 678, 2004.

[16] C. Gopalappa, T. K. Das, St. Enkemann, and S. Eschrich, "Removal of hybridization and scanning noise from microarrays," *IEEE Transactions on Nano Bioscience*, vol. 3, pp. 210-218, 2009.

[17] R. Murugesan and V. Thavavel, "A two-phase scheme for microarray image restoration," *Journal of Information and Computing Science*, vol. 2, no. 4, pp. 317-320, 2007.

[18] A. Kakumani, K. A. Mendhurwar, and R. Kakumani, "Microarray image denoising using independent component analysis," *International Journal of Computer Applications*, vol. 1, no. 11, pp. 87-95, 2010.

[19] R. Lukac, K. N. Plataniotis, B.n Smolka, and A. N. Venetsanopoulos, "A data-adaptive approach to cDNA microarray image enhancement," *ICCS*, pp. 886– 893, 2005.

[20] W. LQ, N. XX, L. ZK, X. F. Zheng, and L. YS, "Enhancing the quality metric of protein microarray image," *Journal of Zhejiang University Science*, vol. 5, no. 12, pp. 1621- 1628, 2004.

[21] N. Saffarian and J. J. Zou, "DNA microarray image enhancement using conditional sub-block bi-histogram equalization," in *Proc. IEEE International Conference on Video and Signal Based Surveillance*, 2006, pp. 86.

Nagaraja J has received B.E degree in2007 from VTU University, Belgaum and M.Tech degree in 2009 from VTU University, Belgaum, Karnataka,India. Currently he is working as a Lecturer at Dayananda Sagar College of Engineering, Karnataka, India and His experience in teaching started from the year 2009. Currently he is pursuing PhD in VTU University. His areas of interest include microarray image processing, medical image segmentation and clustering algorithms.

Dr.Pradeep.BS.B.E(CSE).,MTech(Networking)., Ph.D(CSE)He is working as Professor & HOD in CSE dept. of RRCE, Bangalore, Karnataka. 11+ years experienced in teaching, R&D and Administration. He has published more than 47 papers in International/National Journals and conferences. His areas of interest are image processing, mobile computing.

Manjunath S. S has received B.E degree in 2000from Mysore University, Mysore and M.Tech degree in 2005 from VTU University, Belgaum, and Karnataka India. Currently he is working as an Assistant Professor at Dayananda Sagar College of Engineering, Karnataka, India and His experience in teaching started from the year 2000. Currently he is pursuing PhD in Mysore University. His areas of interests include microarray image processing, medical image segmentation and clustering algorithms.

