APPLICATION OF MATHEMATICAL MORPHOLOGY FOR THE ENHANCEMENT OF MICROARRAY IMAGES

Nagaraja J[1], Manjunath S.S[2], Lalitha Rangarajan[3], Harish Kumar. N[4]
[1, 2 & 4]Department of CSE, Dayananda Sagar College of Engineering, Bangalore, India
[3]Department of CSE, Mysore University, Mysore, India

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, tumour 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. In this paper a fully automatic technique to enhance microarray images is presented using mathematical morphology. Experiments on Stanford and TBDB 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.

KEYWORDS: *Microarray, Dilation, Erosion, Adaptive Threshold and Noisy microarray images.*

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 2 describes the literature survey carried out in the areas of microarray image enhancement. Section 3 presents morphological approach which makes use of top and bottom hat transform to enhance microarray images. Section 4 highlights the results of extensive experimentation conducted on some benchmark images. Finally conclusion is discussed.

II. REVIEW OF LITERATURE

The literature survey carried out has revealed that a fair amount of research has been put in the areas of 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. Figure.1 shows a subgrid of microarray image.

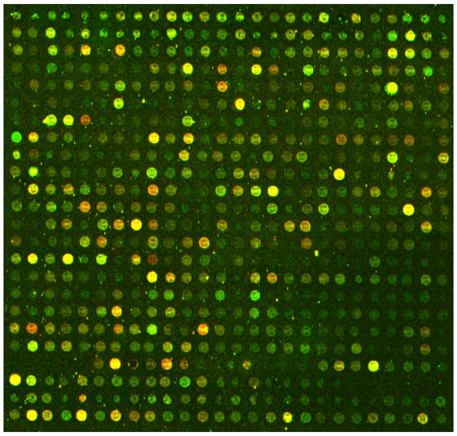


Figure. 1 Subgrid of Microarray image (ID: 32040)

III. ENHANCEMENT MODEL

The image enhancement is the process of improving the interpretability of information in images to provide better input for the higher image processing applications. Enhancement model illustrates phases involved to enhance microarray images as shown in Figure 2.

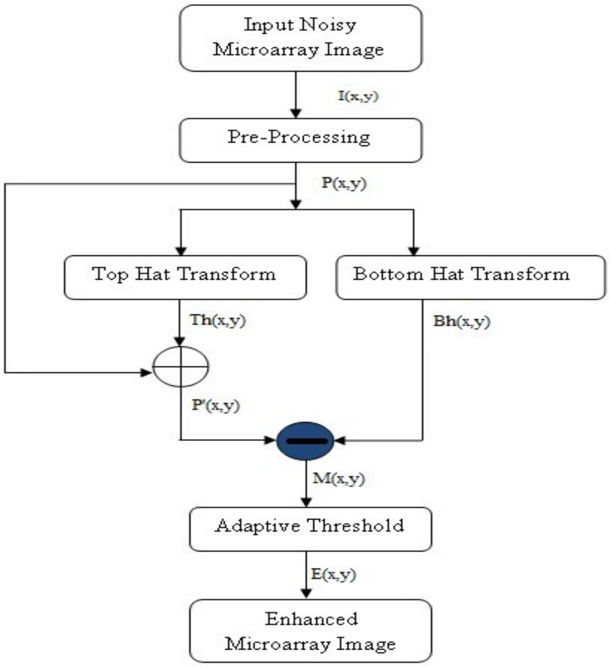
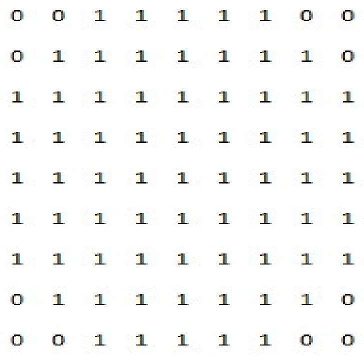


Figure 2. Enhancement Model

Mathematical morphology being used to remove artifacts and insignificant spots in the subgrid. In the pre-processing stage the noisy RGB image is converted in to gray level .The resulted image will be pre-processed image say P(x,y). The tophat and bottom hat transform is performed on P(x,y).The tophat is performed by erosion followed by dilation. The bottom hat is performed by dilation followed by erosion. 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 Figure.3.



0	0	1	1	1	1	1	0	0
0	1	1	1	1	1	1	1	0
1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1
0	1	1	1	1	1	1	1	0
0	0	1	1	1	1	1	0	0

Figure 3. structuring element with radius-5

After performing tophat transform (Th(x,y)) which will be added to pre-processed image (P(x,y)) results P'(x,y) . This is performed to improve quality of the image. P'(x,y) is subtracted from bottom hat transformed image Bh(x,y) to remove artifacts pixels in the microarray image.

To eliminate insignificant spots adaptive threshold being used. Thresholds on spot size are first computed on segments of the image. Insignificant spots are filtered using these thresholds. The gray level image is converted in to binary level with low threshold to reside the information available in the image. Binary image is divided into n segments. Number of segments can be increased depending on the level of noise. The subgrid is divided into 4 segments in the proposed approach as follows.

1st segment Rows=0 to r/2 Columns=0 to c/2	2nd segment Rows= 0 to r/2 Columns= c/2+1 to c
3rd segment Rows=r/2+1 to r Columns= 0 to c/2	4th segment Rows= r/2 +1to r Columns= c/2+1 to c

where r is the number of rows and c is number of columns of skew corrected image.

For each segment, the numbers of connected components are computed. The thresholds on spot size in each segment are calculated using the equation 1.

$$\left(T(i) = \frac{\text{Number of pixels in } i^{\text{th}} \text{ segment}}{\text{Total number of connected Components}} \right) \tag{1}$$

where i ranges from 1 to 4.

For example in the Figure 4, (ID: 32040) the number of bright pixels in the four segments are 5523, 5090, 6075, 2031. Total number of connected components is 894. The thresholds are 5523/894=6, 5090/894= 6, 6075/894=7, 2031/894=2.

The results of the proposed filtering process in removing the insignificant spots using the threshold value and execution time (τf) are reported in Table 1.

Execution time for the filtering process is proportional to number of spots in a noisy microarray image. Adaptive thresholds obtained in the previous step are used to filter insignificant noisy spots in the segments. If the number of pixels in a component are less than threshold value ($_{T(i)}$) in each segment, then remove the spot (insignificant spot) by setting intensity zero to all pixels in that component. The idea behind using adaptive threshold is, if in a sub array, suppose few successive columns or rows have tiny spots filtering using global threshold will eliminate all these spots.

Table 1. Estimated Threshold Values and Execution Time (T_f) of the Proposed Filtering Process.

Noisy Image ID	Thresholds on a subarray	#spots in the subarray	Execution time(τ_f) in seconds
40031 (Stanford)	T1=17 T2=11 T3=10 T4=09	146	10
44004 (TBDB)	T1=14 T2=12 T3=12 T4=10	777	15
17931 (Stanford)	T1=28 T2=23 T3=47 T4=41	721	13
39119 (TBDB)	T1=20 T2=18 T3=15 T4=15	562	8

IV. RESULTS AND 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 TB database. The images are available for free download from website [22, 23]. Figure.4 shows noisy microarray image and in Figure.5 Enhanced image using proposed approach is shown.

Enhancement is very much essential as it helps the biologists to take the decisions on gene expression analysis, gene discovery, and drug analysis etc. with the clear spots the accuracy of analysis improves. Application of mathematical morphology yields a high quality image and it reveals most of the unidentified spots clearly.

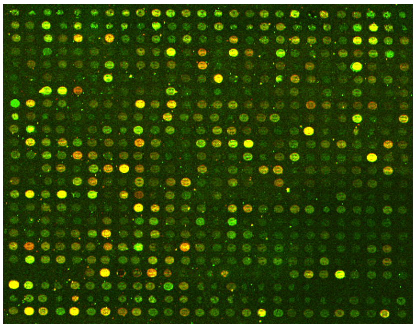


Figure. 4 Noisy subgrid, Image ID: 32040

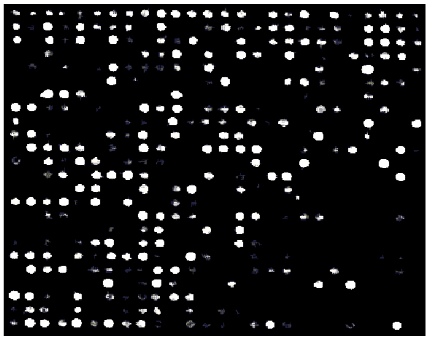


Fig. 5.Enhanced subgrid, Image ID: 32040

Figure.6. shows one subgrid of noisy microarray image. As discussed in section 3, Morphological dilation, erosion and Adaptive threshold are used to perform filtering. Figure.7. shows enhanced image from this observation, it infers that, most of the contaminated (insignificant, noisy) pixels are removed.

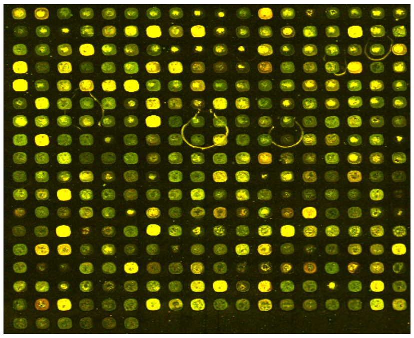


Figure. 6 Noisy subgrid, Image ID: 39119, Database:TBDB

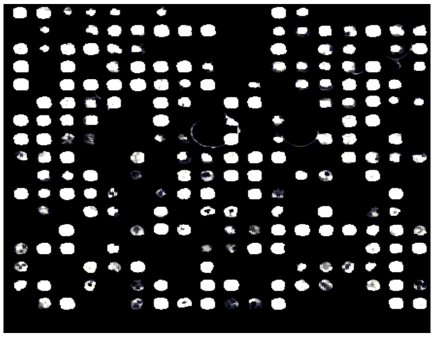


Figure. 7 Enhanced subgrid, Image ID: 39119, Database:TBDB

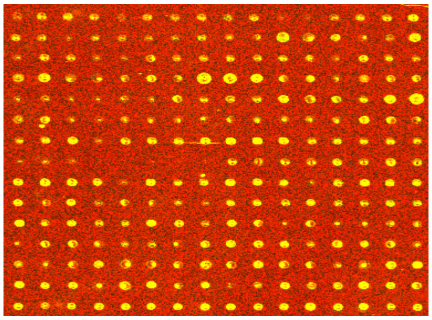


Figure. 8 Noisy subgrid, Image ID: 35964, Database:TBDB

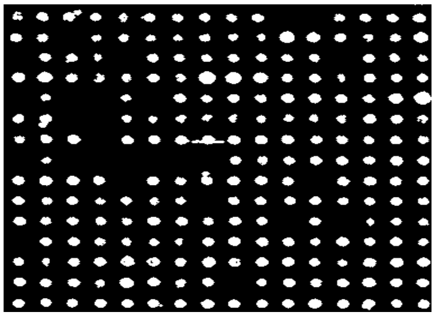


Figure. 9. Enhanced subgrid , Image ID: 35964, Database:TBDB

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 and bilateral works good with removing all the noise content from the image. Performance analysis is shown in Figure 10 and 11. Table 2 and 3 illustrates comparative results of the proposed method with existing filters.

Table 2. Numerical Values on Signal to Noise Ratio for Denoising Methods

Image Id	**Peak Signal to Noise ratio in db**				
	Weiner	**Median**	**Gaussian**	**Bayes**	**Proposed**
34133(TBDB)	79.00	79.06	77.69	82.03	87.73
32070(TBDB)	67.8758	68.5429	67.9479	80.7468	86.7109
422471(Stanford)	72.2693	72.5676	71.7460	83.2749	87.1175
400311(UNC)	70.1570	70.6426	69.5335	82.1275	87.2587

Table 3. Numerical Values on Mean Square Error for Denoising Methods

Image Id	**Mean Square Error**				
	Weiner	**Median**	**Gaussian**	**Bayes**	**Proposed**
34133(TBDB)	23.36	23.96	27.4681	17.8004	10.0690
32070(TBDB)	20.23	22.13	20.12	18.14	11.1429
422471(Stanford)	18.1872	19.1612	22.1912	15.7216	10.7058
400311(UNC)	17.1614	18.2532	14.5212	13.1714	10.556

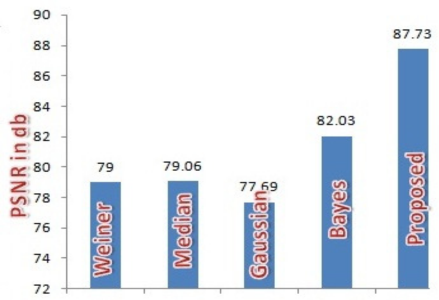


Figure 10. Comparison chart of PSNR of different denoising methods for microarray images

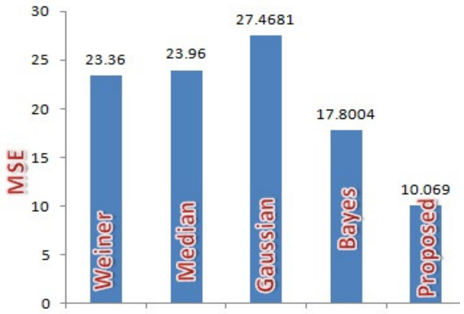


Figure 11. Comparison chart of MSE of different denoising methods for microarray images

V. CONCLUSION

In this work automatic technique for enhancement of microarray image is presented. The noise removal is performed through top hat and bottom hat transform which are implemented using morphological dilation and erosion .To the morphological image adaptive 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] Yuk Fai Leung and Duccio Cavalieri, (2003) "Fundamentals of cDNA microarray data analysis" in TRENDS in Genetics Vol.19 No.11 pp 649-659.

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

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

[4] X. H.Wang, R. S. H. Istepanian, and Y. H. Song, (2003),"Microarray image enhancement by denoising using stationary wavelet transforms". IEEE Transactions Nanobiosciene , 2(4):184 – 189.

[5] Rastislav Lukac and Bogdan Smolka, (2003), "Application of the adaptive center weighted vector median framework for the enhancement of cDNA microarray images", International Journal of applied mathematics and Computer Science (amcs), Vol. 13, No. 3, 369–383.

[6] Rastislav Lukaca, Konstantinos N. Plataniotis, Bogdan Smolka and Anastasios N.Venetsanopoulos, (2005) " cDNA microarray image processing using fuzzy vector filtering frame work", Fuzzy Sets and Systems ,ELSEVIER, 17-35.

[7] Mario Mastriani, and Alberto E.Giraldez, (2005), "Microarray denoising via smoothing of coefficients in wavelet domain", international journal of biological, biomedical and medical sciences, pp7-14.

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

[9] Hara Stefanou, Thanasis Margaritis, Dimitri Kafetzopoulos, Konstantinos Marias and Panagiotis Tsakalides, (2007), " Microarray image denoising using a two- stage multiresolution technique", IEEE International Conference on Bioinformatics and Biomedicine ,pp.383- 389.

[10] Guifang Shao, Hong Mi, Qifeng Zhou and Linkai Luo, (2009), " Noise estimation and reduction in microarray images ", IEEE, World Congress on Computer Science and Information Engineering,pp.564-568.

[11] Ali Zifan , Mohammad Hassan Moradi and Shahriar Gharibzadeh, (2010),"Microarray image nhacment using decimated and undecimated wavelet transforms", SIViP, pp.177- 185.

[12] Tamanna Howlader and Yogendra P. Chaubey, (2010),"Noise reduction of cDNA microarray images using complex wavelets", IEEE Transactions on Image Processing, Vol.19.pp.1953-1967.

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

[14] Weng guirong (2009), "CDNA Microarray image processing using morphological operator and edge-Enhancing diffusion", 3rd international conference on bioinformatics and biomedical engineering, pp. 1-4.

[15] Yoganand Balagurunathan, Naisyin Wang, Edward R. Dougherty, Danh Nguyen, Yidong Chen, (2004), "Noise factor analysis for cDNA microarrays", Journal of Biomedical Optics, Vol. 9 , No. 4 , pp. 663-678.

[16] Chaitra Gopalappa, Tapas K. Das, Steven Enkemann, and Steven Eschrich , (2009)"Removal of Hybridization and Scanning Noise from Microarrays", IEEE Transactions on NanoBioscience, Vol. 3, pp. 210-218.

[17] Ram Murugesan, V.Thavavel, (2007), " A Two-phase Scheme for Microarray Image Restoration" Journal of Information and Computing Science, Vol. 2, No. 4, pp. 317-320.

[18] Arunakumari Kakumani, Kaustubha A. Mendhurwar, Rajasekhar Kakumani, (2010), "Microarray Image Denoising using Independent Component Analysis", Vol. 1, No. 11, pp. 87-95.

[19] Rastislav Lukac, Konstantinos N. Plataniotis, Bogdan Smolka, Anastasios N. Venetsanopoulos, (2005), "A Data-Adaptive Approach to cDNA Microarray Image Enhancement", ICCS 2005, pp. 886–893.

[20] Wang LQ, Ni XX, Lu ZK, Zheng XF, Li YS, (2004), "Enhancing The Quality Metric Of Protein Microarray Image", Journal of Zhejiang University Science, Vol. 5, No. 12, pp. 1621- 1628.

[21] Nader Saffarian, Ju Jai Zou, (2006), "DNA Microarray Image Enhancement Using Conditional Sub-Block Bi-Histogram Equalization", IEEE International Conference on Video and Signal Based Surveillance,pp.86.

[22] https://genome.unc.edu

[23] http://www.tbdb.org/cgi-in/data/clickable.pl.html

Authors

Nagaraja J has received B.E degree in 2007 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 interests include microarray image processing, medical image segmentation and clustering algorithms.

Manjunath S.S has received B.E degree in 2000from Mysore University, Mysore and M.Tech degree in 2005 from VTU University, Belgaum, Karnataka India. Currently he is working as a 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.

Lalitha Rangarajan has received master degree in Mathematics from Madras University, India and from the Department of Industrial Engineering Purdue University. She completed PhD in Computer science from university of Mysore, India. She has been teaching courses in mathematics on operation research and computer science for master degree students for more than 25years. She is presently a Reader at Department of Computer Science, University of Mysore, India. Her current research interests are Image Retrieval Feature Reduction and Bioinformatics. She has more than 40 publications in reputed Journals and conferences.

Harish Kumar .N has received B.E degree in 2009 from VTU University, Belgaum and M.Tech degree in 2011 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 2011. His areas of interests include microarray image processing, medical image segmentation and clustering algorithms.