# Study and Analysis of Microarray Denoising Using Systholic Boolean Orthonormalizer Network in Wavelet Domain

Vishakha P.S.

Dept. of EXTC, MumbaiUniversity SSPM's COE, Kankavli, India. Supriya S.T.

Dept. of EXTC, Mumbai University SSPM's COE, Kankavli, India.

## ABSTRACT

In this paper, we present a new approach to deal with the noise inherent in the microarray image processing procedure. The method is based on the following procedure: We apply 1) Bidimentional Discrete Wavelet Transform (DWT-2D) to the Noisy Microarray, 2) scaling and rounding to the coefficients of the highest subbands (to obtain integer and positive coefficients), 3) bit-slicing to the new highest subbands (to obtain bit-planes), 4) then we apply the Systholic Boolean Orthonormalizer Network (SBON) to the input bit-plane set and we obtain two orthonormal otput bit-plane sets (in a Boolean sense), we project a set on the other one, by means of an AND operation, and then, 5) we apply re-assembling, and, 6) rescaling. Finally, 7) we apply Inverse DWT-2D and reconstruct a microarray from the modified wavelet coefficients. Denoising results compare favorably to the most of methods in use at the moment.

## 1. INTRODUCTION

A microarray is affected by noise in its acquisition and processing. Microarray denoising is used to remove the additive noise while retaining as much as possible the important image features. In the recent years there has been an important amount of research on wavelet thresholding and threshold selection for bioimages denoising, e.g., microarray images [1], [2], because wavelet provides an appropriate basis for separating noisy signal from the image signal. The motivation is that as the wavelet transform is good at energy compaction, the small coefficients are more likely due to noise and large coefficient due to important signal features [3]-[5]. These small coefficients can be thresholded without affecting the significant features of the image.

In general, the results of the microarray processing combine two sample images that after further image processing, gene expression data can be produced for further analysis, such as gene clustering or identification [1], [2]. These three crucial steps, experiment, image processing and data analysis, determine the success or not of the microarray analysis. Image processing plays a potentially large impact on the subsequent analysis. In recent years, a large number of commercial tools have been developed in microarray image processing [1], [2]. The tasks of all these tools mainly focus on two major targets. namely: spot segmentation and spot intensity extraction. However, the quality of the images from the experiments is not always perfect. The gene array experiments involve a large number of error-prone steps which lead to a high level of noise in the resulting images [1], [2]. Hence, the accuracy of the gene expressions derived from these images will largely be affected in the process.

In order to assure the accuracy of the gene expression, normally the replicated experiments and incorporated statistical methods are needed to estimate the errors [1], [2].

These methods deal mainly with measurement error, such as preparation of the sample, cross hybridization, and fluctuation of fluorescence value from gene to gene. But none deals particularly with the effect of the noise [1], [2].

In fact, the thresholding technique is the last approach based on wavelet theory to provide an enhanced approach for eliminating such noise source and ensure better gene expression. Thresholding is a simple non-linear technique, which operates on one wavelet coefficient at a time. In its basic form, each coefficient is thresholded by comparing against threshold, if the coefficient is smaller than threshold, set to zero; otherwise it is kept or modified. Replacing the small noisy coefficients by zero and inverse wavelet transform on the result may lead to reconstruction with the essential signal characteristics and with less noise. Since the work of Donoho & Johnstone [5], there has been much research on finding thresholds, however few are specifically designed for images [3], [4], [6].

## 2. BIDIMENTION DWT AND SBON

## 2.1 Bidimensional Discrete Wavelet Transform

The Bidimensional Discrete Wavelet Transform (DWT-2D) [8]-[17] corresponds to multiresolution approximation expressions. In practice, multiresolution analysis is carried out using 4 channel filter banks composed of a low-pass and a high-pass filter and each filter bank is then sampled at a half rate (1/2 down sampling) of the previous frequency. By repeating this procedure, it is possible to obtain wavelet transform of any order. The down sampling procedure keeps the scaling parameter constant (equal to <sup>1</sup>/<sub>2</sub>) throughout successive wavelet transforms so that is benefits for simple computer implemented in a separable way be filtering the lines and columns.

## 2.2 Synthetic Boolean Orthonormalizer Network

The SBON was introduced by Mastriani [19] as a Boolean Orthonormalization Process (BOP)to convert nonorthonormal Boolean basis, i.e., a set of non-orthonormal binary vectors (in a Boolean sense) to an orthonormal Boolean basis, i.e., a set of orthonormal binary vectors (in a Boolean sense). The BOP algorithm has a lot of fields of applications, e.g.: Steganography, Hopfield Networks, Boolean Correlation Matrix Memories, Bi-level image processing, lossy compression, iris, fingerprint and face recognition, improving edge detection and image segmentation, among others. That is to say, all those applications that need orthonormality in a Boolean sense. It is important to mention that the BOP is an extremely stable and fast algorithm.

## 2.3 Noise Sources and Its Statistical Measurement in Microarray Imaging

It is well known microarray technology can monitor thousand of DNA sequences in a high density array on a glass. The basic procedure for a microarray experiment is simply described as follow. Two mRNA samples are reversetranscribed into cDNA, labeled using different fluorescent dyes (e.g., the red fluorescent dye Cy5 and the green fluorescent dye Cy3), then mixed and hybridized with the arrayed DNA sequences. After this competitive hybridization, the slides are imaged using a scanner which makes fluorescence measurement for each dye. From the differential hybridization of the two samples, the relative abundance of the spotted DNA sequences can be assessed.

Exciting methods to reduce the noise source include using clean glass slide and using a higher laser power rather than higher PMT voltages. However, there are not adequate for the required image qualities and an enhanced software procedure embedded within the process in a much better alter-native. Here, we focus on the implementation of the *SBON* method (in wavelet domain) to the denoising on microarray images [2]. Yet there are some fundamental obstacles that need clarification before the full potential of microarrays can be explored. One of the major problems in interpretation of microarray data is that different clustering techniques produce different results.

On the other hand, the assessment parameters that are used to evaluate the performance of noise reduction [20], [21] are the following ones:

Average Absolute Difference (AAD):

$$AAD = \frac{\sum_{r,c} |I(r,c) - I_d(r,c)|}{R \star C}$$
(1)

Peak Signal to Noise Ratio (PSNR):

$$PSNR = \frac{\mathbf{R} * \mathbf{C} * \max \left( l(r, c) \right)^2}{\sum_{r,c} \left( \mathbf{I}(r, c) - I_d(r, c) \right) \mathbf{I}^2}$$
(2)

Signal to Noise Ratio (SNR):

$$SNR = \frac{\sum_{r,c} I(r,c)^2}{\sum_{r,c} \left( \left[ I(r,c) - I_d(r,c) \right] \right]^2}$$
(3)

Image Fidelity (IFy):

$$IFy = 1 - \frac{1}{SNR}$$
(4)

Correlation Quality (CQy):

$$CQy = \frac{\sum_{r,c} l(r,c) + l_d(r,c)}{\sum_{r,c} l(r,c)}$$
(5)

Structural Content (SCt):

$$SCt = \frac{\sum_{r,c} I(r,c)^2}{\sum_{r,c} I_d([r,c)]^2}$$
(6)

Where for an image of  $R^*C$  (rows-by-columns) pixels, r means row, c means column, I means original image (without

noise), and *Id* means denoised image. Such as, a lower *AAD* gives a "cleaner" image as more noise is reduced; larger *SNR* and *PSNR* indicates a smaller difference between the original (without noise) and denoised image; if *IFy* and *SCt* spread at 1, we will obtain an image *Id* of better quality; and a larger value of CQy usually corresponds to a better quantitative performance [20], [21].

## 3. LITERATURE SURVEY

Microarrays have become the tool of choice for the global analysis of gene expression. Powerful statistical tools are now available to analyze this expression and to gain an understanding of how changes in gene expression patterns impact biological systems. Innovations have been made since the origin of this imaging technique which dates back to the 1970s [23]. The analysis of such data has become a computationally- intensive task that requires technological developments at various stages, from the design of the array, to image analysis, database storage, data processing and clustering and information extraction. Further progress has been made in biology research and genomic clinical diagnosis. In recent years, a large number of commercial tools have been developed for microarray image processing Microarray arrays are a scientific tool that should be viewed in a similar fashion to any other laboratory technique with careful experimental planning, replication, and proper statistical analysis. A lot of research has been conducted to examine these data with statistical techniques to help discern possible patterns in the data.

Among other approaches used for microarray image denoising and enhancement are the iterative normalization of cDNA microarray data [24] and the image reconstruction technique (IRT) introduced by O'Neill and Magoulas [25].

A Method based on fuzzy vector filtering has also been used for processing microarray images [26].

Other methods based on the wavelet transform [27] and cellular networks [28] have also been used for this purpose [29, 30].

Wang et al. [31] proposed enhancing the microarray image using the stationary wavelet transform (SWT) [32].

Another approach based on a variant of the translationinvariant wavelet transform was proposed by Adjeroh [33].

## 4. **DISCUSSION**

The simulations demonstrate that the SBON technique improves the noise reduction performance to the maximum, for bioimages. Here, we present a set of experimental results using two bioimages. Such images were converted to bitmap file format for their treatment [22].

For statistical filters employed, i.e., Median, Lee, Kuan, Gamma-Map, Enhanced Lee, Frost, Enhanced Frost, Wiener, DS, and EDS, we use a reduction scheme [22]. Figure 2 shows the noisy (30 %) and filtered microarray images used in the first experiment of [1], with a 274-by-274 (pixels) by 65536 (gray levels) bitmap matrix. Table 1 summarizes the assessment parameters vs. 19 filters for Fig. 2, where En-Lee means Enhanced Lee Filter, En-Frost means Enhanced Frost Filter, ST means Soft-Thresholding, HT means Hard-Thresholding and SST means Semi-Soft-Thresholding.

The assessment parameters were applied to the whole image. Figure 3 shows the noisy (10 %) and filtered microarray images used in the second experiment of [1], with a 256-by-256 (pixels) by 65536 (gray levels) bitmap matrix. Table 2 summa-rizes the assessment parameters vs. 19 filters for Figure 2. In both cases, the bioimages were processed by using 10 statistical filters, VisuShrink with Daubechies 4 wavelet basis and 1 level of decomposition (improvements were not noticed with other basis of wavelets) [2], [3], [5], [6], [22], SureShrink, Oracle-Shrink, BayesShrink, NormalShrink, TNN [5]-[7], [22], and SBON respectively. Figures 2 and 3 summarize the edge preservation performance of the SBON technique vs. the rest of the filters with a considerably acceptable computational complexity. A 3-by-3 kernel was employed for all statistic noise filters. For *TNN* [7] the empirical function parameter value  $\lambda = 0.01$ .

For Lee, Enhanced Lee, Kuan, Gamma, Frost and Enhanced Frost filters the damping factor is set to 1, see [3], [22]. The quantitative results of Table 1 and 2 shows that the *SBON* technique can eliminate noise without distorting useful image information and without destroying the important image edges.

Also, in the experiment, the *SBON* outperformed the conventional and no conventional noise reducing filters in terms of edge preservation measured by Pratt figure of merit [21]. In nearly every case in every homogeneous region, the *SBON* produced the lowest standard deviation and was able to preserve the mean value of the region.

The numerical results are further supported by qualitative examination, as shown in Fig. 2 and 3.

On the other hand, all filters was applied to complete image, for Figure 2 (274-by-274) pixels and Figure 3 (256- by-256) pixels, and all the filters were implemented in MATLAB® (Mathworks, Natick, MA) on a PC with an Athlon (2.4 GHz) processor.

## 5. CONCLUSION

In this paper we have developed a SBON technique based tools for removing additive noise in microarrays. The simulations show that the SBON have better performance than the most commonly used filters for microarrays (for the studied benchmark parameters) which include statistical filters, wavelets, and a version of TNN. The SBON exploits the local coefficient of variations in reducing noise. The performance figures obtained by means of computer simulations reveal that the SBON technique provides superior performance in comparison to the above mentioned filters in terms of smoothing uniform regions and preserving edges and features. The effectiveness of the technique encourages the possibility of using the approach in a number of ultrasound and radar applications. Besides, the method is computationally efficient and can significantly reduce the noise while preserving the resolution of the original microarray image. Considerably increased Pratt's figure of merit strongly indicates improvement in detection performance. Also, cleaner images suggest potential improvements for classification and recognition. On the other hand, the drawback of applying the developed SBON technique for removing additive noise in microarrays is the increase in the computational complexity, for blame of the slicing process.

Finally, the natural extension of this work is in Synthetic Aperture Radar (SAR) images, as well as in multimedial applications.



## 4. REFERENCES

- E.C. Rouchka. (2004, April). Lecture 12: Microarray Image Analysis. [Online] Available: http://kbrin.abldg.louisville.edu/CECS694/Lecture12.ppt
- X.H. Wang, S.H. Istepanian, and Y.H. Song, "Microarray Image Enhancement by Denoising using Stationary Wavelet Transform," IEEE Transactions on Nanobioscience, vol.2, no. 4, pp. 184-189, December 2003. [Online]. Available: http://technology.kingston.ac.uk/momed/papers/IEEE\_T N\_Microarray\_Wavelet%20Denoising.pdf
- [3] H.S. Tan. (2001, October). Denoising of Noise Speckle in Radar Image. [Online]. Available: http://innovexpo.itee.uq.edu.au/2001/projects/s804294/th esis.pdf
- [4] H. Guo, J.E. Odegard, M. Lang, R.A. Gopinath, I. Selesnick, an C.S. Burrus, "Speckle reduction via wavelet shrinkage with application to SAR based ATD/R," Technical Report CML TR94- 02, CML, Rice University, Houston, 1994.
- [5] D.L. Donoho and I.M. Johnstone, "Adapting to unknown smoothness via wavelet shrinkage," Journal of the American Statistical Association, vol. 90, no. 432, pp. 1200-1224, 1995.
- [6] S.G. Chang, B. Yu, and M. Vetterli, "Adaptive wavelet thresholding for image denoising and compression," IEEE Transactions on Image Processing, vol. 9, no. 9, pp.1532-1546, September 2000.

7-8 April, 2012

Proceedings published by International Journal of Computer Applications® (IJCA)ISSN: 0975 - 8887

- [7] X.-P. Zhang, "Thresholding Neural Network for Adaptive Noise reduction," IEEE Trans. On Neural Networks, vol.12, no. 3, pp.567-584, May 2001.
- [8] I. Daubechies, Ten Lectures on Wavelets, SIAM, Philadelphia, 1992.
- [9] B.B. Hubbard, The World According to Wavelets: The Story of Mathematical Technique in the Making, A. K. Peter Wellesley, Massachusetts, 1996.
- [10] S. G. Mallat, "Multiresolution approximations and wavelet orthonormal bases of L2 (R)," Transactions of the American Mathematical Society, 315(1), pp.69-87, 1989a.
- [11] A. Grossman and J. Morlet, "Decomposition of Hardy Functions into Square Integrable Wavelets of Constant Shape," SIAM J. App Math, 15: pp.723-736, 1984.
- [12] C. Valens. (2004). A really friendly guide to wavelets. [Online]. Available: http://perso.wanadoo.fr/polyvalens/clemens/wavelets/wa velets.html
- [13] G. Kaiser, A Friendly Guide To Wavelets, Boston:Birkhauser, 1994.
- [14] I. Daubechies, "Different Perspectives on Wavelets," in Proceedings of Symposia in Applied Mathematics, vol. 47, American Mathematical Society, USA, 1993.
- [15] J. S. Walker, A Primer on Wavelets and their Scientific Applications, Chapman & Hall/CRC, New York, 1999.
- [16] E. J. Stollnitz, T.D. DeRose, and D.H. Salesin, Wavelets for Computer Graphics: Theory and Applications, Morgan Kaufmann Publishers, San Francisco, 1996.
- [17] J. Shen and G. Strang, "The zeros of the Daubechies polynomials," in Proc. American Mathematical Society, 1996.
- [18] A. K. Jain, Fundamentals of Digital Image Processing, Englewood Cliffs, New Jersey, 1989.
- [19] M. Mastriani, "Enhanced Boolean Correlation Matriz Memory", (RNL02), in Proceedings of X RPIC Reunión de Trabajo en Procesamiento de la Información y Control, San Nicolás, Buenos Aires, Argentina, October 8-10, 2003.
- [20] G. Delfino and F. Martinez. (2000, March). Watermarking insertion in digital images (spanish). [Online]. Available: http://www.internet.com.uy/fabianm/watermarking.pdf
- [21] Y. Yu, and S.T. Acton, "Speckle Reducing Anisotropic Diffusion," IEEE Trans. on Image Processing, vol. 11, no. 11, pp.1260-1270, 2002.

- [22] M. Mastriani and A. Giraldez, "Enhanced Directional Smoothing Algorithm for Edge-Preserving Smoothing of Synthetic-Aperture Radar Images," Journal of Measurement Science Review, vol 4, no. 3, pp.1-11, 2004. [Online]. Available: http://www.measurement.sk/2004/S3/Mastriani.pdf.
- [23] Southern, E.M.: Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98, 503–517 (1975)
- [24] Wang Y., Lu J., Lee R., Gu Z., Clarke R.: Iterative normalization of CDNA microarray data. IEEE Trans. Inf. Technol. Biomed. 6, 29–37 (2000)
- [25] O'Neill, P., Magoulas, G.D.: Improved processing of microarray data using image reconstruction techniques. IEEE Trans. Nanobiosci. 2(4), 176–183, December (2003)
- [26] Lukac, R., Plataniotis, K.N., Smolka, B., Venetsanopoulos, A.N.: cDNA microarray image processing using fuzzy vector filtering framework. J. Fuzzy Sets Syst. Special Issue on Fuzzy Sets and Systems in Bioinformatics 152(1), 17–35, May (2005)
- [27] Mallat, S.: A theory for multiresolution signal decomposition: the wavelet representation. IEEE Trans. Pattern Anal. Machine Intel. 11, 674–693, July (1989)
- [28] Chua,L.O.,Yang, L.:Cellular neural networks: theory. IEEETrans. Circuits Syst. 35, 1257-1272 (1988)
- [29] Zhang, X.Y., Chen, F., Zhang, Y.T., Agner, S.G., Akay, M., Lu, Z.H., Waye, M.M.Y., Tsui, S.K.W.: Signal processing techniques in genomic engineering. In: Proceedings of the IEEE90(12), 1822–1833, December (2002)
- [30] Arena, P., Bucolo, M., Fortuna, L., Occhipinty, L.: Celular neural networks for real-time DNA microarray analysis. IEEE Engl.Med. Biol. 21, 17–25 (2002)
- [31] Wang, X.H., Istepanian, R.S.H., Song, Y.H.: Microarray image enhancement by denoising using stationary wavelet transform. IEEE Trans. Nanobiosci. 2(4), 184– 189, December (2003)
- [32] Pesquet, J.C.,Krim, H.,Carfantan,H.:Time-invariant Orthonormal Wavelet Representations. IEEE Trans. Signal Process. 44, 1964–1970, August (1996)
- [33] Adjeroh, D.A., Zhang, Y., Parthe, R.: On denoising and compression of DNA microarray images. Pattern Recogn., Special Issue on Bioinformatics, pp. 2478– 2493, December (2006)

MPGI National Multi Conference 2012 (MPGINMC-2012) "Advancement in Electronics & Telecommunication Engineering" Proceedings published by International Journal of Computer Applications® (IJCA)ISSN: 0975 - 8887 7-8 April, 2012



Original

Noisy

SBON



SURE Shrink

Bayes Shrink

Normal Shrink



#### SURE Shrink

....

Bayes Shrink

......

-

## Fig. 3 Original, noisy and filtered images TABLE I ASSESSMENT PARAMETERS VS. FILTERS FOR FIGURE 2

Filter	Assessment Parameter								
	AAD	SNR	PSNR	IF	CQ	SC	FOM		
En-Frost	38.2653	3.4464	33.7364	0.7109	150.7467	0.5663	0.39857		
En-Lee	39.7437	3.3363	33.8373	0.7112	150.7472	0.5632	0.49876		
Frost	38.4374	3.2423	33.7033	0.7106	150.5244	0.5689	0.48756		
Lee	39.2427	3.4242	32.6363	0.7015	150.4141	0.5924	0.43447		
Gamma	39.6252	3.1112	33.2703	0.7063	150.1918	0.5751	0.44235		
Kuan	39.8224	3.1243	31.8272	0.7041	149.3121	0.5715	0.45342		
Median	39.5252	3.1131	32.7916	0.6852	148.9172	0.5896	0.40704		
Wiener	39.1829	3.4557	33.7033	0.7106	150.5244	0.5689	0.44236		
DS	38.7332	3.4657	33.9997	0.7169	150.9898	0.5599	0.64111		
EDS	38.1484	3.6969	34.1315	0.7182	151.5252	0.5612	0.64324		
VisuShrink (ST)	39.1450	3.4596	33.7412	0.7109	151.1527	0.5657	0.44382		
VisuShrink (HT)	38.8612	3.5283	34.4115	0.7166	151.3316	0.5666	0.44324		
VisuShrink (SST)	38.1829	3.5557	34.7033	0.7196	151.9202	0.5612	0.46432		
SureShrink	38.1612	3.5751	34.7193	0.7198	151.9244	0.5611	0.43322		
OracleShrink	38.1189	3.6957	34.7233	0.7198	151.9844	0.5619	0.45534		
BayesShrink	38.1145	3.6968	34.7237	0.7199	151.9953	0.5612	0.46329		
NormalShrink	38.1098	3.6998	34.8734	0.7199	151.9983	0.5609	0.59333		
TNN	38.1008	3.7157	34.8833	0.7199	151.9992	0.5600	0.65432		
SBON	37.7155	3.7772	36.8388	0.7353	155.4613	0.5513	0.69123		

### TABLE II ASSESSMENT PARAMETERS VS. FILTERS FOR FIGURE 3

Filter	Assessment Parameter								
	AAD	SNR	PSNR	IF	CQ	SC	FOM		
En-Frost	12.4747	290.1324	363.6712	0.9830	226.4744	0.8972	0.41265		
En-Lee	12.8474	290.2522	363.9321	0.9883	226.8373	0.8932	0.51986		
Frost	12.1847	290.2772	363.0233	0.9828	226.3272	0.8923	0.55312		
Lee	12.3733	290.2333	363.0238	0.9838	226.2822	0.8943	0.44421		
Gamma	12.3830	290.8331	363.3433	0.9882	226.8383	0.8934	0.51235		
Kuan	12.3833	290.8272	363.4923	0.9887	226.8381	0.8934	0.54129		
Median	12.9973	289.1212	361.8374	0.9673	225.9287	0.8734	0.51286		
Wiener	11.9042	290.8635	363.5568	0.9866	226.8901	0.8954	0.56413		
DS	11.4572	290.9950	363.9393	0.9898	226.9723	0.8993	0.64213		
EDS	11.5792	290.9998	363.9865	0.9899	226.9975	0.8993	0.64449		
VisuShrink (ST)	11.9055	289.2367	361.5523	0.9761	222.7564	0.8872	0.51228		
VisuShrink (HT)	11.9042	290.8673	363.5615	0.9966	226.8909	0.8976	0.56424		
VisuShrink (SST)	11.7864	290.9546	363.9822	0.9975	226.8937	0.8984	0.56389		
SureShrink	11.7074	291.0753	363.8343	0.9991	226.8942	0.8991	0.57432		
OracleShrink	11.8436	290.9332	363.7363	0.9968	226.8963	0.8983	0.55234		
BayesShrink	11.9353	290.9363	363.7361	0.9923	226.8942	0.8962	0.56328		
NormalShrink	11.6875	290.9992	363.9353	0.9992	226.9021	0.8999	0.59611		
TNN	11.4447	291.7243	363.9991	0.9994	226.9732	0.9002	0.62900		
SBON	10.9071	294.9237	383.1090	0.9992	229.8972	0.9173	0.69322		