

Variational Mode Decomposition and Multiple Feature Segmentation on Microarray Images

G.Sai Chaitanya Kumar*

PhD Research Scholar, Department of CSE, JNTU Hyderabad, India.

Dr. R.KiranKumar

Assistant Professor, Department of Computer Science, Krishna University, Machilipatnam, India.

Dr. G.Apparao Naidu

Professor, Dept of CSE, J.B. Institute of Engineering and Technology, Hyderabad, India.

Abstract

Microarray technology allows the simultaneous monitoring of thousands of genes in parallel. Based on the gene expression measurements, microarray technology have proven powerful in gene expression profiling for discovering new types of diseases and for predicting the type of a disease. Enhancement, Gridding, Segmentation and Intensity extraction are important steps in microarray image analysis. This paper presents a noise removal method in microarray images based on Variational Mode Decomposition (VMD). VMD is a signal processing method which decomposes any input signal into discrete number of sub-signals (called Variational Mode Functions) with each mode chosen to be its band width in spectral domain. First the noisy image is processed using 2-D VMD to produce 2-D VMFs. Then Discrete Wavelet Transform (DWT) thresholding technique is applied to each VMF for denoising. The denoised microarray image is reconstructed by the summation of DWT filtered VMFs and the reconstructed image is segmented using FCM

*Corresponding Author

Clustering method with each pixel having multiple features. This filtering mechanism is named as 2-D VMD-DWT thresholding method and segmentation mechanism is named as multiple feature FCM method. The proposed filtering method is compared BEMD-DWT thresholding method. The qualitative and quantitative analysis shows that 2-D VMD-DWT thresholding method produces better noise removal than BEMD-DWT thresholding method and produces better segmentation results with multiple features for each pixel than a single feature.

Keywords: Empirical Mode decomposition; Variational Mode Decomposition; Discrete Wavelet Transform; Image Enhancement; Microarray Images.

I. INTRODUCTION

The most powerful tool in molecular genetics for biomedical research is Microarray, which allows parallel analysis of the expression level of thousands of genes. The most important aspect in microarray experiment is image analysis. The output of image analysis is a matrix consisting of a measure of intensity of each spot in the image. This measure denotes gene expression ratio (transcription abundance) between the test and control samples for the corresponding gene. The positive expression indicates the over-expression, while negative expression indicates under-expression between the control and treatment genes. The main components in microarray image analysis are localization, segmentation and spot quantification [1]. This paper mainly focuses on Gaussian noise removal in microarray images. The main applications of microarray technology are Gene discovery, Drug discovery, Disease diagnosis, Toxicological research etc [2]. The microarray image analysis is shown in figure 1.

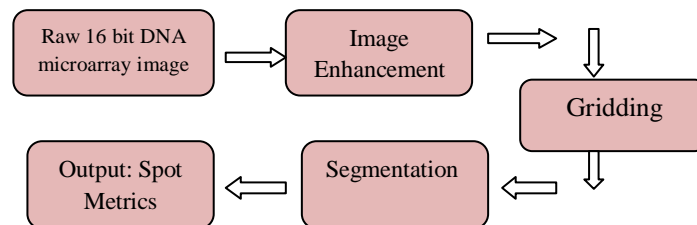


Figure 1: Microarray Image Analysis

The evaluation of microarray images is a difficult task as the fluorescence of the glass slide adds noise floor to the microarray image [3]. The processing of the microarray image requires noise suppression with minimal reduction of spot edge information that derives the segmentation process. Thus the task of microarray image

enhancement is of paramount importance [4]. This paper presents a Gaussian noise removal in microarray images using Variational Mode Decomposition. First the noisy image is decomposed into VMFs using 2-D VMD. Each VMF is denoised using DWT thresholding technique. After denoising, the image is reconstructed by combining all 2-D VMFs. After noise removal, gridding, segmentation and Expression ratio calculation are the important tasks in microarray image analysis. Any noise in the microarray image will affect the subsequent analysis [5].

Many microarray image segmentation approaches have been proposed in literature. Fixed circle segmentation [6], Adaptive circle Segmentation Technique [7], Seeded region growing methods [8] and clustering algorithms [9] are the methods that deal with microarray image segmentation problem. This paper mainly focuses on clustering algorithms. These algorithms have the advantages that they are not restricted to a particular spot size and shape, does not require an initial state of pixels and no need of post processing. These algorithms have been developed based on the information about the intensities of the pixels only (one feature). But in the microarray image segmentation problem, not only the pixel intensity, but also the distance of pixel from the center of the spot and median of intensity of a certain number of surrounding pixels influences the result of clustering. In this paper, multiple feature fuzzy c-means clustering algorithm is proposed, which utilizes more than one feature. The qualitative and quantitative results show that multiple feature fuzzy C-means clustering algorithm has segmented the image better than single feature FCM clustering algorithm.

The paper is organized as follows: Section II presents Denoising using Bi-dimensional Empirical Mode Decomposition (BEMD)-DWT thresholding method, Section III presents Denoising using 2-D VMD -DWT thresholding method, Section IV presents Multiple Feature FCM Clustering Algorithm, Section V presents Experimental results and Section VI report conclusions.

II. BI-DIMENSIONAL EMPIRICAL MODE DECOMPOSITION-DWT THRESHOLDING METHOD

Empirical mode decomposition [10] is a signal processing method that nondestructively fragments any non-linear and non-stationary signal into oscillatory functions by means of a mechanism called shifting process. These oscillatory functions are called Intrinsic Mode Functions (IMF), and each IMF satisfies two properties, (a) the number of zero crossings and extrema points should be equal or differ by one. (b) Symmetric envelopes (zero mean) interpret by local maxima and minima [11]. The signal after decomposition using EMD is non-destructive means that the original signal can be obtained by adding the IMFs and residue. The first IMF is a high frequency component and the subsequent IMFs contain from next high

frequency to the low frequency components. The shifting process used to obtain IMFs on a 2-D signal (image) is summarized as follows:

- a) Let $I(x,y)$ be a Microarray image used for EMD decomposition. Find all local maxima and local minima points in $I(x,y)$.
- b) Upper envelope $Up(x,y)$ is created by interpolating the maxima points and lower envelope $Lw(x,y)$ is created by interpolating minima points. This interpolation is carried out using cubic spline interpolation method.
- c) Compute the mean of lower and upper envelopes denoted by $Mean(x,y)$.

$$Mean(x, y) = \frac{(Up(x, y) + Lw(x, y))}{2} \quad (1)$$

- d) This mean signal is subtracted from the input signal.

$$Sub(x, y) = I(x, y) - Mean(x, y) \quad (2)$$

- e) If $Sub(x,y)$ satisfies the IMF properties, then an IMF is obtained .

$$IMF_i(x, y) = Sub(x, y) \quad (3)$$

- f) Subtract the extracted IMF from the input signal. Now the value of $I(x,y)$ is

$$I(x, y) = I(x, y) - IMF_i(x, y) \quad (4)$$

Repeat the above steps (b) to (f) for the generation of next IMFs.

- g) This process is repeated until $I(x,y)$ does not have maxima or minima points to create envelopes.

Original Image can be reconstructed by inverse EMD given by

$$I(x, y) = \sum_{i=1}^n IMF_i(x, y) + res(x, y) \quad (5)$$

The mechanism of de-noising using BEMD-DWT is summarized as follows

- a) Apply 2-D EMD for noisy microarray to obtain IMF_i ($i=1, 2, \dots, k$). The k th IMF is called residue.
- b) The first intrinsic mode function (IMF1) contains high frequency components and it is suitable for denoising. This IMF1 is denoised with DWT Thresholding technique [12]. This de-noised IMF1 is represented with DNIMF1.
- c) The denoised image is reconstructed by the summation of FIMF and remaining IMFs given by

$$RI = DNIMF1 + \sum_{i=2}^k IMF_i \quad (6)$$

Where RI is the reconstructed band. The flow diagram of BEMD-DWT filtering is shown in figure 2.

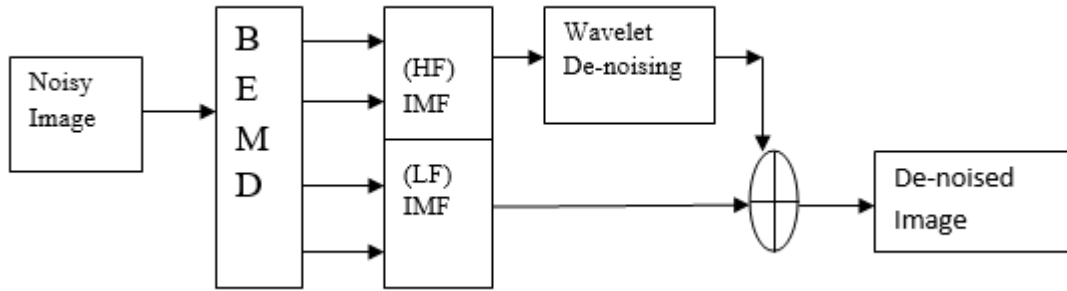


Figure 2: Flow Diagram of BEMD-DWT filtering method

III. VARIATIONAL MODE DECOMPOSITION- DWT THRESHONDING METHOD

In the Variational Mode Decomposition (VMD) framework, the signal is decomposed into k mode components, where each mode u_k is required to be mostly compact around a center pulsation ω_k determined along with the decomposition [13]. For a one-dimensional Signal s, the algorithm to assess the bandwidth of a mode u_k is as follows:

- (a) Compute the analytic signal associated with u_k by means of the Hilbert transform to obtain a unilateral frequency spectrum.
- (b) Shift the mode’s frequency spectrum to baseband by mixing with an exponential tuned to the respective estimated center frequency.
- (c) Estimate the bandwidth through the H1 Gaussian smoothness of the demodulated signal; for example, the squared L2-norm of the gradient.

Then, the constrained variational problem to find all the modes is defined by

$$\min_{u_k, \omega_k} = \left\{ \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|_2 \right\}$$

Subject to:

$$\sum_k u_k = s \tag{7}$$

where s stands for the signal to decompose, u_k is the k^{th} mode, ω_k is a frequency, δ is the Dirac distribution, t is a time, and $*$ denotes convolution. Higher values of k indicate modes with lower frequency components. The same procedure will be done for 2-D VMD.

The mechanism of de-noising using 2-D VMD-DWT is summarized as follows

- Apply 2-D VMD for noisy microarray to obtain VMFi ($i=1, 2, \dots, k$). The k th VMF is called residue.
- The VMFs are denoised with DWT Thresholding technique. This de-noised VMF is represented with DNVMF.
- The denoised image is reconstructed by the summation of VMFs given by

$$RI = \sum_{i=1}^k VMF_i \quad (8)$$

Where RI is the reconstructed image. The flow diagram of 2-D VMD-DWT filtering is shown in figure 3.

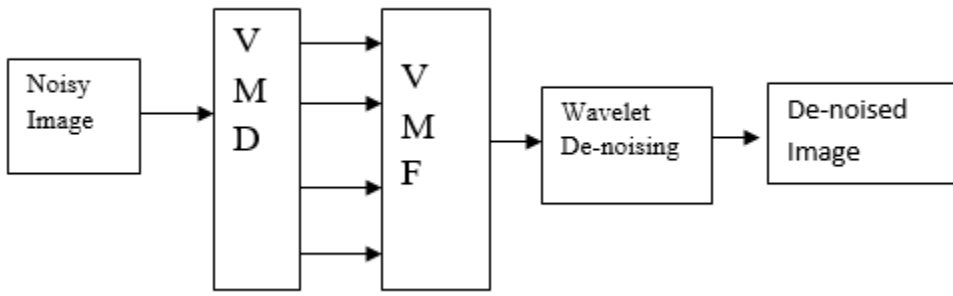


Figure 3: Flow diagram of 2-D VMD-DWT filtering method

IV. MULTIPLE FEATURE FUZZY C-MEANS CLUSTERING ALGORITHM

The FCM algorithm for segmentation of microarray image is described below [14]:

- Take randomly K initial clusters from the $m \times n$ image pixels.
- Initialize membership matrix u_{ij} with value in range 0 to 1 and value of $m=2$.

Assign each pixel to the cluster C_j $\{j=1,2,\dots,K\}$ if it satisfies the following condition $[D(\cdot, \cdot)]$ is the Euclidean distance measure between two values.

$$u_{ij}^m D(I_i, C_j) < u_{iq}^m D(I_i, C_q), q = 1, 2, \dots, K$$

$$j \neq q \quad (9)$$

The new membership and cluster centroid values as calculated as [15]

$$u_{ik} = \frac{1}{\sum_{j=1}^K \left(\frac{D(C_i, I_k)}{D(C_j, I_k)} \right)^{\frac{1}{m-1}}}, \text{ for } 1 \leq i \leq K$$

$$C_j^{\wedge} = \frac{\sum_{j=1}^n u_{ij}^m I_j}{\sum_{j=1}^n u_{ij}^m} \quad (10)$$

3. Continue 2-3 until each pixel is assigned to the maximum membership cluster [16].

The clustering algorithms used for microarray image segmentation are based on the information about the intensities of the pixels only. But in microarray image segmentation, the position of the pixel and median value of surrounding pixels also influences the result of clustering and subsequently that leads to segmentation. Based on this observation, multiple feature clustering algorithm is developed for segmentation of microarray images. To apply fuzzy c-means clustering algorithm on a single spot, we take all the pixels that are contained in the spot are, which is obtained after gridding process, and create a dataset $D = \{x_1, x_2, x_3, x_4, x_5, \dots, x_n\}$, where $x_i = [x_i^{(1)}, x_i^{(2)}, x_i^{(3)}]$ is a three dimensional vector that represents the i th pixel in the spot region. We use three features, defined as follows

$x_i^{(1)}$: Represents the pixel intensity value.

$x_i^{(2)}$: Represents the distance from pixel to the center of the spot region.

The spot center is calculated as follows:

1. Apply edge detection to the spot region image using canny method.
2. Perform flood-fill operation on the edge image using imfill method.
3. Obtain label matrix that contain labels for the 8-connected objects using bwlabel function.
4. Calculate the centroid of each labeled region (connected component) using region props method.

$x_i^{(3)}$: Represents the median of the intensity of surrounding pixels.

For each pixel in the spot region, once the features are obtained forming the dataset D , then the fuzzy c-means clustering algorithm is applied.

V. EXPERIMENTAL RESULTS

In this section, the proposed filtering method is performed on a sample microarray slide drawn from the standard microarray database corresponds to breast category aCGH tumor tissue that has 48 blocks, each block consisting of 110 spots. A sample block has been chosen and 36 spots of the block have been cropped for simplicity. The sample image is a 198*196 pixel image that consists of a total of 38808 pixels. The RGB colored image microarray image have been converted to grayscale image to specify a single intensity value that varies from the darkest (0) to the brightest (255) for each pixel. To check the performance of the filtering methods presented in this paper Gaussian noise with different values of σ is added to the image. The noisy

image is filtered using methods presented in this paper. The quantitative results are evaluated using PSNR values are shown in table 1 and qualitative results are shown in figure 4. Both qualitative and quantitative results show that 2-D VMD-DWT filtering method has better denoising effect. After denoising the image is segmented using multiple feature FCM clustering algorithm.

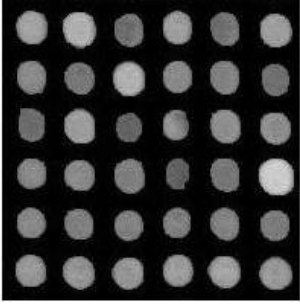
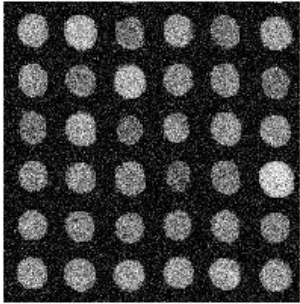
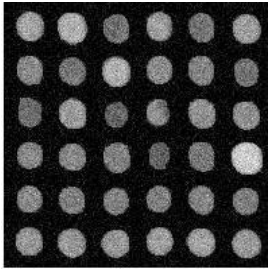
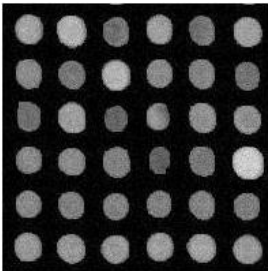
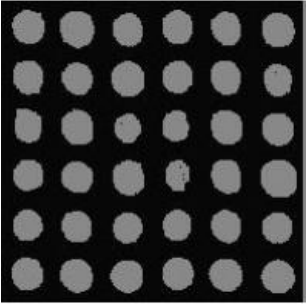
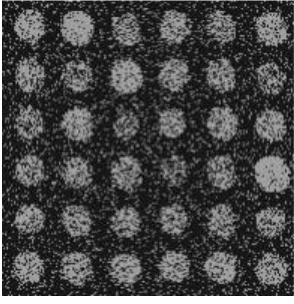
Original Image	Gaussian Noise
	
Filtering Using BEMD+DWT	Filtering Using 2-D VMD+DWT
	
Segmentation using Multiple feature FCM on filtered image	Segmentation using FCM on noisy image
	

Figure 4: Qualitative analysis of filtering method

Quantitative Analysis: Quantitative analysis is a numerically oriented procedure to figure out the performance of algorithms without any human error. The Mean Square Error (MSE) [17] is significant metric to validate the quality of image. It measures the square error between pixels of the original and the resultant images. The MSE is mathematically defined as

$$MSE = \frac{1}{N} \sum_{j=1}^k \sum_{i \in C_j} \|v_i - c_j\|^2 \tag{11}$$

Where N is the total number of pixels in an image and xi is the pixel which belongs to the jth cluster. The lower difference between the resultant and the original image reflects that all the data in the region are located near to its centre. Table 2 shows the quantitative evaluation of clustering algorithm. The results confirm that multiple feature fuzzy c-means algorithm produces the lowest MSE value for segmenting the microarray image.

TABLE 1: PSNR values of filtering method

Method	$\sigma = 0.015$	$\sigma = 0.025$	$\sigma = 0.036$
BEMD + Wavelet (Universal Threshold)	34.11	29.86	23.86
BEMD + Wavelet (SURE shrink)	32.89	28.74	22.11
2-D VMD+ Wavelet (Universal Shrink)	36.16	30.98	24.43
2-D VMD+ Wavelet (Sure Shrink)	38.91	32.34	26.98

Table 2: MSE values

Method	MSE Values	MSE Values
K-means	282.781	346.47
Fuzzy c-means (Single Feature)	216.392	228.69
Multiple feature Fuzzy C-means	198.327	186.276

VI. CONCLUSIONS

Microarray technology is used for parallel analysis of gene expression ratio of different genes in a single experiment. The analysis of microarray image is done with gridding, segmentation and information extraction. The expression ratio of each and every gene spot denotes the transcription abundance between two genes under experiment. This paper presents a new method for microarray image denoising and segmentation. Clustering algorithms have been used for microarray image segmentation with an advantage that they are not restricted to a particular shape and size for the spots. Instead of using single feature clustering algorithm, this paper presents multiple feature clustering algorithm with three features for each pixel such as pixel intensity, distance from the center of the spot and median of surrounding pixels. Spot information includes the calculation of Expression Ratio in the region of every gene spot on the microarray image. The expression-ratio measures the transcription abundance between the two sample genes. The proposed method performs better noise suppression and producing better segmentation results.

REFERENCES

- [1] J.Harikiran, Dr.P.V.Lakshmi, Dr.R.Kirankumar, "Automatic Gridding Method for microarray images", *Journal of Theoretical and Applied Information Technology*, volume 65, Number 1, pp.235-241, 2014.
- [2] J.Harikiran, Dr.P.V.Lakshmi, Dr.R.Kiran Kumar, "Fast Clustering Algorithms for Segmentation of Microarray Images", *International Journal of Scientific & Engineering Research*, Volume 5, Issue 10, pp 569-574, 2014.
- [3] J.Harikiran, et.al. "Vector Filtering Techniques for Impulse Noise Reduction with Application to Microarray images", *International Journal of Applied Engineering Research*", volume 10, Number 3, pp. 7181-7193, 2015.
- [4] Emmanouil i.Athanasiadis, et.al."Complementary DNA Microarray Image Processing Based on the Fuzzy Gaussian Mixture Model", *IEEE Transactions on Information Technology in Biomedicine*, Volume 13, issue 4, pp.419-425, 2009.
- [5] J.Harikiran, et.al. "Denoising based clustering algorithm for segmentation of Microarray Image", *International Journal of Electronics Communication and Computer Engineering*, Vol. 3, Issue 6, pp-1400-1406, 2012.
- [6] M.Eisen, *ScanAlyze User's manual*, 1999,
- [7] J.Buhler, T.Ideker and D.Haynor, "Dapple:Improved Techniques for Finding spots on DMA Microarray Images", *Tech. Rep. UWTR 2000-08-05*, University of Washington, 2000.
- [8] R.Adams and L.Bischof, "Seeded Region Growing", *IEEE Transactions on*

Pattern Analysis and Machine Intelligence, Vol 16,no. 6, pp.641-647, 1994.

- [9] D.RamaKrishna, J.Harikiran et.al.” “Various Versions of K-means Clustering Algorithm for Segmentation of Microarray Image”, *International Journal of Electronics Communication and Computer Engineering*, Volume 4, Issue 1, pp.1554-1558, 2012.
- [10] J.Harikiran, et.al. “Multiple Feature Fuzzy C-means Clustering Algorithm for Segmentation of Microarray image”, *IAES International Journal of Electrical and Computer Engineering*”, Vol. 5, No. 5, pp. 1045-1053, 2015.
- [11] B.Saichandana et.al. “Image Fusion in Hyperspectral Image Classification using Genetic Algorithm”, *IAES Indonesian Journal of Electrical Engineering and Computer Science*, Volume 2, No. 3, June 2016, pp.703-711.
- [12] D.Ramakrishna, J.Harikiran et.al. “Noise Removal using Empirical Mode Decomposition and Wavelet Transform in Microarray Images”, *International journal of Computer Applications*, Volume 68, No 2, April 2013, pp.1-7.
- [13] Konstantin Dragomiretskiy, Dominique Zosso, “Variational Mode Decomposition”, *IEEE Transactions on Signal Processing*, Volume 62, Number 3, February 2014, pp. 531-544.
- [14] Dr.R.Kiran Kumar et.al. “Dimensionality Reduction and Classification of Hyperspectral Images using Genetic Algorithm”, *IAES Indonesian Journal of Electrical Engineering and Computer Science*, Volume 3, No 3, pp.503-511, September 2016.
- [15] B.Saichandana, J.Harikiran, Dr.K.Srinivas and Dr.R.Kiran Kumar, “Application of BEMD and Hierarchical Image Fusion in Hyper spectral Image Classification”, *International Journal of Computer Science and Information Security*, volume 14, number 5, May 2016, pp.437-445.
- [16] J.Harikiran et.al. “Fuzzy C-means with Bi-dimensional empirical Mode decomposition for segmentation of Microarray Image”, *International Journal of Computer Science Issues*, volume 9, Issue 5, Number 3, pp.273-279, 2012.
- [17] B.Saichandana, et.al, “Clustering Algorithm combined with Hill climbing for Classification of Remote Sensing Image”, *IAES International Journal of Electrical and Computer Engineering*, volume 4, No 6, December 2014, pp.923-930.

