A Fully Automated Spot Detection Approach for cDNA Microarray Images Using Adaptive Thresholds and Multi-Resolution Analysis

MARY MONIR SAEID[ID][1]**, ZAKI B. NOSSAIR**[2]**, AND MOHAMED ALI SALEH**[2]
[1]Department of Information Systems, Faculty of Computer and Information, Fayoum University, Fayoum 63514, Egypt
[2]Department of Electronics, Communications, and Computer Engineering, Faculty of Engineering, Helwan University, Helwan 11795, Egypt

Corresponding author: Mary Monir Saeid (mmh04@fayoum.edu.eg)

ABSTRACT The problem of gridding microarray images remains a challenging task. This is because microarray images are usually contaminated with noise and artifacts, such as low intensity and poor quality spots. In this paper, a new gridding technique for microarray images is introduced. The proposed technique includes both global gridding (sub-array detection) and local gridding (individual spot detection). Our technique is developed based on multi-resolution analysis and a new adaptive threshold method. The proposed framework is fully automated in the sense that it does not need any user intervention and the only input required is the microarray image. The presented technique can be applied to images with different specifications, such as resolution, number of sub-arrays, number of spots in each sub-array, and noise levels. The experimental results show that the proposed method is highly accurate when compared with the existing software tools as well as with recently published techniques. Our results also show that the presented approach is very effective for gridding microarray images with low intensity, poor quality spots, and missing/irregular spots. The spot detection accuracy of the proposed method is improved by up to 5.48% compared with that of the other published algorithms.

INDEX TERMS cDNA microarray images, fully automatic gridding, sub-array detection, spot detection, gene expression.

I. INTRODUCTION

Deoxyribonucleic acid (DNA) microarray technology is a powerful tool for evaluating the expression levels for several thousands of genes simultaneously. Microarray technology has been widely used in many applications such as genetic research, understanding and diagnosis of diseases, and pharmacology research. A DNA microarray is a tiny glass slide on which known DNA sequences or genes in solution have been robotically spotted in a rectangular or square grid; one gene per spot. Often, these slides are referred to as gene chips or DNA chips.

DNA microarrays are prepared by obtaining a control tissue sample and a test tissue sample. Then, from each sample, mRNA sequences are extracted and labeled with different fluorescent dye. Cy3 (green) is usually used for labeling the mRNA sequences of the control sample and Cy5 (red) is used for labeling the mRNA sequences of the test sample. The labeled mRNA sequences of the two samples are hybridized to the array simultaneously. The array is then washed to remove the mRNA sequences that are not hybridized. After washing, a laser scanner is used to capture the microarray image [1]. Then, fluorescent measurements of red and green channels will return the expression level for every gene represented by a spot in the microarray image. Fig. 1 illustrates the preparation steps for a microarray slide and the acquisition of its microarray image. As it can be seen from Fig.1, a microarray image is usually composed of rectangular or square areas called sub-arrays; each sub-array contains a number of spots. Each spot represents the expression level of a single gene.

Processing of microarray images consists of three steps: gridding, segmentation and quantification [2]. Gridding is the first main step in extracting data from a microarray image. It is the process of identifying the area of each sub-array within an image (global gridding) and then identifying the area of each spot within each sub-array (local gridding). Segmentation is the process that separates spot pixels from background pixels within each spot area. Quantification, which is

---

The associate editor coordinating the review of this manuscript and approving it for publication was Prakasam Periasamy.
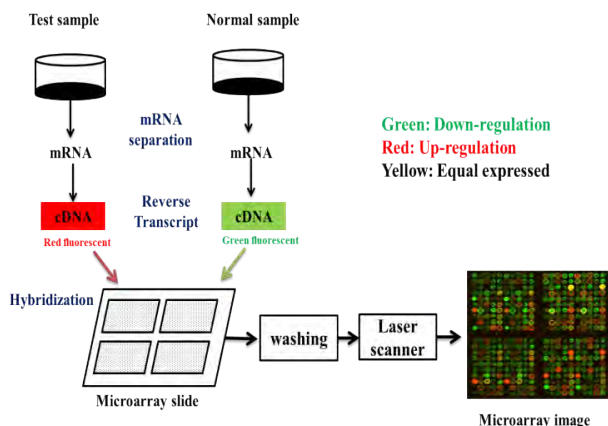
FIGURE 1. The preparation steps for a cDNA microarray slide and the acquisition of its microarray image.

FIGURE 2. The right side image is a typical microarray image containing 16 sub-arrays arranged in a grid of 4 × 4 sub-arrays. Each sub-array has 8 × 10 features (spots). Also the right side image shows some image rotation problem and the left side image shows some spot problems within the sub- array.

the last step, is to calculate intensity values for the foreground and background pixels for each spot of the microarray image. These intensity values are used to estimate the expression levels of the genes represented in the microarray image.

Gridding (also known as addressing) is the most fundamental and important step in microarray image analysis. Accurate gridding helps substantially in improving the efficiency of the segmentation and quantification steps. To extract each spot from the microarray image, first the image has to be divided into sub-arrays (global gridding) and each sub-array has to be separated into spot regions (local gridding).

There are several problems with microarray images that make the gridding process a difficult task. First, Even though microarrays arrange spots on a relatively regular grid, in the real situation, the exact location of the grid may vary due to mechanical constraints in the spotting process and hybridization inconsistencies [3]. Second, very often microarray images have some tilt or rotation that occurs during printing or scanning. Third, the spots of a microarray image can vary in size, shape and position due to noise in the sample preparation and hybridization processes. In addition, spot intensity levels are highly variable since different genes may express differently and weak spots are often difficult to detect. Fig. 2 illustrates some of these microarray image problems.

These problems make processing of microarray images is often hard to automate. Thus, there are three different gridding methodologies according to the degree of human intervention. These methods are manual gridding, semiautomatic gridding and fully automated gridding. These techniques are described in the following subsections.

A. MANUAL GRIDDING
This is the first method used for gridding microarray images. In this method, all the parameters required for gridding are provided manually. These parameters are the number of sub-arrays in the microarray image, the number of spots in each sub-array, and the spot size.
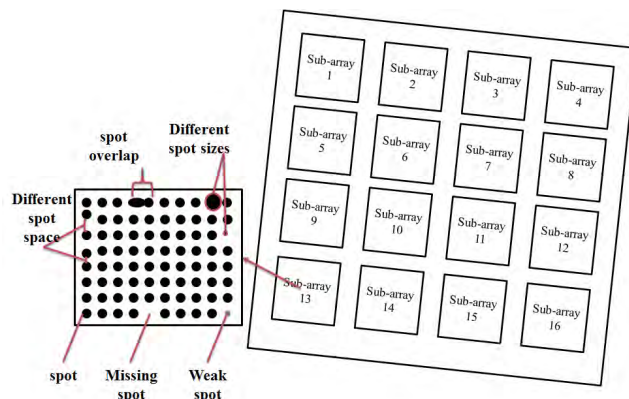
There are several software packages such as ScanAlyze [4], Spot Finder [5], ImaGene [6] and GenePix [7] that are used for manual gridding of microarray images. These systems provide Graphical User Interface (GUI) tools for assisting the user to manually provide the number of sub-arrays, the number of spots in each sub-array, and the spot size. For example, in GenePix, the user must assign the layout of blocks, the number of spots in each block, and the distances of spot centers from adjacent rows and columns. The approach in [8] presents a method for gridding microarray images based on axis projections. This method requires user intervention in order to manually adjust the locations of the grids.

In general, the advantage of this method is that it can provide 'perfect' grid alignment. The disadvantage of this method is introducing considerable inaccuracy due to human errors, particularly with arrays having irregular spacing between the spots and large variation in spot sizes. In addition, the method is very time consuming and is not applicable for high density microarray images [9].

B. SEMIAUTOMATIC GRIDDING
This gridding method reduces the human intervention in the sense that it requires from the user to partially provide some of the gridding parameters in order to achieve correctness of gridding results. For example, the approach in [10] performs spot detection based on Markov random field (MRF) method and some heuristic criteria. This method requires the number of rows and the number of columns per sub-grid to be manually provided. The approach in [11] describes a semi-automatic system which mainly focuses on the problem of finding individual spots with high accuracy. In general, semi-automatic gridding methods might not be sufficient to meet the requirements for high throughput processing of microarray images [9].

C. FULLY AUTOMATIC GRIDDING
Automatic gridding algorithms identify all spots of microarray images without any human intervention. Many automatic

gridding approaches are proposed for sub-array gridding and spot detection. For example, the approach in [12] describes a fully automatic gridding methodology using intensity projection profiles of microarray images. This gridding method is sensitive to contaminations and large number of missing spots.

The approach in [13] proposes automatic gridding of DNA microarray images using optimum sub-array. The approach is based on the selection of an optimum sub-image and then the intensity projection profiles of this optimum sub-image are used to calculate the parameters for gridding. On the other hand, the approach in [14] describes a global gridding methodology in which Radon transform is used for detecting and correcting rotations and then the algorithm applies morphological operators to separate the sub-arrays in a cDNA microarray image. Another method in [15] describes an approach for cDNA microarray gridding which is based on a hierarchical refinement algorithm. The approach applies rotation correction and local fine-tuning before performing the gridding process.

The approach in [16] presents a fully automatic method for microarray gridding. The method first finds the locations of sub-arrays of the entire image and then determines the co-ordinates of spots in each sub-array. The method detects and corrects rotation in the image by applying affine transformation.

The approach in [17] is an automatic gridding method that applies a genetic algorithm (GA) to accurately determine the line segments constituting borders between adjacent blocks or spots.

Many approaches are developed for local gridding which basically assume that the sub-grids are already identified such as the method in [18]. This method presents an automatic approach for finding spot locations. The method applies a hill-climbing approach to maximize the intensities of the spots. One or more starting pixels must be predefined that will affect the final result.

The gridding approach in [19] uses improved Otsu method to obtain an optimal threshold that separates spot pixels from background pixels. In this method, the grid lines are optimized by a heuristic technique with the help of estimating the distribution of the spots. The approach in [20] applies Otsu method optimized by the multilevel thresholds to achieve accurate microarray gridding.

The approach in [21] also describes an automatic gridding technique based on intensity projection profiles for the special-domain microarray images. The method first applies Radon transform to correct any rotation in the microarray images. Also, as a preprocessing step, a histogram based threshold method is used to enhance the intensity of the image spots. Finally, a refinement technique is applied to detect and correct grid line errors.

Many of the above mentioned approaches perform gridding for microarray images in the spatial domain. A preprocessing step is usually applied in these algorithms in order to reduce the effect of noise. In particular, there are many filters which are usually used to reduce noise such as mean, median, and morphological filters. Mean filters introduce blurring to images when are used for reducing noise. The median filter is a nonlinear filter which can remove noise without blurring sharp edges, but it will eliminate some useful spot information. The morphological filter can achieve fairly good noise reduction but the design of the structure element depends on actual image contents.

The problem of automatic gridding is complicated by the fact that microarray images are usually highly contaminated with noise and artifacts due to the wet lab processes. Some of the artifacts, which often occur in microarray images and contribute in complicating the gridding process, are rotation, misalignment and local deformation of the ideal rectangular grid. Therefore, there is a high need for development of gridding methods which are accurate and robust against these problems.

In this paper, a new fully automatic gridding technique is introduced. The presented technique is developed based on multi-resolution analysis and an adaptive threshold method.

The main reason of using multi-resolution analysis is that huge amount of image data will be eliminated because the coarse-scale sub-band of the microarray image is the only sub-band that will be used in performing global and local gridding. Also the use of multi-resolution analysis will eliminate the need for pre-processing operations for noise removal. Especially, discrete wavelet transform plays an important role in reducing noise from images and it has been shown to be more effective than filtering methods [22]–[24]. Furthermore, the advantage of using the wavelet transform is that it allows us to perform multiple decompositions in order to distinguish the distance between sub-arrays and spots. That is, the multi-resolution capability of the DWT makes it possible to detect features at a resolution that may go undetected at another. DWT can often compress or de-noise a signal without appreciable degradation.

The proposed adaptive threshold method will be applied in local gridding because the projection profiles of the microarray images are often non-uniform due to some spots may be missed or the spots may have low intensity and poor quality [25], [20].

This paper is organized as follows. Introduction is given in Section 1. Section 2 describes the different datasets (materials) that are used to test the proposed gridding technique. The proposed technique is presented in Section 3. Experimental results are given in Section 4. Conclusions are provided in Section 5.

II. DATASETS
Four different databases of cDNA microarray images are used for testing the proposed gridding method. The images are selected from different sources to have different spot sizes, shapes, and scanning resolutions, in order to accurately evaluate the efficiency of the proposed method.

The first dataset consists of 14 microarray images selected from Computational Cancer Genomics (CCG) group of the

Swiss Institute of Bioinformatics (SIB) [26]. The images selected are those with IDs from 661 to 667. Each experiment ID includes two channels (Cy3 and Cy5). The microarray images are stored in TIFF format, with image resolution of 1000×1000 pixels, and spot resolution of 18×18 pixels. Each image has 4 sub arrays and each sub array has different number of spots. The number of spots in each sub-array is ranging from 35 to 49 spots.

The second dataset of images is drawn from the University of California; San Francisco [27]. The microarray images are stored in TIFF format, with image resolution of 1512×1488 pixels, and spot resolution of 8×8 pixels. This dataset contains two images. Each image has 36 sub-grids arranged in 6 rows and 6 columns. Each sub-grid has 210 spots.

The third dataset consists of images which are taken from Gene Expression Omnibus [28]. Thirteen images are selected for testing the proposed method. These images correspond to channels 1 and 2 for experiments IDs GSM15898, GSM16101, GSM16389, GSM16391 and experiments IDs GSM17137, GSM17163, GSM17186, GSM17190 and GSM17192. The resolution of these images ranges from 1802×1942 to 5997×2200 pixels. The images are in TIFF format. The spot resolution is 12×12 pixels. Also each image contains 48 sub-grids, arranged in 12 rows and 4 columns; each sub array has 182 spots.

The fourth dataset of images are drawn from [29]. Fourteen images were selected for testing the proposed method. These images correspond to channels 1 and 2 for experiments IDs 1302, 1303, 1309, 1310, 1311, 1312, and 1313. The images have resolution of 1024×1024 pixels and are in TIFF format. The spot resolution is 8×8 pixels. Each image contains four sub arrays and each sub array contains 1,600 spots.

III. THE PROPOSED GRIDDING TECHNIQUE

In this section, the proposed fully automatic method for gridding microarray images is presented. The method is capable of processing microarray images without any user intervention (does not require any input parameters).

The proposed method consists of two major steps: (1) global gridding, and (2) local gridding. These two steps are presented in the following two subsections.

A. THE PROPOSED GLOBAL GRIDDING APPROACH

The main idea of the proposed global gridding method is to decompose the microarray image, using Haar DWT, into three decomposition levels and then use the coarse scale sub-band of the third decomposition level to define the sub-arrays within the microarray image. The reason for performing three decomposition levels is to ensure that the distance between spots becomes very small such that it will not appear in the projection profile trajectories of the microarray image specially after smoothing the trajectories. However, the distance between sub-arrays will stay well pronounced in the trajectories even after smoothing. This will make the outer sides of sub-arrays are very clear in the trajectories and

therefore the grid lines between sub-arrays can be accurately determined.

The detailed steps of our proposed automatic sub-gridding algorithm are shown in Fig. 3 and are explained as follows:
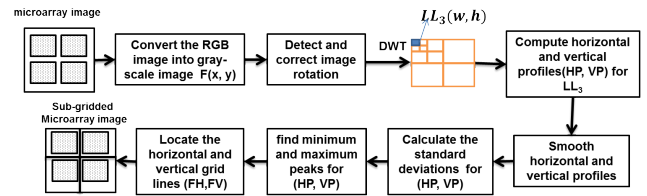

FIGURE 3. A block diagram showing the steps of our global gridding algorithm.

Step 1: Convert the RGB microarray image into a gray-scale image. The gray-scale image will be referred to as $F(x, y)$ and it will be considered of size $M \times N$ pixels.

Step 2: Detect and correct any tilt in the image using Radon transform method presented in [21].

Step 3: Decompose the microarray image, $F(x, y)$, using Haar DWT into three levels. Fig. 4 shows the coarse scale sub-band of the third decomposition level for a microarray image taken from the third dataset. As it can be observed from the coarse scale sub-band, the spots within each sub-array are almost merged together. That is, the distances between spots become very small compared to distances between sub-arrays. The coarse scale sub-band of the third decomposition level will be referred to as $LL_3(w, h)$.

FIGURE 4. The coarse scale sub-band after the third decomposition level for a microarray image taken from the third dataset.

Step 4: Calculate the horizontal projection profile (HP) and the vertical projection profile (VP) for LL_3 as follows:

$$HP(w) = \sum_{h=0}^{h=H-1} LL_3(w, h)$$

$$\text{where, } H = \frac{N}{8}, \quad w = \left(0, 1, \ldots \ldots, \frac{M}{8} - 1\right) \quad (1)$$

$$VP(h) = \sum_{w=0}^{W=W-1} LL_3(w, h)$$

$$\text{where, } W = \frac{M}{8}, \quad h = \left(0, 1, \ldots \ldots, \frac{N}{8} - 1\right) \quad (2)$$

Step 5: Perform smoothing for HP and VP profiles using a three-point moving average filter. The smoothed HP and VP are shown in Fig. 5.
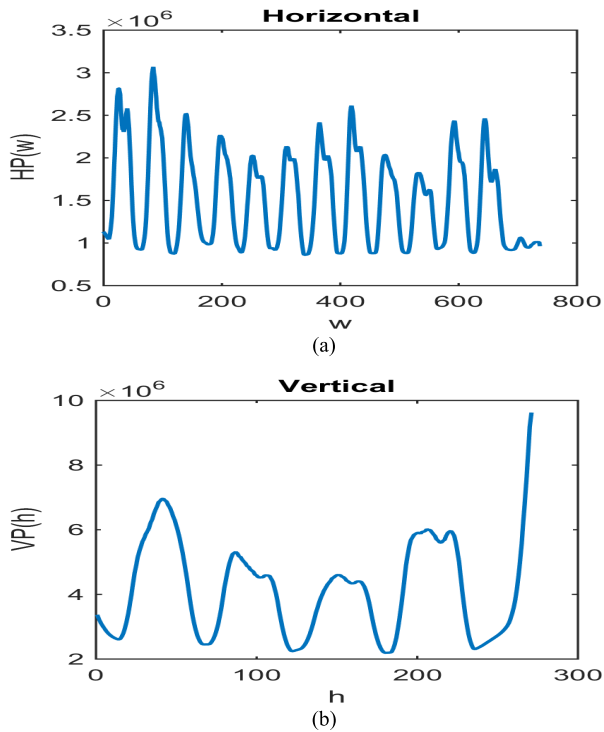

FIGURE 5. (a) The horizontal projection profile of the LL3 sub-band after smoothing. (b) The vertical projection profile of the LL3 sub-band after smoothing.

Step 6: Calculate the standard deviations for HP and VP profiles.

Step 7: Find the minimum and maximum peaks in the projection profiles HP and VP using the algorithm given in [30]. This algorithm implements a simple idea which is that the local maxima are detected between valleys. Thus, the algorithm looks for the highest point, around which there are points lower by a "peak threshold" on both sides. However, the proposed global gridding algorithm uses the standard deviations of HP and VP as peak thresholds in place of the delta value used in [30]. The standard deviation is used as the peak threshold because it is a measure for the variations of the projection profile around its mean. Therefore, small peaks which arise due to noise will not be detected. As it can be seen from Fig. 6, this method is able to select the proper peaks and ignore the minor peaks.

Step 8: Allocate the vertical and horizontal grid lines at the minimum points which are determined in step 7. The locations of the horizontal grid lines and the locations of the vertical grid lines will be denoted by H_L and V_L respectively.

Step 9: Determine the locations of the horizontal and vertical grid lines for the original microarray image (the special domain microarray image) as following:
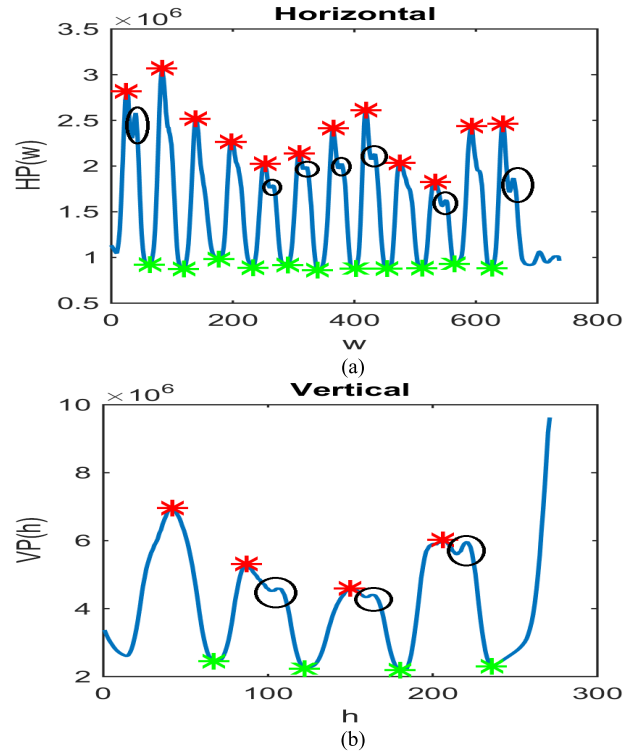
$$FH = H_L^* 8 \qquad (3)$$

FIGURE 6. (a) The minimum and maximum peaks in the horizontal projection profile. (b) The minimum and maximum peaks in the vertical projection profile.

$$FV = V_L^* 8 \qquad (4)$$

Step 10: Map these grid lines FH and FV onto the original microarray image for producing sub-gridding. Fig. 7 shows a microarray image from the third dataset (GEO) after performing sub-array detection using the proposed global gridding approach.
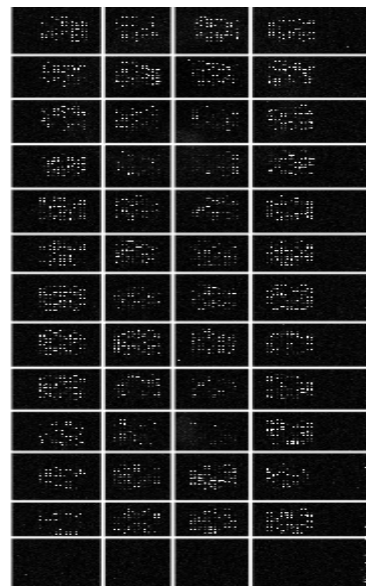

FIGURE 7. A microarray image taken from the third dataset (GEO) after performing sub-array detection using the proposed global gridding approach.

B. THE PROPOSED LOCALGRIDDING APPROACH

As it has been stated previously, a microarray image usually contains a number of sub-arrays where each sub-array is a rectangular block of spots. The sub-array detection is performed by the method presented in the previous section. After sub-array detection, the aim is to separate each sub-array into spot regions (local gridding). The proposed local gridding method proceeds as follows: Each sub-array image is decomposed using Haar DWT into one decomposition level to remove the noise but to keep the spacing between spots. Then, the horizontal and vertical projection profiles are computed and a local adaptive threshold technique is applied to each sub-array to determine the grid lines between spots. This local adaptive threshold method is used because applying one global threshold is not always suitable for local gridding. The adaptive local threshold method performs better than one global threshold because some maximum peaks might come lower than the global threshold and also some minimum peaks might come larger than the global threshold. This often occurs because some spots may be missed or the spots may have low intensity and poor quality, therefore the projection profiles often become non-uniform. In our local threshold method, the value of the threshold is calculated based on the mean of the local region. Fig. 8 shows a block diagram for the steps of the proposed local gridding method.
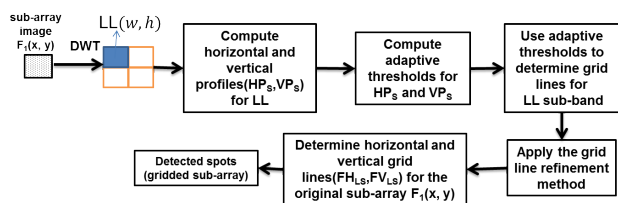
FIGURE 8. A block diagram showing the steps of our local gridding algorithm.

The steps of the local gridding algorithm are explained as follows:

Step 1: Decompose the sub-array image using Haar DWT into one decomposition level. The sub-array image will be referred to as $F_1(x, y)$ and is assumed to be of size $M_1 \times N_1$.

Step 2: Calculate the horizontal projection profile (HP_S) and vertical projection profile (VP_S) for the coarse scale sub-band (LL) of $F_1(x, y)$; these profiles are calculated as follows:

$$HP_S(w) = \sum_{h=0}^{h=H-1} LL(w, h)$$

$$where, H = \frac{N_1}{2}, \quad w = \left(0, 1, \ldots\ldots, \frac{M_1}{2} - 1\right) \quad (5)$$

$$VP_S(h) = \sum_{w=0}^{w=W-1} LL(w, h)$$

$$where, W = \frac{M_1}{2}, \quad h = \left(0, 1, \ldots\ldots, \frac{N_1}{2} - 1\right) \quad (6)$$

Step 3: Compute an adaptive threshold value for the horizontal projection profile (also an adaptive threshold value for the vertical projection profile is computed the same way) as follows:

First, compute the mean value ($\overline{m}$) for the horizontal projection profile as shown in Fig. 9(a) (Fig. 10(a) for vertical projection profile). Second, segment the horizontal projection profile into two regions (R_1, R_2) as follows:

$$R_1 = HP_S(w), \quad w < floor\left(\frac{M_1}{4}\right) \quad (7)$$

$$R_2 = HP_S(w), \quad w \geq floor\left(\frac{M_1}{4}\right) \quad (8)$$

and then compute a mean value ($\overline{m_1}$) for R_1 and a mean value ($\overline{m_2}$) for R_2 as shown in Fig. 9(b) (Fig. 10(b) for vertical). If $\overline{m_1} > k^*$, where k is a constant that has been determined
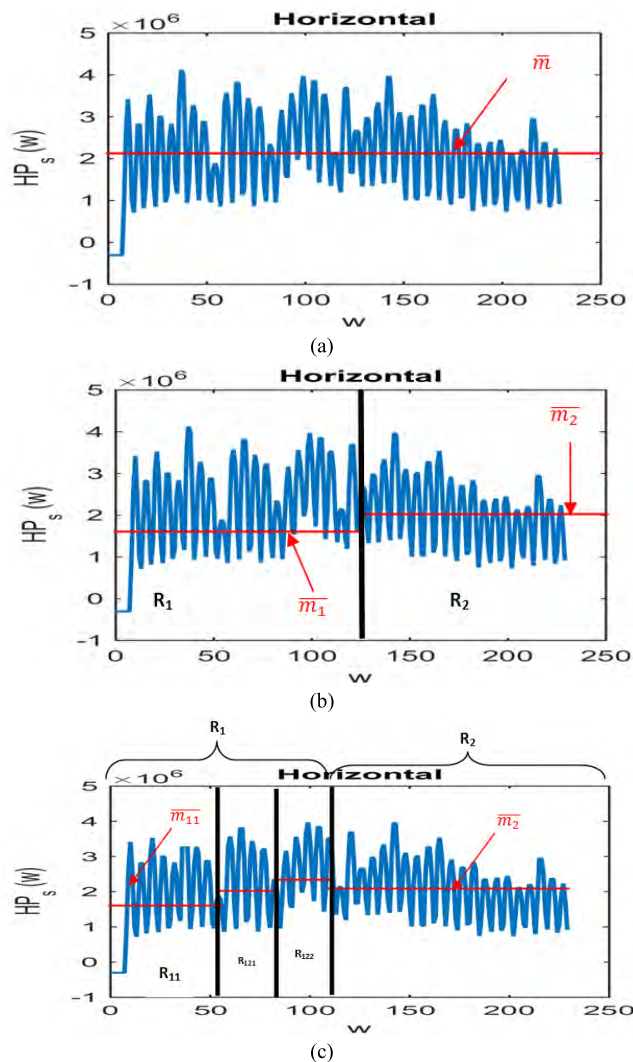
FIGURE 9. (a) The horizontal projection profile and the global mean value ($\overline{m}$) of this profile is indicated. (b) The horizontal projection profile after it has been segmented into two regions (R_1, R_2) and the mean values ($\overline{m_1}, \overline{m_2}$) of the two regions. (c) The horizontal projection profile after segmenting it into four regions ($R_{11}, R_{121}, R_{122}, R_2$) and the mean values of the four regions.
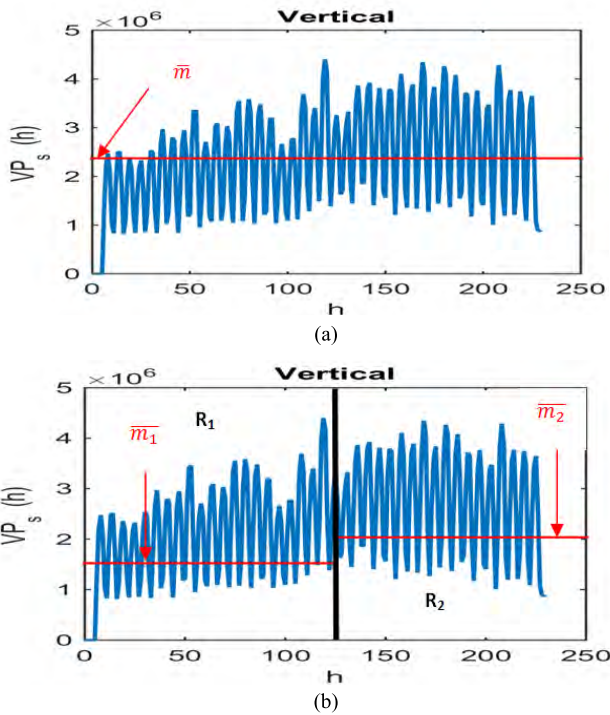
FIGURE 10. (a) The vertical projection profile and the global mean value ($\overline{m}$) of this profile is indicated. (b) The vertical projection profile after it has been segmented into two regions (R_1, R_2) and the mean values ($\overline{m_1}, \overline{m_2}$) of the two regions.

empirically ($k = 1.2$ for the images of the datasets used in this study). Then divide R_1 region into two regions: R_{11} and R_{12}. Then compute mean values ($\overline{m_{11}}$), ($\overline{m_{12}}$) for these two new regions and compare these mean values with the mean value ($\overline{m_1}$) of region R_1 for possible further division and so on. This division process is illustrated in Fig. 9(c).

Step 4: Subtract the mean value from the corresponding projection profile region. Let the outputs after subtracting the mean values be FD_{HP} and FD_{VP} vectors.

Step 5: Convert FD_{HP} and FD_{VP} vectors into binary profile trajectories as follows:

$$F_H = \begin{cases} 1, & FD_{HP} > 0 \\ 0, & \text{otherwise} \end{cases} \qquad (9)$$

$$F_V = \begin{cases} 1, & FD_{VP} > 0 \\ 0, & \text{otherwise} \end{cases} \qquad (10)$$

Step 6: Compute the middle of each interval of zeros in the binary profile trajectories F_H and F_V. These middle points are considered the locations of grid lines (boundaries between spots). These middle points are written into the H_{Ls} and V_{Ls} vectors.

Step 7: Apply the grid line optimization method presented in [21] to modify any line errors in our local gridding method.
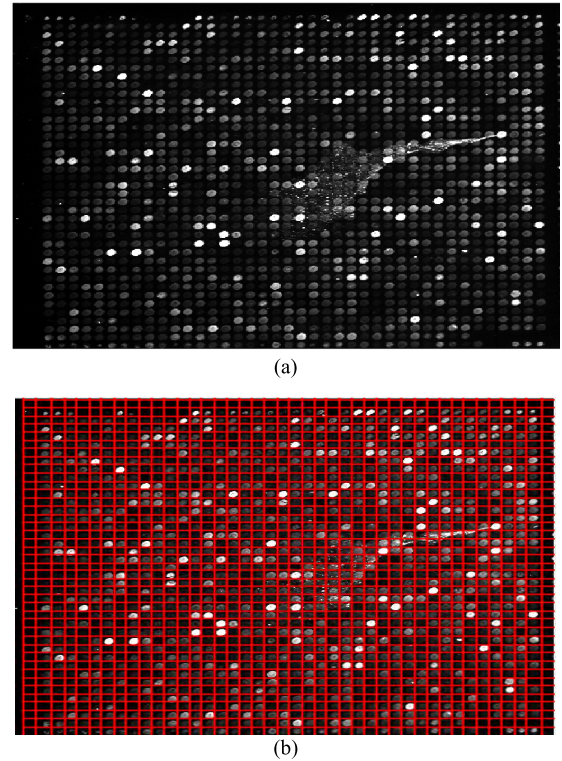
Step 8: Determine final horizontal and vertical gridding for the original sub-array (the special domain sub-array) as

follows:

$$FH_{Ls} = H_{Ls} * 2 \qquad (11)$$
$$FV_{Ls} = V_{Ls} * 2 \qquad (12)$$

Step 9: Map these grid lines FH_{Ls} and FV_{Ls} onto the sub-array to produce the spot regions as shown in Fig. 11.


FIGURE 11. Sub-array taken from an image of the 4th dataset: (a) before local gridding. (b) After applying the proposed local gridding method.

C. EXPERIMENTAL RESULTS AND PERFORMANCE ANALYSIS

This section demonstrates the output of each of the two steps of the proposed gridding method with typical microarray images. That is, the performance of the proposed method for global gridding and the performance of the proposed method for local gridding are investigated and the accuracy of each of the two steps is evaluated. In all experiments, the only input given to each of the proposed steps is the image itself and no other data is needed. Our gridding algorithm is implemented using Matlab R2014b running on laptop with Windows 8 platform.

1) PERFORMANCE OF THE SUB-ARRAY DETECTION APPROACH

The proposed sub-array detection approach (global gridding algorithm) is applied to 29 images taken from the first three datasets. It is to be noticed that the images of these datasets are different in resolution, number of sub-arrays, quality, and noise level. Also some of these images contain large number

(a)

(b)
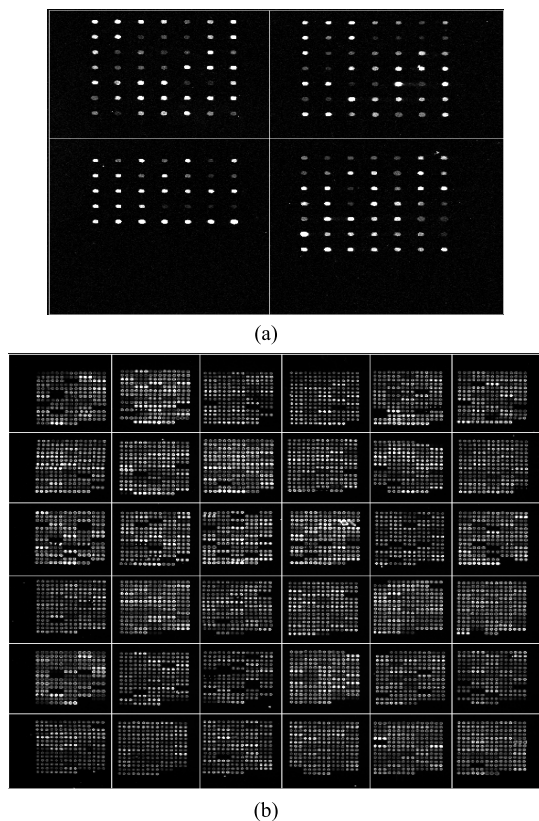
FIGURE 12. Examples of successfully obtained Sub-grids after applying the proposed global gridding method. (a) Sub-grid detection for an image taken from 1st dataset (SIB). (b) Sub-grid detection for an image taken from 2nd dataset (UCSF).

of poorly expressed spots. Fig. 12 illustrates the results of the proposed global gridding method for images taken from first and second datasets. Fig. 7 shows the global gridding for an image of the third dataset.

To evaluate the effectiveness of the proposed sub-array gridding approach, the accuracy of the method is computed according to the following expression:

$$Accuracy = \frac{N_{Correct\_sub-arrays}}{N_{Total\_Sub-arrays}} \times 100\% \qquad (13)$$

The accuracy of the global gridding for each microarray image of the three datasets is computed and presented in Table. 1.

TABLE 1. Accuracy of the proposed global gridding approach for three different databases.

Data Base	Number of images	Total number of sub-arrays	Number of correctly detected sub-arrays	Accuracy of the proposed method
SIB	14	56	50	89.29%
UCSF	2	72	72	100%
GEO	13	464	464	100%

From Table 1, it is observed that the accuracy of the first dataset (SIB) is lower than that of the other two datasets. This

is because the covering glass of image 664 is misplaced and the hybridization is very bad, and in image 665 the covering glass is broken.

2) PERFORMANCE OF THE PROPOSED SPOT DETECTION METHOD

The proposed spot detection method is tested on images taken from four databases. The images of these databases vary in image resolution, spot size, spot shape, spot resolution and sub-array layout. Each microarray spot, after applying our local gridding method, was visually classified into one of the following three categories:

- Perfectly (P): A spot was ''perfectly'' gridded if all its pixels were resided within the equivalent compartment of the grid cell.

- Marginally (M): A spot was ''marginally'' gridded when more than 80% of its pixels were resided within the equivalent compartment of the grid cell.

- Incorrectly (I): A spot was ''incorrectly'' gridded when less than 80% of the spot pixels were resided within its respective grid cell.

The spot detection results are shown in Table. 2. In this table, the results of the proposed method are compared with the results of two existing methods: maximum between class variance method [19] and Grid line refinement method [21]. The accuracy of the perfectly detected spots by the proposed method and the other two methods is shown graphically in Fig. 13.

TABLE 2. Spot detection accuracy obtained for the method presented in [19], the method presented in [21], and the proposed method.

| Data Base	Method of [19]			Method of [21]			Proposed method		
	I	M	P	I	M	P	I	M	P
SIB	2.0	12.9	85.1	5.48	5.414	89.11	4.08	1.33	94.59
UCSF	3.4	1.0	95.6	2.08	1.360	96.56	0	2.4	97.6
GEO	3.2	0.8	96	1.271	0.859	97.87	1.11	1.02	97.87
Dersi	5.238	1.33	93.4	4.625	0.563	94.81	0	1.67	98.33

From Table 2, it is observed that the proposed method has improved the spot detection accuracy compared to the other two methods. In database 1, the proposed method has achieved 5.48% accuracy improvement for spot detection compared to the other two methods. In database 2 and database 3, the percentage of spots which are correctly placed in their corresponding grid cells is very high, exceeding 97%.

For database 4, the proposed method has achieved 3.52% accuracy improvement for spot detection compared to the other two methods. Fig. 14 shows the local gridding results for an image taken from database 4. This figure shows that the proposed gridding method is able to detect a grid line which is missed in the gridded image given in [19]. This performance analysis indicates that the proposed method is robust against
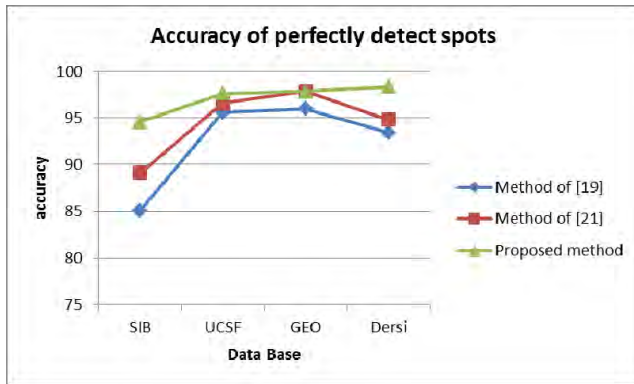
FIGURE 13. Accuracy of the perfectly detected spots by the proposed method and the methods in [19] and [21].

(a)

(b)
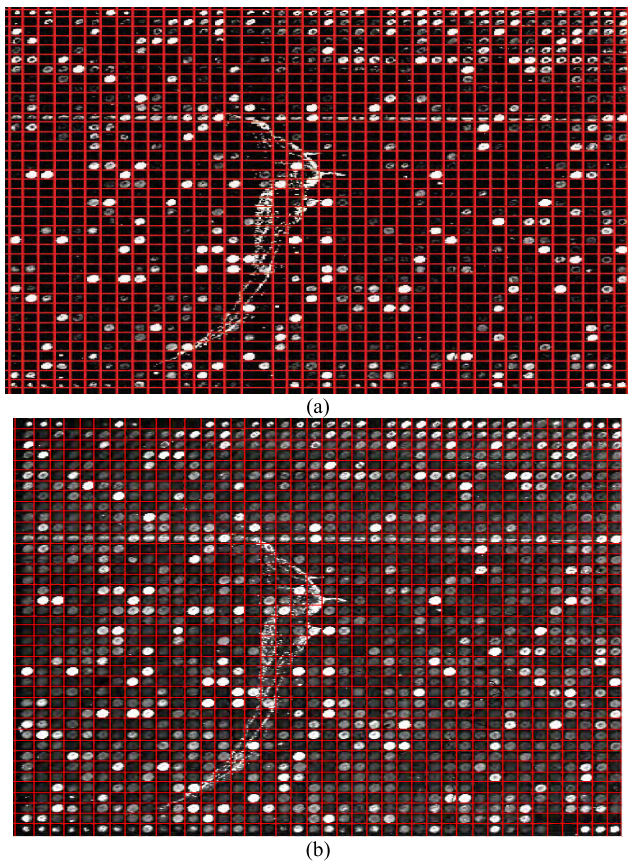
FIGURE 14. Sub-grid taken from image 1310_ch1_OD046_green of the 4th dataset.(a) Gridding result presented in [19]. (b) Gridding result after applying the proposed method.

various noises and contaminations that are commonly found in microarray images.

IV. CONCLUSION

The DNA microarray imaging technology has led to enormous progress in the life sciences by allowing scientists to analyze the expression of thousands of genes at a time. In this paper, a fully automatic gridding method for cDNA microarray images using multi-resolution analysis is presented. The proposed approach makes no assumptions about the spot size,

the number of rows and columns of sub-arrays or the number of spots in the microarray image. The proposed method is performed using two main steps. First, it extracts the sub-arrays of the entire microarray image (global gridding). Second, it identifies the locations of the spots in each sub-array (local gridding). The use of multi-resolution analysis in our approach has eliminated the need of pre-processing operations for noise removal and the use of local adaptive threshold method has produced optimal grid lines for separating the spot regions.

The proposed method has been tested on microarray images drawn from four sources. The experimental results show that the proposed method improves the spot detection accuracy compared to other published methods.

REFERENCES

[1] A. Baxevanis and B. Ouellette, *A Practical Guide to Analysis of Genes and Proteins*, 2nd ed. Hoboken, NJ, USA: Wiley, 2001.
[2] O. S. Baans and A. B. Jambek, "Software profiling analysis for DNA microarray image processing algorithm," in *Proc. IEEE Int. Conf. Signal Image Process. Appl. (ICSIPA)*, Kuching, Malaysia, Sep. 2017, pp. 129–132.
[3] S. Draghici and A. Kuklin, *Data Analysis Tools for DNA Microarrays*. Boca Raton, FL, USA: Chapman Hall, 2003.
[4] M. B. Eisen, *ScanAlyze User Manual*. Stanford, CA, USA: Stanford Univ., 1999.
[5] P. Hegde, R. Qi, K. Abernathy, C. Gay, S. Dharap, R. Gaspard, J. E. Hughes, E. Snesrud, N .Lee, and J. Quackenbush, "A concise guide to cDNA microarray analysis," *BioTechniques*, vol. 29, no. 3, pp. 548–562, Sep. 2000.
[6] *6.1 User Manual*, I. ImaGene, Los Angeles, CA, USA, 2008.
[7] *GenePix 4000A User's Guide*, Axon Instrum., Inc., Union City, CA, USA, 1999.
[8] N. Deng and H. Duan, "The automatic gridding algorithm based on projection for microvirray image," in *Proc. Int. Conf. Intell. Mechatron. Automat.*, Aug. 2004, pp. 254–257.
[9] P. Bajcsy, "An overview of DNA microarray grid alignment and foreground separation approaches," *J. Adv. Signal Process.*, vol. 1, pp. 1–13, Dec. 2006.
[10] M. Katzer, F. Kummert, and G. Sagerer, "A Markov random field model of microarray gridding," In *Proc. ACM Symp. Appl. Comput.*, Melbourne, FL, USA, Mar. 2003, pp. 72–77.
[11] J. Buhler, T. Ideker, and D. Haynor, "Dapple: improved techniques for finding spots on DNA microarrays," UW CSE, Madison, WI, USA, Tech. Rep. 112, Aug. 2000.
[12] Y. Wang, F. Shih, and M. Ma, "Precise gridding of microarray images by detecting and correcting rotations in subarrays," in *Proc. 8th Joint Conf. Inf. Sci.*, Jul. 2005, pp. 1195–1198.
[13] J. Deepa and T. Tessamma, "A new gridding technique for high density microarray images using intensity projection profile of best sub image," *Comput. Eng. Intell. Syst.*, vol. 4, no. 1, pp. 7–19, 2013.
[14] L. Rueda, "Sub-grid detection in DNA microarray images," in *Proc. Pacific-Rim Symp. Image Video Technol.*, 2007, pp. 248–259.
[15] Y. Wang, Q. M. Marc, K. Zhang, and Y. F. Shih, "A hierarchical refinement algorithm for fully automatic gridding in spotted DNA microarray image processing," *Inf. Sci.*, vol. 177, no. 4, pp. 1123–1135, Feb. 2007.
[16] L. Rueda and I. Rezaeian, "A fully automatic gridding method for cDNA microarray images," *BMC Bioinf.*, vol. 12, no. 1, Dec. 2011, Art. no. 113.
[17] E. Zacharia and D. Maroulis, "An original genetic approach to the fully automatic gridding of microarray images," *IEEE Trans. Med. Imag.*, vol. 27, no. 6, pp. 805–813, Jun. 2008.
[18] L. Rueda and V. Vidyadharan, "A hill-climbing approach for automatic gridding of cDNA microarray images," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, vol. 3, no. 1, pp. 72–83, Jan./Mar. 2006.
[19] G.-F. Shao, F. Yang, Q. Zhang, Q.-F. Zhou, and L.-K. Luo, "Using the maximum between-class variance for automatic gridding of cDNA microarray images," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, vol. 10, no. 1, pp. 181–192, Jan./Feb. 2013.

[20] Z. Gan, N. Zeng, F. Zou, J. Chen, M. Du, L. Liao, H. Li, and Y. Zhang, "Multilevel segmentation optimized by physical information for gridding of microarray images," *IEEE Access*, vol. 7, pp. 32146–32153, 2019.

[21] V. G. Biju and P. Mythili, "Microarray image gridding using grid line refinement technique," *J. Image Video Process.*, vol. 5, no. 4, pp. 1010–1016, 2015.

[22] M. Srivastava, C. L. Anderson, and J. H. Freed, "A new wavelet denoising method for selecting decomposition levels and noise thresholds," *IEEE Access*, vol. 4, pp. 3862–3877, 2016.

[23] D. A. Adjeroh, Y. Zhang, and R. Parthe, "On denoising and compression of DNA microarray images," *Pattern Recognit.*, vol. 39, no. 12, pp. 2478–2493, Dec. 2006.

[24] Z. Gan *et al.*, "Wavelet denoising algorithm based on NDOA compressed sensing for fluorescence image of microarray," *IEEE Access*, vol. 7, pp. 13338–13346, Jan. 2019.

[25] P. Nandihal, V. S. Bhat, and J. Pujari, "Adaptive min max thresholds algorithm of microarray image denoising based on nonsubsampled contourlet transform," *IJERT J.*, vol. 6, pp. 1–5, 2018.

[26] *SIB*. Accessed: Jan. 15, 2015. [Online]. Available: http://www.isrec.isb-sib.ch/

[27] *UCSF*. Accessed: Jan. 15, 2015. [Online]. Available: http://cancer.ucsf.edu/research/cores/array/array-sampledata

[28] *GEO*. Accessed: 2019. [Online]. Available: http://www.ncbi.nlm.nih.gov/geo/

[29] (2010). *DeRisi*. Accessed: 2019. [Online]. Available: http://www.bio.davidson.edu/projects/magic/magic.html

[30] E. Billauer. (2011). *Peak Detection Using MATLAB*. [Online]. Available: http://www.billauer.co.il/peakdet.html

ZAKI B. NOSSAIR received the B.S. degree in electronics and communications engineering from Helwan University, Cairo, Egypt, in 1978, the M.Sc. degree in electrical engineering from the Stevens Institute of Technology, Hoboken, NJ, USA, in 1985, and the Ph.D. degree in electrical engineering from Old Dominion University, Norfolk, VA, USA, in 1989. From 1990 to 1992, he was an Assistant Professor with the Department of Electronics and Communications Engineering, Helwan University, where he was an Assistant Professor, from 1994 to 1996, an Assistant Professor, from 2004 to 2006, and an Associate Professor, from 2006 to 2014. From 1992 to 1994, he was a Research Scientist with the Department of Electrical and Computer Engineering, Old Dominion University. From 1996 to 2004, he was an Assistant Professor with the Department of Computer Technology, College of Technology, Ryadh, Saudi Arabia. Since 2014, he has been a Professor Emeritus at Helwan University. His main research interests include signal processing and artificial intelligence.

MARY MONIR SAEID received the B.S. degree in electronics and communication from Fayoum University, Fayoum, Egypt, in 2006, the M.Sc. degree in electronics, communication, and computer engineering Helwan University, Helwan, Egypt, in 2011, where she is currently pursuing the Ph.D. degree in electronics, communication, and computer engineering.

Since 2011, she has been an Assistant Lecturer with the Department of Information Systems, Faculty of Computer and Information, Fayoum University. Her research interest includes signal and image processing and its medical applications.

MOHAMED ALI SALEH received the B.S. degree in electronics and communications engineering from Helwan University, Cairo, Egypt, and the M.Sc. and Ph.D. degrees from USA. He is currently an Assistant Professor with the Department of Electronics, Communications, and Computer Engineering, Faculty of Engineering, Helwan University. His research interests include signal processing and artificial intelligence.

● ● ●