Buletinul Ştiinţific al Universităţii "Politehnica" din Timişoara

Seria ELECTRONICĂ şi TELECOMUNICAŢII
TRANSACTIONS on ELECTRONICS and COMMUNICATIONS

Tom 52(66), Fascicola 1, 2007

FPGA based digital device for microarray image processing

Bogdan BELEAN[1], Monica BORDA[1], Raul MALUTAN[1], Victoria BIARGE-RODELLAR[2], Pedro GOMEZ [2]

Abstract - The present paper proposes a digital device architecture for microarray image processing and acquisition which takes advantage of parallel computation capabilities offered by FPGA technology. A microarray image process described by the steps filtering, image enhancement and segmentation was implemented using FPGA technology.

Keywords: microarray image, FPGA technology, parallel processing, genomic signal processing

I. INTRODUCTION

Gene expression represents the transformation of gene's information into proteins. For identifying a gene in a biological sequence and predicting the function of the identified gene, quantifying gene expression is mandatory. Measurement of gene expression can provide clues about regulatory mechanism, biochemical pathways and broader cellular function. The methods for detecting and quantifying gene expression levels include: key recombinant DNA technologies such as the PCR (Polymeric Chain Reaction), Serial Analysis of Gene Expresion (SAGE) and Complementary DNA (cDNA) microarray [1].

The last mentioned, microarray technology is based on creating DNA microarrays which represents gene specific probes arrayed on a matrix such as a glass slide or microchip. Usually samples from two sources are labeled with two different fluorescent markers and hybridized on the same array (glass slide). After hybridization, the array is scanned using two light sources with different lengths (red and green) to determine the amount of labeled sample bound to each spot through hybridization process. The light sources induce fluorescence in the spots which is captured by a scanner and a composite image is produced [2].

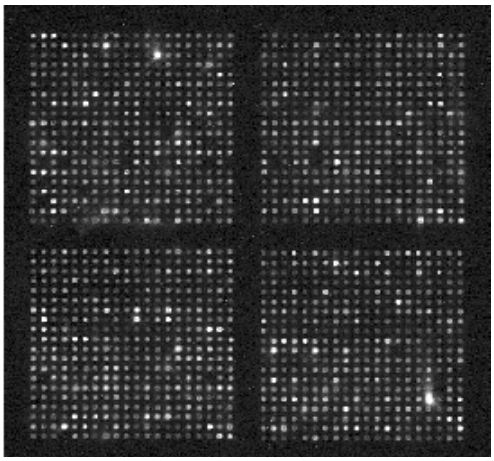

Figure 1. Microarray image.

While producing the composite image, the glow of the slide and also the unwashed sample material contribute to the background noise of the image. The surface of the spots is not regullar so in this way false spot boundaries appear. The mouvement of the glass slide while image acquisition may also affect spot positions and also a rotated image can be obtained. Other errors may appear because of the presence of dust or small fibers. To decrease the background noise the scanners can be adjusted by decreasing their sensitivity but this also weeks the precision in detecting weekly expressed spots. Due to all these inconveniences and to the large amount of data produced by a microarray experiment the main parameters to be taken into consideration during data analyses are accuracy and processing time.

The accuracy is given by the quality of the image processing techniques. The flow of processing a microarray image is generally separated in three tasks: addressing, segmentation and intensity extraction, which are described in the next chapter [3]. Before all these a preprocessing step is introduced which include noise removal techniques and also image

[1] Facultatea de Electronică şi Telecomunicaţii, Catedra de Comunicaţii, Str. Baritiu Nr. 26, 4027 Cluj Napoca
[2] Universidad Politécnica de Madrid, Campus de Montegancedo, 28660, Boadilla del Monte, Madrid

enhancement strategies to improve the microarray image quality.

In order to increase the accuracy and the processing speed, this paper proposes an FPGA based digital device especially conceived for processing microarray images. Parallel processing strategies are to be applied in image processing using this device. Also, its only purpose being microarray image processing, the device will have the advantage of using all its resources for the same cause: speeding up the processing by parallelizing image processing techniques.

II. MICROARRAY IMAGE PROCESSING

The major tasks of microarray image processing are to identify the array format including the array layout, spot size and shape, spot intensities and distances between spots. In order to do this we have to follow the three steps specific to the processing of a microarray image: addressing, segmentation and intensity extraction. The first step associates an address to each spot of the image. In the second one, pixels are classified either as foreground, representing the DNA spots, or as background. The last step calculates the intensities of each spot and also estimates background intensity values. All these steps can be preceded by a preprocessing step in order to remove noise and also to improve image quality by enhancing weekly expressed spots.

The main goal of the cDNA microarray image processing strategies presented in this paragraph is to eliminate the shortcoming of the previous processing methods, which needed user intervention to manually adjust a grid to determine the spot addresses in the image. The first strategy is based on generating image vertical and horizontal profiles by summing up intensities in each row and column [4]. By calculating the minimum and the maximum values we can determine a grid positioned around each spot, and also we can estimate spot dimensions. This method works well except in case that weakly expressed spots exist. The problem of weakly expressed spots is eliminated using image transformations to enhance spot intensities. Another strategy used in microarray image processing is to take advantage of the fact that a microarray image can be divided into independent sub-images. Considering this, a global addressing step is introduced which separates the microarray image into independent sub-images. This strategy provides the benefit of parallelizing further processing of microarray image, due to the fact that the independent sub-images can be processed in the same time using an appropriate technology. The technology proposed by this paper is FPGA and the architecture for a FPGA based digital device for microarray image processing is described in the next chapter.

Taking into consideration the strategies mentioned above, a microarray image I_0 is treated as follows:

Step 1: Eliminate shot noise using median filters

Step 2: Enhance the image I_0 using the following transformation [5]:

$$I_N = \frac{Io(x,y)^{\lambda_x+\lambda_y}}{\lambda_x+\lambda_y} \quad (1) \quad \text{where}$$

I_N is the resulted image after the enhancement and x,y are the horizontal and vertical pixel indices, the origin coresponding to the upper left corner,

$$\lambda_x = \frac{1}{\frac{1}{y}\cdot\sum_{i=0}^{Y-1} Io(x,n)} \quad (2)$$

$$\lambda_y = \frac{1}{\frac{1}{x}\cdot\sum_{i=0}^{X-1} Io(n,y)} \quad (3)$$

Step 3: Calculate the horizontal and vertical profiles of the image. Look for the minimum in the profiles. Separate the image into independent sub-images based on the obtained minimum.

Step 4: Each sub-image is treated as an independent image and based on the minimums from horizontal and vertical profiles a grid is estimated for each subimage. In the figure 3 a grid is overlaped on a sub-image.

Step 5: A segmentation method based on spot border detection called Mexican Hat is applied.

As shown in figure 2, to detect spot borders, a 3x3 matrix is overlapped on the enhanced image. Based on the value of the surrounding pixel we determine if the current pixel belongs to the spot's contour or not.

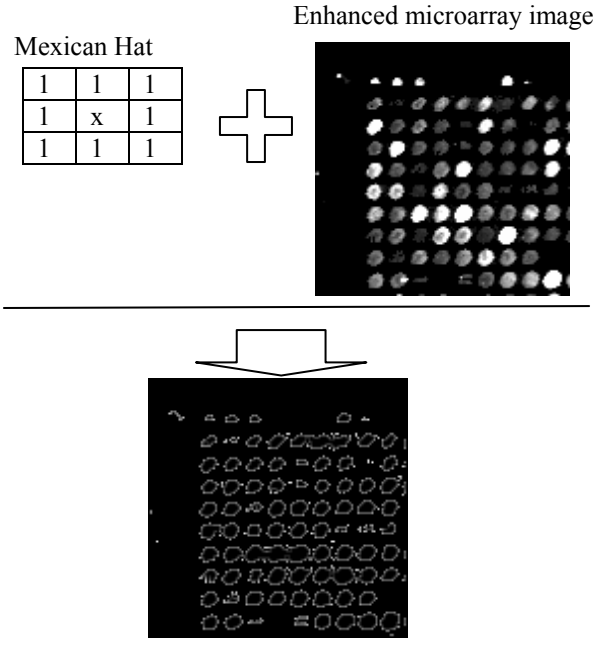

Figure 2. Spot border detection using Mexican Hat algorithm. (MatLab results)

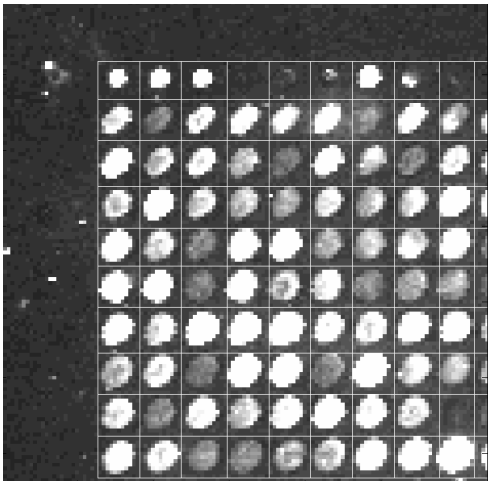

Figure 3. Overlaped grid on a microarray sub image. (MatLab results)

The steps mentioned above describe the workflow of a microarray data analyses process from image filtering to image segmentation.

III. DIGITAL DEVICE FOR MICROARRAY IMAGE PROCESSING

In a microarray image acquisition and processing system many decisions are left to an operator who may not have the adequate skills in statistical image processing. Also, the necessary equipment and technology for a microarray analysis is not portable and, due to the large amount of data contained in a microarray image, the existent technology is not efficient in field application were fast results are needed. The main objective of the present work is to provide the specialist with a digital device that can scan a cDNA microarray glass slide. After scanning, image processing algorithms are applied to the image and the results are delivered to the specialist through a general purpose interface.

The parameters to be taken into consideration when designing such a digital device are: processing time, accuracy and portability. To obtain the best processing time and accuracy, the FPGA technology was chosen to implement the design, due to the possibility of parallelizing the workflow of a microarray image process and the calculation specific to this type of image processing. The portability of the device is given by the USB connection between the FPGA based digital device and any personal computer.

The architecture under design fulfilling these objectives may be seen in Figure [5]. First a microarray glass slide obtain after hybridization is swap by a double–laser scanning device. Due to the laser excitation fluorescence levels are produced. A highly sensitive CCD (charge coupled device) sensor is used to store the microarray image composed by the fluorescence levels found along the glass slide, after hybridization. The CCD device with the help of an ADC will deliver serially to the FPGA an array of integer values which represent a digitized microarray image. Using a memory controller the microarray image is stored into a Fast Asynchronous SRAM as a grey-scale image data object for further processing. Using FPGA technology algorithms specific to microarray image data analyses are parallelized and applied to the gray-scale images. Results are deliver trough a USB 2.0 interface to the user.

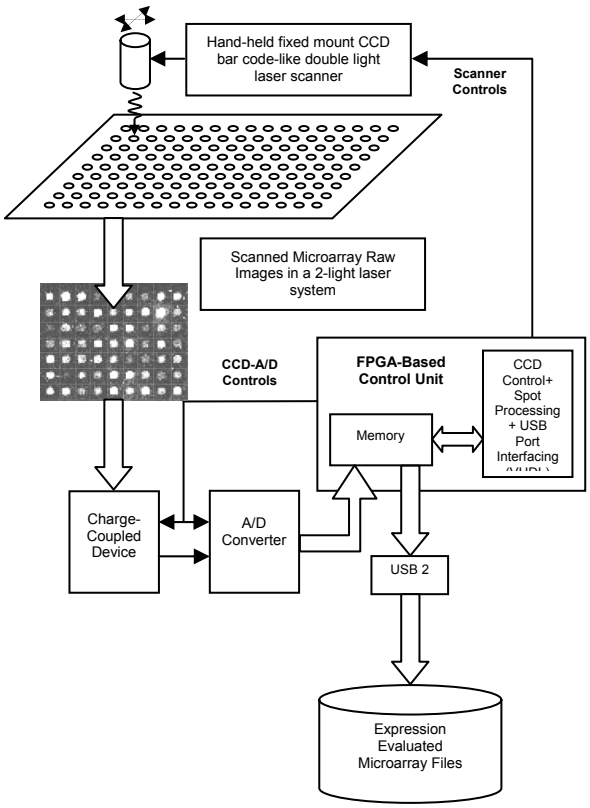

Figure 4. System architecture of the digital device. The CCD device gathers the cDNA microarray image after laser scanning. The image is stored in memory and processed by the synthesized implementation of the spot detection algorithms from VHDL code. Results are exported by the USB2 interface to a host computer.

The development is based on the popular Spartan 3 from Xilinx. The USB 2.0 module provided by Digient realizes the user interface with the Spartan 3 board. The central processing unit of the whole system is implemented on the FPGA Spartan 3 board and has the following functions:

a) Scanner control – FPGA control unit sets the light, speed and level for the double light laser scanner;

b) CCD and A/D control - sets the speed and synchronize the data transfer between the CCD sensor and Analog to Digital Converter. The transfer rate was estimated at 12.5 Mbytes/sec.;

c) Memory Storage – realized using a memory controller described in VHDL language by a Finite State Machine;

d) USB port control – FPGA control unit uses the Cypress CY7C68013 controller integrated on the USB 2.0 module for modeling the USB 2.0 user interface based on an Enhanced Parallel Port protocol.

e) Microarray spot processing (grid and border detection) – represents a set of routines implementing the functions described in section 3 translated to HDL for fast synthesis to FPGA programming. Due to the FPGA technology feature of accessing hundreds of memory addresses and multipliers in the same time, calculations specific to microarray spot processing algorithms are done in parallel. Data types are integers and the most problematic processing is posed by logarithm estimation, which is done through look-up-tables.

IV. EXPERIMENTAL RESULTS

In Chapter II were presented algorithms for grid detection, spot border detection and for cDNA microarray image filtering and transformation implemented in MatLab. The same strategies were used for the VHDL implementation of the border detection algorithm together with some preprocessing methods. The difference between these two types of implementation is the way in which the calculations specific to microarray image processing are made. In the case of VHDL implementation, the whole image is written into FPGA and the calculations are made in parallel due to FPGA's possibility of accessing hundreds of memory addresses in the same time. On the other hand, for the MatLab implementation, calculations are made sequentially, processing time being greater compared with the VHDL implementation. Statistics on the use of FPGA resources for a typical implementation strategy are given in Table 1.

Table 1. Statistics of FPGA resource usage

Device Utilization Summary

Logic Utilization	Used	Available	Utilization	Note(s)
Number of Slice Flip Flops:	1,288	15,360	8%	
Number of 4 input LUTs:	5,932	15,360	38%	
Logic Distribution:				
Number of occupied Slices:	3,933	7,680	51%	
Number of Slices containing only related logic:	3,933	3,933	100%	
Number of Slices containing unrelated logic:	0	3,933	0%	
Total Number 4 input LUTs:	6,136	15,360	39%	
Number used as logic:	5,932			
Number used as a route-thru:	204			
Number of bonded IOBs:	12	173	6%	
Number of GCLKs:	1	8	12%	

As a general idea the processing of a 12x12 pixel square to implement border detection requires 60 ns of processing time, which amounts to a total of 17 msec to process the global microarray image referred to in the last section.

The board used for the implementation of the cDNA microarray process described in the previous paragraph is Spartan 3 with XC3S1000.

V. CONCLUSIONS AND FUTURE WORK

The results presented in this paper show the importance of using automatic methods less dependent on the operator skills in cDNA microarray image processing. The technology chosen to implement an automatic digital device for microarray image processing was FPGA, due to the parallel computation capabilities and to the posibility of reconfiguration. The experimental results made this hardware technologie a solution for realising a fast and accurate automated system for cDNA microarray image processing.

An ideea for future implementations is to take into consideration the posibility of croping a microarray image into independent sub-images. This way, designing a multiprocessor system which processes in parallel two or more microarray sub-images becomes a challenge in cDNA microarray image processing.

ACKNOLEDGEMENTS

This work was supported by the Socrates Erasmus 2006 grant (student mobility) and also by PNII IDEI Nr. 909/2007 grant.

REFERENCES

[1] Paul A. Clarke, Robert te Poele, Richard Wooster, Paul Workman, "Gene expression microarray analysis in cancer biology, pharmacology and drug development: progress and potential", Biochemical Pharmacology NO 62, pp. 1311-1336, 2001

[2] Peter Bajcsy, "An Overview of DNA Microarray Image Requirements for Automated Processing", IEEE Transactions on Image Processing VOL 13, NO 1, pp. 15-25, January 2004

[3] Y.H. Yang, M.J. Buckley, "Comparison of methods for image analysis on cDNA microarray data", Technical Report 586, University of California – Berkley, 2001

[4] Q. Li, C. Fraley, R. E. Baumgarner, K. Y. Yeung and A. E. Raftery, Donuts, Scratches and Blanks. Robust, Technical Report no. 473, University of Washington, January 2005.

[5] P. Gomez, F. Diaz, B. Belean,, R. Malutan, B. Stetter, R. Martinez, V. Rodellar, "Robust cDNA microarray image processing on a hand-held device"