



Real-time lossless compression of microarray images by separate compaction of foreground and background



N. Karimi^a, S. Samavi^{a,b,*}, S. Shirani^b, A. Banaei^c, E. Nasr-Esfahani^a

^a Dept. of Electrical and Computer Engineering, Isfahan University of Technology, Isfahan, Iran

^b Dept. of Electrical and Computer Engineering, McMaster University, Hamilton, ON, Canada

^c Dept. of Electrical and Computer Engineering, University of Texas, Arlington, TX, USA

ARTICLE INFO

Article history:

Received 5 December 2012

Received in revised form 24 August 2014

Accepted 2 December 2014

Available online 9 December 2014

Keywords:

Image compression

High content screening

Microarray

Hardware implementation

ABSTRACT

Microarray technology is a powerful tool in molecular biology which is used for concurrent monitoring of a large number of gene expressions. Each microarray experiment produces hundreds of images. In this paper, we present a hardware scheme for lossless compression of 16-bit microarray images. The most significant parts of the image pixels are compressed by classifying them into foreground and background regions. This increases the spatial redundancy of the image. The foreground regions are packed together using a novel compaction unit. Real-time compression of these images is achieved while the bit-per-pixel values are comparable with the standard offline compression tools.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

During the past two decades the field of molecular genetics has attracted growing interest of scientists. Biomedical research evolves and advances not only through the compilation of knowledge but also through the development of new technologies. Using traditional methods to assay gene expression, researchers were able to survey a relatively small number of genes at a time. The emergence of microarray technology has enabled researchers to address previously intractable problems and to uncover novel potential targets for therapies. Microarrays allow scientists to analyze expression of many genes in a single experiment quickly and efficiently. They represent a major methodological advance and illustrate how the advent of new technologies provides powerful tools for researchers. Scientists are using microarray technology to try to understand fundamental aspects of growth and development as well as to explore the underlying genetic causes of many human diseases [1,2].

A sample gray scale microarray image is shown in Fig. 1. The size of this sample image is 455×455 pixels. In microarray images, the regions of interest are the subset of pixels that correspond to the spots (the foreground). The spots are approximately circular but significant variations are possible due to experimental fluctuations in the spotting procedure. Perturbations on the spot position, irregularities in the spot shape and size, holes in spots, unequal distribution of the probe within spots,

variable background and global problem that affect multiple spots (like scratches, contamination, and dust) are common. In general, the shape and the size of the spot can fluctuate significantly across the array. The background is usually much less crucial for the downstream image processing step, although it is used as a reference for background correction operations [3,4]. Each microarray experiment produces thousands of images. Microarray images are large in size and also each experiment is costly. Various organizations share their microarray databases. For efficient storage and for sharing a large number of these images, image compression is essential [5,6].

Compression of microarray images could be done by using lossless or lossy methods. Even though compression ratios of lossless methods are less than those of lossy methods, for biomedical applications lossless methods are the more appropriate choice.

In microarray images, the intensity level of each spot contains important information about the gene expression. Thus, to preserve the intensity and the size of the spots the lossless compression schemes are more suitable. Several compression methods have specifically been proposed for the microarray images. Most of these algorithms are software based while a number of them have been realized by hardware.

The method that Jørnsten et al. [7] have proposed performs gridding of the whole image at the beginning and then the spots are segmented. The segmentation of the spots is achieved by using the approximate center of each spot and a seeded region growing procedure followed by a two-component Gaussian mixture model. The segmentation mask is encoded using chain-coding and the interior of the regions is encoded using a modified version of the LOCO-I (LOW COMPLEXITY LOSSLESS COMPRESSION for IMAGES) algorithm. Hence the algorithm is called SLOCO.

* Corresponding author at: Isfahan University of Technology, Iran.
E-mail address: samavi96@cc.iut.ac.ir (S. Samavi).

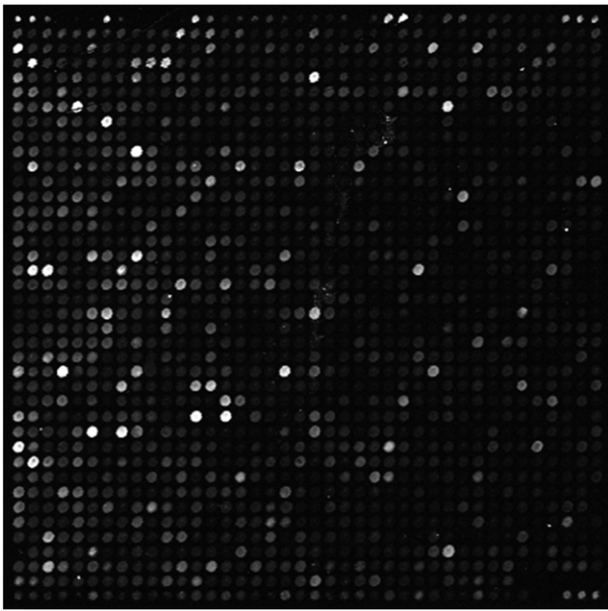


Fig. 1. A sample gray scale microarray image.

Hua et al. [8] proposed a method called BASICA (Background Adjustment, Segmentation, Image Compression and Analysis). The algorithm is based on wavelet transformation. Initially, a segmentation is performed using the Mann–Whitney algorithm, and the segmentation mask is encoded separately. The wavelet-transforms of the foreground and the background are formed separately and the wavelet coefficients are coded using object-based EBCOT (Embedded Block Coding Optimized Truncation). It is possible to code the background in a lossy manner while the foreground is losslessly compressed.

Faramarzpour et al. [9] proposed a software-based lossless compression method. Their method isolates each spot into an individual region of interest. A spiral path is fit on each spot such that its center coincides with the center of mass of the spot. The idea is to convert the 2D structure of the image into a 1D sequence which can scan the image in a highly correlated manner while preserving its spatial continuity. It is expected that the DPCM coding of this one dimensional signal possesses low entropy. Then, predictive coding is applied along this path, with a separation between residuals belonging to the foreground area and those belonging to the background area.

Lonardi and Luo [10] proposed lossless and lossy compression algorithms for microarray images (MicroZip). The method uses a fully automatic gridding procedure, similar to that of Faramarzpour's method, for separating spots from the background (which can be lossy compressed). Through segmentation, the image is split into two channels: foreground and background. Then, for entropy coding, each channel is divided into two 8 bit sub-channels and arithmetic encoded, with the option of being previously processed by a Burrows–Wheeler transform.

The method presented by Zhang [11] performs lossless compression of microarray images. This method uses a known spot segmentation procedure [12] to categorize pixels into foreground and background sets. Compression is then performed separately on each set. Each pixel value is considered as consisting of a MSB part and a LSB part. A prediction step is performed on the MSB part of the pixels. Then PPAM (Prediction by Partial Approximation Matching) compression algorithm is used on the predicted MSB parts and the LSB parts of the foreground and the background, separately.

Reference [13] presents a lossless method for compression of microarray images based on arithmetic coding using a 3D context model. The 3D context modeling is performed on consecutive bit-planes of an image. The compression starts from the MSB plane and works its way toward the LSB plane. In each plane the modeling of each bit is

performed using its neighboring bits and the corresponding bit from the previous compressed plane. In [14,15] the lossless method of [13] is improved. In the improved version, for each plane a context is defined based on the causal neighbors in the current plane and non-causal neighbors in the previous plane. Then the current bit, based on the mentioned contexts, is coded using arithmetic coding.

The algorithm presented in [16] is a lossless image compression method which segments the pixels of the image into three categories of background, foreground, and spot edges. The segmentation is performed by finding a threshold value which minimizes the weighted sum of the standard deviations of the foreground and background pixels. Each segment of the image is compressed using a separate predictor.

It has been shown that the quality of DNA microarray data has high variations. Before using the images for biological purposes, usually, extensive gridding and image processing algorithms are needed. Hence, hardware implementations of these algorithms have been proposed to speed up the lengthy software procedures. In [17] cellular neural networks (CNNs) are employed, for microarray image analysis, in a dedicated machine called the CNN universal machine. In [18] two pipeline architectures are proposed for the lossy microarray image compression. The image is first processed and through morphological operations noise and very small spots are eliminated. For compression purposes they process the pixels of the image in a raster scan order. In the first architecture, “pseudo-RLE” method is applied and in the second architecture, “residual Huffman coding” method is used for the compression of images. In the residual Huffman coding method, codes are calculated for the difference between every two neighboring pixels. In [19] a lossy compression scheme is applied which uses real-time techniques. In their method the microarray image is first processed to get an all-zero background. By doing so, relatively high compression ratios were obtained. The main purpose of hardware schemes of [18] and [19] is real-time processing of microarray images. The compression part of these schemes uses simple methods that work on previously processed images with backgrounds that are free of noise. Hardware image compression schemes do exist for many different applications, such as [20], which are proposed for lossy compression of images in a camera sensor networks. Such designs cannot be used for biomedical imagery where changing of pixel values is not tolerable.

The motivation behind the hardware that is proposed in this paper is that right now hardware solutions have been proposed to speed up the complex and time consuming gridding of the microarray images [21]. Also, to speed up the analysis of these images hardware implementations have been employed [18,19]. The last stage in the life cycle of microarray images, before storing them, is their compression. Hence, in this paper we are proposing a hardware solution for the compression of microarray images to complement the existing hardware mechanisms of gridding and analysis. Even though the technology of microarray is a relatively mature one, it is a part of the growing high content screening (HCS) technology [22]. High-throughput cell-based screening has become a field of interest because it allows the advent and test of new drugs in extremely fast pace. Microarrays, fluorescent microscopy, RNAi (RNA interference), and a number of related tools are giving researchers a clear view of cellular activities and a means of creating new drugs [22]. The hardware solutions that are being proposed are needed remedies due to the high-throughput nature of the technologies. Furthermore the hardware solutions that are proposed for the microarray images are all applicable, with modifications, to other HCS technologies.

We propose to perform a classification on the most significant 8 bits of the pixels. The least significant part of the pixels is not compressed. The proposed hardware first classifies the significant part of the image into the foreground and background regions. The classification is done by an efficient thresholding scheme. The foreground pixels, the segmentation map and the background pixels are separately packed by a proposed compaction block. By separate treatment of the foreground and background pixels better exploitation of correlations is possible.

Download English Version:

<https://daneshyari.com/en/article/454695>

Download Persian Version:

<https://daneshyari.com/article/454695>

[Daneshyari.com](https://daneshyari.com)