DNA microarray image coding

Miguel Hernández-Cabronero[†], Juan Muñoz-Gómez[†], Ian Blanes[†]
Michael W. Marcellin[†‡] and Joan Serra-Sagristà[†]

[†] Department of Information and Communications Engineering,
Universitat Autònoma de Barcelona, Barcelona, Spain.

[‡] Department of Electrical and Computer Engineering
University of Arizona, Tucson, AZ, USA.

Abstract

DNA microarrays are useful to identify the function and regulation of a large number of genes in a single experiment, even whole genomes. In this work, we analyze the relationship between DNA microarray image histograms and the compression performance of lossless JPEG2000. Also, a reversible transform based on histogram swapping is proposed. Intensive experimental results using different coding parameters are discussed. Results suggest that this transform improves previous lossless JPEG2000 results on all DNA microarray image sets.

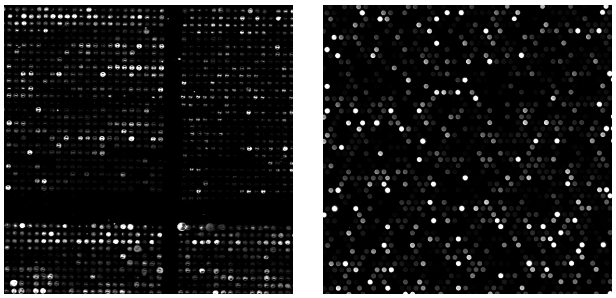
1 Introduction

1.1 DNA microarrays

DNA microarrays are a state of the art tool in medicine and biology for the study of genetic function, regulation and interaction [1]. Genome-wide monitoring is possible with existing DNA microarrays, which are used in research against cancer [2] and HIV [3], among many other applications. DNA microarrays consist of a solid surface on which thousands of different known genetic sequences, the oligonucleotides, are bound. Each sequence is contained in a single microscopic hole or *spot* and all spots are arranged conforming to a regular pattern, usually a rectangular or hexagonal grid. Example images for these two layouts are shown on Figure 1. Two samples dyed with fluorescent markers, usually Cy3 and Cy5 of the cyanine family, are made to react on the microarray. When one sample has expressed a gene, part of it is hybridized and adhered to the spot corresponding to that gene. The rest is washed away so that each dye is present in a spot proportionally to the activity of a gene in the corresponding sample. After the hybridization, the microarray is exposed to laser beams and the emissions from the fluorescent Cy3 and Cy5 dyes are recorded independently as so-called green and red channel images, respectively. Comparing the relative intensity of the green and red channels, it is possible to detect expression differences between two samples, which can be employed to make hypothesis about the function and regulation of thousands of individual genes.

Each microarray experiment outputs a pair of monochrome, single component images corresponding to the green and red channels. Due to the microscopic size of the spots, the produced images have a high spatial resolution: images from 1000×1000 onwards are typically described in the literature, but sizes over 4000×13000 are common nowadays. Since gene expression can vary in a very wide range and a high degree of precision is desired, DNA microarray images have a intensity resolution of 16 bits per pixel (bpp).

1068-0314/12 $26.00 © 2012 IEEE
DOI 10.1109/DCC.2012.11



(a) (b)

Figure 1: Example DNA microarray image: 600×600 crop with different spot layouts. a) *array3* image from the MicroZip set with square grid spot layout; b) *slide_1-red* from the Arizona set with hexagonal grid spot layout. Gamma levels have been adjusted for better viewing.

After the images have been recorded, they are computer analyzed to extract the genetic information present in them. However, analysis techniques are not fully mature or universally accepted, so it is preferable to keep the original images and not only the extracted genetic data because repeating an experiment is expensive and not always possible. Because of the high spatial and intensity resolutions, raw data for a single DNA microarray image can require from a few to hundreds of Megabytes. Most experiments are carried out under several different conditions, and with the increasing interest in DNA microarrays, very large amounts of data are created each year around the world. DNA microarray images need to be kept and shared, so efficient storage and transmission methods are required. In consequence, compression emerges as a natural approach.

Both lossy and lossless techniques have been proposed in the literature. Lossy approaches exhibit better compression performance on microarray images, but information loss is not globally accepted because it could affect reanalysis with future techniques. On the other hand, purely lossless methods guarantee perfect fidelity of the data, which is preferable for future reanalysis, at the cost of poorer compression performance as compared to lossy techniques. The efficient lossless compression of this type of image has proved to be a difficult task. This is partly due to the considerable amount of noise and the abundance of high frequencies present in this type of image. For this reason, original approaches like the one proposed in this work are needed to achieve the storage and transmission requirements for DNA microarray images.

1.2 State of the art in lossless compression

Many different techniques have been proposed for the lossy and lossless compression of DNA microarray images. In this subsection, we discuss lossless schemes that have been published in the literature. The typical image compression process consists of up to five stages: preprocessing, transform, quantization, entropy coding and postprocessing. Microarray image compression can be modeled likewise, but not all stages are equally relevant if we focus only on lossless compression: the quantization stage, which consists of dividing sets of values or vectors into groups, effectively reducing the total number of symbols needed to represent them, is not usually considered for lossless compression; the same happens for the postprocessing stage, consisting of processing images after compression to enhance their visual quality, to provide new features or to analyze their properties. Lossless techniques belonging to the rest of the stages are addressed next. A more exhaustive review of the state of the art can be found in the literature [4].

1.2.1 Preprocessing

The preprocessing stage comprises any computation performed on an image to prepare it for the compression or analysis processes. It is very important in DNA microarray images because many of the existing techniques rely heavily on the results of this stage to obtain competitive coding performance. The main preprocessing method is segmentation and consists in determining which of the image pixels belong to spots (i.e., the foreground), as opposed to those that do not (i.e., the background).

In 2003, Faramarzpour et al. proposed a segmentation stage consisting of two steps [5]: first spot regions are located by studying the row intensity sum minima, and then region centroids are used to estimate the spot centers. Simpler versions of this spot region location idea had already been used by Jörnsten and Yu in 2002 [6]. Later, in 2004, Lonardi and Luo presented their MicroZip software [7], which used a variation of Faramarzpour's spot region finding idea, but considering the existence of subgrids, which can be appreciated in Figure 1a. In 2004, Hua et al. proposed a scheme with a segmentation technique based on the Mann-Whitney U test [8]. In 2006, Bierman et al. described a simple thresholding method for dividing microarray images into low and high intensities [9], determining the lowest of the threshold values from 2^8, 2^9, 2^{10} or 2^{11} such that at least 90% of the pixels fall within it. In 2007, Neekabadi et al. proposed another threshold-based technique for segmentation [10] in three subsets (background, edge and spot pixels), using a threshold that minimizes the total standard deviation of pixels above and below it. In 2009, Battiato and Rundo published an approach based on Cellular Neural Networks (CNNs) [11].

1.2.2 Transform

The transform stage consists of changing the image domain from the spatial domain to a domain where it can be more efficiently processed or coded. However, transform based compression is not typically as efficient for DNA microarray images as it is for other types of images not containing such sharp edges [12]. For this reason, transformations are not frequently researched in microarray image compression, although they are used in some works.

In 2004, Hua et al. [8] published a modification of the EBCOT algorithm, the basis of the JPEG2000 standard [13], that included a tailored integer odd-symmetric transform. In 2004, Lonardi and Luo [7] made use of the Burrows-Wheeler transform [14] for lossy or lossless compression in their MicroZip software.

Table 1: Classification of lossless microarray-specific techniques discussed on Subsection 1.2, sorted chronologically.

Preprocessing	**Transform**	**Entropy coding**	
Segmentation		Segmentation	Context
[6], 2002	[8], 2004	[5], 2003	[15], 2005
[5], 2003	[7], 2004	[9], 2006	[16], 2006
[8], 2004		[11], 2009	[17], 2006
[7], 2004			[18], 2009
[9], 2006			
[10], 2007			
[11], 2009			

1.2.3 Entropy coding

In this stage, data obtained from previous stages are expressed in an efficient manner to produce a compact bitstream. Many techniques segment the image before compression,

while others build contexts or try to predict the intensity of the next pixels based on the previous ones. Purely lossless techniques using each approach are described next.

At least three different works that use segmentation can be found on lossless compression of DNA microarrays. In 2003, Faramarzpour et al. presented a prediction-based technique [5]. The image is gridded, and a linear prediction scheme is applied after creating a spiral path from the estimated spot center. In 2006, Bierman et al. presented their MACE (Micro Array Compression and Extraction) software [9]. The image is divided first using a threshold-based method. The low intensity pixels are coded using standard dictionary-based techniques, while the high intensity pixels are processed with a sparse matrix algorithm and then compressed. In 2009, Battiato and Rundo published an algorithm [11] based on image color reindexing after segmentation. Segmentation is made by means of a CNN-based system to produce two complementary subimages. The foreground image is compressed with a generic lossless algorithm and stored separately. The background image is first transformed into an indexed image. Then its color palette is reindexed with an algorithm that reduces the zero-order entropy of local differences, which are losslessly coded.

In no less than four publications, context building is used to perform lossless DNA microarray image compression. In 2005 and in 2006 Zhang et al. [15, 16] proposed a context-based lossless approach that also employs segmentation. Once the image is divided, a simple predictive scheme is used for the most significant bytes of each pixel, while the least significant bytes are coded using prediction by partial approximate matching (PPAM), also proposed by Zhang and Adjeroh [19]. In 2006, Neves and Pinho [17] proposed another context-based lossless approach. It is a bitplane-based technique that uses 3D finite-context models to drive an arithmetic coder. In 2009, they improved this scheme so that specific contexts are built for each image [18].

Table 1 presents a summary of all discussed methods classified attending to the stage of the image compression process in which they make their contribution.

1.3 Paper structure

This paper is organized as follows. We discuss the use of lossless JPEG2000 on DNA microarray images in Section 2. In Section 3, we analyze the typical histogram of a DNA microarray image and propose a point transform based on histogram swapping. Results for the application of this transform with lossless JPEG2000 are presented. Finally, we draw some conclusions in Section 4.

2 Lossless JPEG2000 coding of DNA microarray images

In this section, we study the performance of lossless JPEG2000 compression on DNA microarray images. We describe the image sets used for benchmarking in the literature and some of their properties in Subsection 2.1. We show lossless JPEG2000 compression results and compare them to previous data and other techniques in Subsection 2.2. We analyze the impact of the number of DWT decomposition levels and quality layers on the compression performance in Subsection 2.3.

2.1 Benchmark image sets

A number of different DNA microarray image sets have been used for benchmarking compression performance. No set has been used across all publications on DNA microarray image compression, but the MicroZip, ApoA1 and ISREC sets are employed more

Table 2: Image sets used in the literature.

Image set	Images	Size (px)
MicroZip [21]	3	$> 1800 \times 1900$
Yeast [22]	109	1024×1024
ApoA1 [23]	32	1044×1041
ISREC [24]	14	1000×1000
Stanford [20]	20	$> 2000 \times 2000$
Arizona	6	4400×13800

Table 3: Compression results in bpp for the Kakadu and JJ2000 implementation of the JPEG2000 standard, and for some other generic and microarray-specific techniques. Best results are highlighted in **green** and worst results in **red**.

Algorithm	MicroZip	Yeast	ApoA1	ISREC	Stanford	Arizona
Kakadu (5 levels)	9.508	**9.082**	11.052	11.360	8.007	9.099
JJ2000 (5 levels)	**9.515**	9.079	11.063	**11.366**	**8.010**	**9.106**
Bzip2	9.394	**6.075**	**11.067**	10.921	7.503	8.944
CALIC	9.281	8.502	10.515	10.615	7.248	8.767
JBIG	9.297	6.888	10.851	10.925	7.411	8.858
JPEG-LS	8.974	8.580	10.608	11.145	**7.204**	**8.646**
Battiato's [11]	**8.369**	–	**9.52**	**9.49**	–	–

frequently. The Stanford set was obtained from the Stanford Microarray Database public FTP [20] and the Arizona set has been kindly provided by David Galbraith and Megan Sweeney from the University of Arizona. Table 2 shows key properties of all sets documented in the literature. All images are monochrome, unsigned, 16 bits per pixel (bpp), and contain a single component per red/green channel.

2.2 Lossless compression performance

In this section, we report an experiment that we have conducted to test lossless JPEG2000 compression performance on DNA microarray images. We have compressed all images from the sets described in Subsection 2.1 using the Kakadu v6.0 [25] and the JJ2000 v5.1 [26] implementations of the JPEG2000 standard. In both cases, we have used lossless compression, 33 quality layers and 5 DWT decomposition levels. The number of quality layers was chosen to be the same as in a previous work by Pinho [27]. All codestreams are JPEG2000 part 1 compliant.

Table 3 shows compression results for the Kakadu and the JJ2000 implementation employing the mentioned configuration, and also for a generic compressor (Bzip2), some general image compressors (CALIC, JBIG and JPEG-LS) and the best microarray-specific compressor (Battiato's), as reported in a previous work [4]. These results have been computed dividing the total size in bits for all compressed files by the total number of pixels in the images.

It can be seen that both JPEG2000 implementations exhibit very similar compression performances, which are poor compared to the best microarray-specific technique, and generally to the other compressors as well. The results obtained with the JJ2000 implementation are consistent to the ones published by Pinho [27].

2.3 DWT decomposition levels and quality layers

In our experiments, we analyze the impact of varying the number of DWT decomposition levels and quality layers when using Kakadu JPEG2000 on DNA microarray images.

Table 4: Lossless compression results in bpp for different DWT decomposition levels using Kakadu JPEG2000. Best results are highlighted in **green** and worst results in **red**.

DWT levels	MicroZip	Yeast	ApoA1	ISREC	Stanford	Arizona
0	**10.027**	**6.829**	**11.525**	**10.888**	**8.567**	**9.548**
1	9.542	**9.089**	11.088	**11.476**	8.146	9.221
3	9.472	9.042	**10.999**	11.312	7.985	9.068
5	**9.467**	9.038	**10.999**	11.314	**7.969**	**9.064**

Table 4 shows results for 0, 1, 3 and 5 DWT decomposition levels and 1 quality layer.

For most sets, compression is improved by approximately 0.5 bpp when the number of decomposition levels is increased from 0 to 5, but most of that improvement is yielded when increasing from 0 to 1 level. Only the Yeast and ISREC sets break that pattern: for these two sets, using 0 decomposition levels produces the best results. Increasing to 1 decomposition level degrades measurably the performance, but further level increments improve the performance slightly as happens for the other sets. For all sets, using more than 5 decomposition levels does not modify the compression performance.

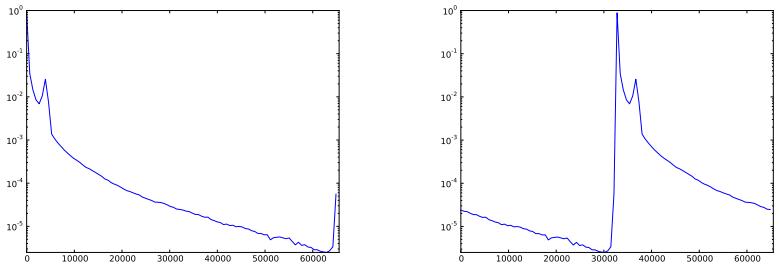
Increasing the number of quality layers from 1 to 33 decreases compression performance slightly, between 0.03 bpp and 0.04 bpp for all tested DWT decomposition levels and sets.

3 DNA microarray image histograms and JPEG2000

This section is organized as follows. In Subsection 3.1, we describe the typical DNA microarray image histogram and compare it to the histogram that JPEG2000 implicitly assumes. In Subsection 3.2 we propose an original transform based on histogram swapping and report compression results when employing lossless JPEG2000 after this transform.

3.1 Typical DNA microarray image histograms

DNA microarray images exhibit similar pixel intensity distributions across all data sets in Table 2. In a typical image, most pixels belong to the background or to low-activity spots and have very low values. Higher intensities are several orders of magnitude less frequent, but among them the intensities close to the maximum are usually over ten times more abundant. The histogram of a representative DNA microarray image from the Arizona set, *slide_1-red*, is shown in Figure 2a.



(a) Original histogram. (b) Histogram after the proposed transform.

Figure 2: Pixel value distribution for original and transformed image *slide_1-red* from the Arizona set using a semilog scale.

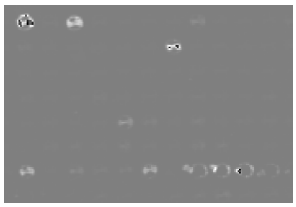
The characteristic distribution of this type of images partly explains the poor lossless compression results of JPEG2000 shown on Table 4. For unsigned image data with bit-depth B, the first step carried out by a JPEG2000 Part 1 encoder is to subtract 2^{B-1} from the value of each pixel [13]. This is known as the level offset stage and typically results in pixel values nominally distributed symmetrically about the origin. DNA microarray images are unsigned. Thus, the subtraction is performed resulting in a highly asymmetrical histogram, with the majority of pixel intensities taking values near -32768. Another problem is that microarray images have considerable high frequency content due to the many edge discontinuities between spots and background. This type of data is not well treated by the wavelet transform. In conclusion, JPEG2000 is receiving an input for which it is not designed, so a high compression performance cannot be expected.

3.2 Histogram swapping and lossless JPEG2000 compression

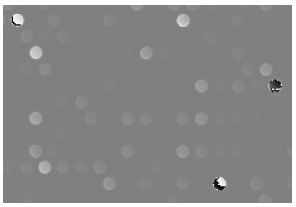
DNA microarray images possess pixel value distributions that diverge from natural images. However, DNA microarray images can be modified so that their intensity histograms become more similar to what JPEG2000 implicitly assumes. If the most significant bit of each pixel of an image is flipped, the right half of the histogram is swapped for the left half. This transformation, which we will call the *histogram swap transform* (HST), can be easily reversed by flipping again the most significant bit of each pixel. Figure 2b shows the pixel distribution of the transformed version of image *slide_1-red*. This histogram is much more nearly symmetric about the origin.

We have conducted an experiment to test JPEG2000 lossless compression on DNA microarray images after applying the proposed transform. We have used 1 quality layer and 0, 1, 3, and 5 DWT decomposition levels, and we report the results on Table 5. It can be observed that the compression performance is always improved when using the HST. Comparing the results for the best choice of decomposition levels before and after the HST, rate improvements from 0.213 bpp to 0.918 bpp (1.97% to 15.53%) can be measured. It is also noteworthy that the compression results follow a different trend after altering the image histograms. When compressing the original images, the performance is generally improved as the number of DWT decomposition levels is increased, as previously shown on Table 4. However, after applying the HST, this pattern is reversed and performance is generally degraded when the number of DWT decomposition levels is increased. This behavior can be explained via two observations. First, the histogram of Figure 2b is very peaked near the origin, reminiscent of the Laplacian distribution often assumed for wavelet transform coefficients [28]. This suggests that the bitplane coder of JPEG2000 may work well when applied directly to the data obtained via the HST (without further transformation). Second, when the HST is applied on an unsigned image with bit-depth B, pixel values slightly smaller than 2^{B-1} become close to 2^B, while values slightly greater than 2^{B-1} become close to zero. In other words, mid-gray values (pre HST) result in abrupt intensity differences between near black and near white (post HST). Examples of this behavior can be appreciated in Figure 3. This adds to the abundant discontinuities already present in DNA microarray images. Thus, increasing the number of DWT decomposition levels, which is not very effective when high frequencies are present [7], results in a performance reduction.

There are other point transforms that lead to histograms similar to Figure 2b. In fact, minor additional performance improvements might be obtained in this fashion. However, the HST has a distinct advantage in terms of implementation. In fact, it is possible to implement the HST without any explicit changes in JPEG2000 or the image data itself. Indeed, if the original unmodified data are simply interpreted as twos-complement signed values, they have exactly the same decimal values that result when the unsigned values are subjected to the HST followed by the JPEG2000 level offset process. Specifically when



(a) Detail of *array3* image from the MicroZip set after the HST.



(b) Detail of *slide-1_red* image from the Arizona set after the HST.

Figure 3: Details of sudden intensity changes after the HST.

Table 5: Lossless compression results in bpp applying Kakadu JPEG2000 after applying the HST, using 1 quality layer and different DWT decomposition levels. Rate differences in bpp as compared to compression of unmodified images. Differences between the best results before and after HST are shown at the bottom.

DWT levels	MicroZip	Yeast	ApoA1	ISREC	Stanford	Arizona
0	9.157 (-0.870)	5.911 (-0.918)	10.786 (-0.739)	10.624 (-0.264)	7.685 (-0.882)	8.795 (-0.753)
1	9.297 (-0.245)	8.862 (-0.227)	10.917 (-0.171)	11.238 (-0.238)	7.851 (-0.295)	8.967 (-0.254)
3	9.455 (-0.017)	9.026 (-0.016)	11.003 (+0.004)	11.300 (-0.012)	7.950 (-0.035)	9.058 (-0.010)
5	9.466 (-0.001)	9.035 (-0.003)	11.012 (+0.013)	11.313 (-0.001)	7.958 (-0.011)	9.070 (+0.006)
Best	9.157 (-0,310)	5.911 (-0,918)	10.786 (-0,213)	10.624 (-0,264)	7.685 (-0,284)	8.795 (-0,269)

interpreted as twos-complement values, pixels between 0x0000 and 0x7FFF yield decimal values between 0 and 32767. Pixels between 0x8000 and 0xFFFF yield values between -32768 and -1. On the other hand, interpreting the data as unsigned and applying the HST results in values between 0x0000 and 0x7FFF being transformed to values from 0x8000 to 0xFFFF, with decimal equivalents 32768 to 65535. After the JPEG2000 level offset, values from 0 to 32767 are obtained. Similarly, values from 0x8000 to 0xFFFF become 0x0000 to 0x7FFF (or 0 to 32767) after the HST and -32768 to -1 after the level offset. Thus, HST followed by JPEG2000 can be performed by simply applying JPEG2000 to the data as if it were signed, even though it is unsigned. Encoding, decoding and the resulting codestreams are all JPEG2000 Part 1 compliant.

4 Conclusion

DNA microarray images are becoming commonplace for genome-wide monitoring, employed intensively in many medical treatments and biological research. The large size of these images motivates the use of coding techniques to help storing and transmitting them. Lossy coding approaches provide better performance than lossless coding techniques, but they are not always accepted because of the possible negative influence on later classification processes.

Lossless compression results for two different JPEG2000 implementations as well as for other schemes have been discussed. We have tested the performance impact of using different numbers of quality layers and DWT decomposition levels and we have concluded that, for most image sets, the best parameters choice is 1 quality layer and 5 DWT decomposition levels. However, it has been observed that lossless JPEG2000 performance is poor when compared to the best microarray-specific technique, and even to some general image compressors.

A reversible transform based on histogram swapping, which draws images closer to JPEG2000 assumptions for context modeling, has been proposed. With this modifica-

tion, the performance of lossless JPEG2000 compression is improved for all image sets. Rate improvements from 0.213 bpp to 0.918 bpp, corresponding to percentage increases of, respectively, 1.97% and 15.53%, have been measured. The histogram swap transform is easily implemented in a JPEG2000 part 1 compliant manner.

Acknowledgements

MicroZip corpus was kindly provided by Neves and Pinho from the University of Aveiro. The Arizona image set was provided by David Galbraith and Megan Sweeney from the University of Arizona.

This work has been partially supported by the European Union, by the Spanish Government (MICINN), by FEDER, and by the Catalan Government under Grants FP7-PEOPLE-2009-IIF FP7-250420, TIN2009-14426-C02-01, UAB-BI3INT2006-08, and 2009-SGR-1224.

References

[1] S. Moore, "Making chips to probe genes," *IEEE SPECTRUM*, vol. 38, no. 3, pp. 54–60, MAR 2001.

[2] S. Satih, N. Chalabi, N. Rabiau, R. Bosviel, L. Fontana, Y.-J. Bignon, and D. J. Bernard-Gallon, "Gene Expression Profiling of Breast Cancer Cell Lines in Response to Soy Isoflavones Using a Pangenomic Microarray Approach," *OMICS-A JOURNAL OF INTEGRATIVE BIOLOGY*, vol. 14, no. 3, pp. 231–238, JUN 2010.

[3] M. S. Giri, M. Nebozhyn, L. Showe, and L. J. Montaner, "Microarray data on gene modulation by HIV-1 in immune cells: 2000-2006," *JOURNAL OF LEUKOCYTE BIOLOGY*, vol. 80, no. 5, pp. 1031–1043, NOV 2006.

[4] M. Hernández-Cabronero, I. Blanes, M. W. Marcellin, and J. Serra-Sagristà, "Standard and specific compression techniques for dna microarray images," *MDPI Algorithms*, vol. 4, 2012, In press.

[5] N. Faramarzpour, S. Shirani, and J. Bondy, "Lossless DNA microarray image compression," in *In Proceedings of the 37th Asilomar Conference on Signals, Systems and Computers*, vol. 2, November 2003, pp. 1501–1504.

[6] R. Jornsten, Y. Vardi, and C. Zhang, "On the bitplane compression of microarray images," in *In Proceedings of the 4th International Conference on Statistical Data Analysis Based on the L1-Norm and Related Methods*, 2002.

[7] S. Lonardi and Y. Luo, "Gridding and compression of microarray images," in *In Proceedings of the Computational Systems Bioinformatics Conference*. IEEE, 2004, Proceedings Paper, pp. 122–130.

[8] J. Hua, Z. Liu, Z. Xiong, Q. Wu, and K. Castleman, "Microarray basica: Background adjustment, segmentation, image compression and analysis of microarray images," *EURASIP Journal on Applied Signal Processing*, vol. 2004, no. 1, pp. 92–107, January 2004.

[9] R. Bierman, N. Maniyar, C. Parsons, and R. Singh, "MACE: lossless compression and analysis of microarray images," in *In Proceedings of the ACM symposium on Applied computing*, ser. SAC '06. ACM, 2006, pp. 167–172.

[10] A. Neekabadi, S. Samavi, S. A. Razavi, N. Karimi, and S. Shirani, "Lossless microarray image compression using region based predictors," in *In Proceedings of the*

International Conference on Image Processing, vol. 1-7. IEEE, 2007, Proceedings Paper, pp. 913–916.

[11] S. Battiato and F. Rundo, "A bio-inspired CNN with re-indexing engine for lossless dna microarray compression and segmentation," in *In Proceedings of the 16th International Conference on Image Processing*, vol. 1-6, IEEE. IEEE, 2009, Proceedings Paper, pp. 1717–1720.

[12] Y. Luo and S. Lonardi, "Storage and transmission of microarray images," *Drug Discovery Today*, vol. 10, no. 23-24, pp. 1689 – 1695, 2005.

[13] D. S. Taubman and M. W. Marcellin, *JPEG2000: Image Compression Fundamentals, Standards and Practice*. Kluwer Academic Publishers, Boston, 2002.

[14] M. Burrows and D. J. Wheeler, "A block-sorting lossless data compression algorithm." HP, Tech. Rep. 124, 1994.

[15] Y. Zhang, R. Parthe, and D. Adjeroh, "Lossless compression of DNA microarray images," in *In Proceedings of the IEEE Computational Systems Bioinformatics Conference*, August 2005, pp. 128 – 132.

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

[17] A. J. R. Neves and A. J. Pinho, "Lossless compression of microarray images," in *In Proceedings of the International Conference on Image Processing, ICIP*. IEEE, 2006, Proceedings Paper, pp. 2505–2508.

[18] ——, "Lossless compression of microarray images using image-dependent finite-context models," *IEEE Transactions on Medical Imaging*, vol. 28, no. 2, pp. 194–201, February 2009.

[19] Y. Zhang and D. Adjeroh, "Prediction by partial approximate matching for lossless image compression," *IEEE Transactions on Image Processing*, vol. 17, no. 6, pp. 924 –935, June 2008.

[20] "Stanford Microarray Database public FTP." [Online]. Available: ftp://smd-ftp.stanford.edu/pub/smd/transfers/Jenny

[21] "MicroZip test image set." [Online]. Available: http://www.cs.ucr.edu/~yuluo/MicroZip

[22] "Stanford Yeast Cell-Cycle Regulation Project." [Online]. Available: http://genome-www.stanford.edu/cellcycle/data/rawdata/individual.html

[23] "ApoA1 experiment data." [Online]. Available: http://www.stat.berkeley.edu/users/terry/zarray/Html/apodata.html

[24] "ISREC image set." [Online]. Available: http://www.isrec.isb-sib.ch/DEA/module8/P5_chip_image/images/

[25] D. Taubman, "Kakadu Software home page." [Online]. Available: http://www.kakadusoftware.com/

[26] "JJ2000 Software home page." [Online]. Available: http://code.google.com/p/jj2000/

[27] A. Pinho, A. Paiva, and A. Neves, "On the use of standards for microarray lossless image compression," *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 3, pp. 563–566, March 2006.

[28] B. E. Usevitch, "A tutorial on modern lossy wavelet image compression: foundations of jpeg 2000," *IEEE Signal Processing Magazine*, vol. 18, no. 5, pp. 22–35, 2001.