Multicomponent compression of DNA microarray images

Miguel Hernández-Cabronero[1], Francesc Aulí-Llinàs[1], Joan Bartrina-Rapesta[1], Ian Blanes[1], Leandro Jiménez-Rodríguez[1], Michael W. Marcellin[1,2], Juan Muñoz-Gómez[1], Victor Sánchez[1], Joan Serra-Sagristà[1] and Zhongwei Xu[1]

Abstract— **In this work, the correlation present among pairs of DNA microarray images is analyzed using Pearson's r as a metric. A certain amount of correlation is found, especially for red/green channel image pairs, with averages over 0.75 for all benchmark sets. Based on that, the lossless multicomponent compression features of JPEG2000 have been tested on each set, considering different spectral and spatial transforms (DWT 5/3, DPCM, R-Haar and POT). Improvements of up to 0.6 bpp are obtained depending on the transform considered, and these improvements are consistent to the correlation values observed.**

Keywords— **microarray images, microarray image compression, JPEG2000, multicomponent compression**

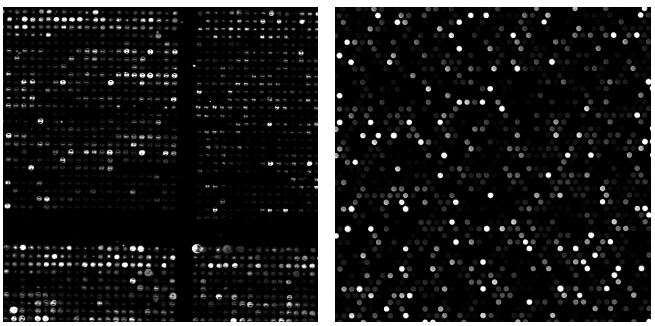
I. Introduction

A. 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 hypotheses about the function and regulation of thousands of individual genes.

[1]Dept. of Information and Communications Engineering, Universitat Autònoma de Barcelona, Barcelona, Spain. Contact e-mail: miguel.hernandez@uab.es

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



(a) (b)

Fig. 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.

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, with sizes over 4000 × 13000 being common nowadays. Since gene expression can vary over a wide range, a high degree of precision is desired, DNA microarray images have a intensity resolution of 16 bits per pixel (bpp).

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 rather than 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. Many DNA microarray studies consist of running several experiments at different moments on similar biological samples that have been exposed to different physical and chemical conditions. 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 re-

quired. 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, but at the cost of poorer compression performance as compared to lossy techniques. The efficient lossless compression of this type of images has proved to be a difficult task. Transform-based coding methods do not perform well on microarray images due to the considerable amount of noise and the abundance of high frequencies present in this type of images [4].

B. 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. For example, 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. On the other hand, the postprocessing stage is always addressed for microarray images, since they are always analyzed to extract genetic data, but is not described in this document. Lossless techniques belonging to the preprocessing, transform and entropy coding stages are addressed next. A more exhaustive review of the state of the art can be found in the literature [5].

B.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 image compression 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 [6]: first, spot regions are located by studying the minimum values of the pixel intensity sum by rows, 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 [7]. Later, in 2004, Lonardi and Luo presented their MicroZip software [4], 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 a segmentation approach based on Cellular Neural Networks (CNNs) [11].

B.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 [4] made use of the Burrows-Wheeler transform [14] for lossy or lossless compression in their MicroZip software. In 2012, we proposed a novel reversible point transform consisting in swapping the left and right halves of the image histogram [15].

TABLE I: Classification of lossless microarray-specific techniques discussed in Subsection I-B, sorted chronologically.

Preprocessing	**Transform**	**Entropy coding**	
Segmentation		Segmentation	Context
[7], 2002	[8], 2004	[6], 2003	[16], 2005
[6], 2003	[4], 2004	[9], 2006	[17], 2006
[8], 2004	[15], 2012	[11], 2009	[18], 2006
[4], 2004			[19], 2009
[9], 2006			
[10], 2007			
[11], 2009			

B.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 coding proposals that are based on segmentation can be found on lossless

compression of DNA microarrays. In 2003, Faramarzpour et al. presented a prediction-based technique [6]. 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. [17], [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 [20]. In 2006, Neves and Pinho [18] 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 [19].

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

C. Paper structure

The rest of this paper is organized as follows. In Section II, the correlation present among microarray images that belong to the same set is analyzed. In Section III, several multicomponent compression experiments are described and discussed. Finally, in Section IV, some conclusions are drawn.

II. Correlation between DNA microarray images

As it was pointed out in Section I, many DNA microarray studies consist of running several experiments at different moments on similar biological samples that have been exposed to different physical and chemical conditions. In addition, each microarray experiment produces two monochrome images which are obtained by scanning the same microarray chip. For this reason, it is natural to assume that some kind of correlation is present among the images produced within a study. Subsection II-A describes the microarray image benchmark sets studied in this paper and Subsection II-B analyzes the correlation present among the images of each set.

TABLE II: Image sets used in the literature.

Image set		Images	Size (px)
Yeast	[21]	109	1024×1024
ApoA1	[22]	32	1044×1041
ISREC	[23]	14	1000×1000
Arizona		6	4400×13800

A. Datasets

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 frequently. Unfortunately, images of the MicroZip set do not have the same size, so they have not been used in our correlation and multicomponent compression experiments. We have included the Yeast set, which has been employed in a few publications. Furthermore, we have gathered additional larger images, closer to what is employed in laboratories today. These images, which come from the Arizona set, have been kindly provided by David Galbraith and Megan Sweeney from the University of Arizona.

Table II 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.

B. Image correlation

In this subsection, the correlation present among images of the same set is analyzed. Two configurations for this experiment have been tested. First, all possible pairs of images from the same set are considered. After that, only red and green channels obtained from the same microarray chip are analyzed. The Pearson product-moment correlation coefficient is the correlation metric used in our experiments. It takes values in $[-1, 1]$ and is defined as

$$r = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2}\sqrt{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}},$$

where X_i and Y_i are the sequences of pixels obtained by scanning two images in the same order, and $\bar{X}$ and $\bar{Y}$ are the average pixel values of the first and second images, respectively.

Figures 2 and 3 show the distribution of the measured Pearson's r values among all pairs and red/green channel pairs, respectively. In these figures it can be observed which ranges of Pearson's r values are more common. Additionally, Tables III and IV display statistical information about each of

the experiment configurations. It can be easily observed that, in general, pairs of images of the same set are not very correlated: most Pearson's r values are under 0.2 except for the Arizona set, which shows larger measurements, but most of them are still below 0.8. On the other hand, when considering only red/green channel pairs from the same chip, correlation is considerably larger. Attending to the average Pearson's r values of the sets, tenfold increases can be observed, while variance is reduced. Comparing the *Max* column in Tables III and IV, we see that the largest correlation value in each set corresponds to a red/green channel pair.

These results show that a certain amount of correlation actually exists among images from the same data set, to a larger extent when considering only red/green channel image pairs, for which average values larger than 0.75 are consistently observed for all sets.

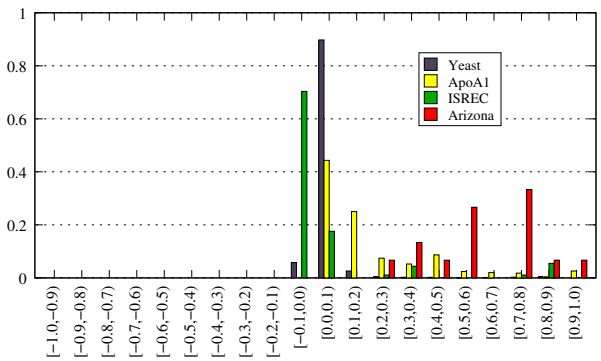


Fig. 2: Pearson's r value distribution for all image pairs from the same chip. The sum of the bars referring to each set is one.

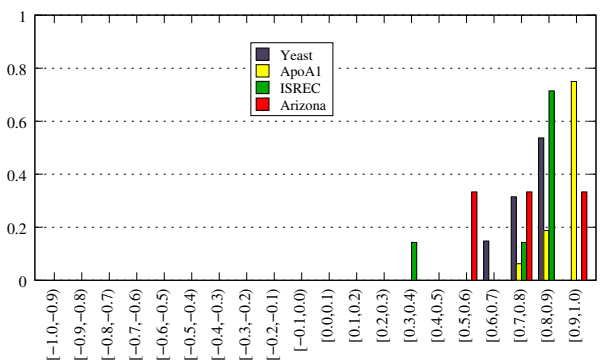


Fig. 3: Pearson's r value distribution for red/green channel image pairs from the same set.

III. Multicomponent compression of DNA microarray images

Attending to the results discussed in Section II, it seems reasonable to exploit the correlation present among pairs of images from the same set, especially among red/green channel image pairs. To do so, we have designed several experiments using lossless JPEG2000, since it is the only progressive lossy-to-lossless scheme supported by the DICOM medical image standard.

TABLE III: Statistical properties of the Pearson's r values obtained comparing all pairs of images in each set. The sum of the bars referring to each set is one.

Set	Max	Min	Average	Variance
ISREC	0.8901	-0.0186	0.0735	0.0501
ApoA1	0.9678	0.0105	0.2043	0.0468
Yeast	0.8961	-0.0329	0.0390	0.0070
Arizona	0.9332	0.2845	0.6153	0.0383

TABLE IV: Statistical properties of the Pearson's r values obtained comparing red/green channel pairs of images in each set.

Set	Max	Min	Average	Variance
ISREC	0.8901	0.3250	0.7822	0.0359
ApoA1	0.9678	0.7969	0.9229	0.0029
Yeast	0.8961	0.6084	0.7821	0.0048
Arizona	0.9332	0.5937	0.7508	0.0195

In these experiments, we have benchmarked the lossless multicomponent compression performance of Kakadu JPEG2000 when applied to DNA microarray images using different spatial and spectral transforms. For the spatial transformations, we have tested different numbers of 5/3 DWT decomposition levels. For the spectral transform, we have also used different numbers of 5/3 DWT decomposition levels as well as the reversible Haar transform (R-Haar), differential pulse code modulation (DPCM) and the pairwise orthogonal transform (POT [24]). Table V shows the average compression performance expressed in bits per pixel when compressing all images of a set together as a single multicomponent image. The order in which the images are arranged for the multicomponent compression affects the performance only to a little extent, generally less than 0.5 bpp. In this table, we show results only for the $R_1G_1 \cdots R_NG_N$ arrangement, where R_i and G_i are the red and green channel images from the i-th pair of a set, respectively. Table VI shows the average results for compressing red/green channels together as a 2-component image. For brevity's sake, these tables display only a representative set of results from all the data obtained in our experiments. A full description of the experiment with data for all tested configurations can be downloaded at `http://deic.uab.es/~mhernandez/media/reports/multicomponent_compression.pdf`. Table VII of this document displays compression performance results of other lossless schemes, including the best-performing microarray-specific technique [11], for ease of reference.

It can be observed that one level of DWT spectral transform does not generally improve compression performance, as compared to using zero wavelet decomposition levels. When it does, the gain does not exceed 0.3 bpp. The POT shows slightly better

results compared to using zero DWT levels for the ApoA1 set, with improvements up to 0.6 bpp. For the Yeast and Arizona sets, the DPCM transform produces improvements of up to 0.4 bpp, better than any of the other transforms. For the ISREC set, no transform is able to improve upon zero DWT decomposition levels. It is also noteworthy that all spectral transforms work better when considering only red/green channel image pairs that when all images in the set are compressed as a single multicomponent image. This is consistent to the results discussed in Section II because red/green channel image pairs exhibit larger amounts of correlation.

IV. Conclusions and future work

In this paper, we have described DNA microarray images and we have motivated the importance of their compression. We have briefly described the state of the art of the lossless compression of this type of imagery. We have analyzed the correlation present on pairs of microarray images from the same set using Pearson's r as a metric. From that we have concluded that there exists a certain amount of correlation among image pairs, especially among green/red channel image pairs. Based on these results, several lossless multicomponent compression tests have been run and described. The DWT 5/3, DCPM, the reversible Haar transform and the POT have been employed using several different configurations in the experiments. No single spectral transform is able to improve upon zero wavelet decomposition levels in the spatial and spectral domain for all the sets, even though the DPCM transform does so except for the ISREC set. For all sets and transforms, the observed compression performance is better when considering only red/green channel image pairs. This is consistent to the correlation values observed.

As future work, we plan to quantitatively analyze the relationship between correlation and multicomponent compression. In addition, we also plan to apply the histogram swap transform [15] together with multicomponent compression.

Acknowledgements

This work has been partially supported by the European Union, by the Spanish Government (MECD) and by the Catalan Government, under Grants TIN2009-14426-C02-01, SGR2009-1224, TEC2010-11776-E, FPU AP2010-0172, China Scholarship Council, UAB-BI3INT2006-08, UAB-472-01-2/09, RYC-2010-05671 and FP7-PEOPLE-2009-IIF FP7-250420.

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] 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.

[5] 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, pp. 30–49, 2012.

[6] 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.

[7] 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.

[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] M. Hernández-Cabronero, J. Muñoz-Gómez, I. Blanes, J. Serra-Sagristà, and M. W. Marcellin, "DNA microarray image coding," in *In Proceedings of the IEEE International Conference on Data Compression, DCC*, IEEE, Ed., 2012.

[16] 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.

[17] 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.

[18] 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.

[19] ——, "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.

[20] 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.

[21] "Stanford Yeast Cell-Cycle Regulation Project: (http://genome-www.stanford.edu/cellcycle/data/rawdata/individual.html)."

[22] "ApoA1 experiment data: (http://www.stat.berkeley.edu/users/terry/zarray/Html/apodata.html)."

[23] "ISREC image set: (http://www.isrec.isb-sib.ch/DEA/module8/P5_chip_image/images/)."

[24] I. Blanes and J. Serra-Sagristà, "Pairwise orthogonal transform for spectral image coding," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 3, pp. 961–972, March 2011.

TABLE V: Average lossless multicomponent compression results in bpp considering all images of each set as a single multicomponent image. Zero spatial DWT decomposition levels are assumed. In every case, using more than zero levels of spatial DWT yields worse performance.

Set	**Spectral transform**	**Spectral levels**	**Bpp**
Yeast	DWT 5/3	0	6.828
Yeast	DWT 5/3	1	8.877
Yeast	POT	1	9.883
Yeast	DPCM	1	6.447
Yeast	R-Haar	1	6.999
ApoA1	DWT 5/3	0	11.524
ApoA1	DWT 5/3	1	11.417
ApoA1	POT	1	11.267
ApoA1	DPCM	1	11.260
ApoA1	R-Haar	1	11.160
ISREC	DWT 5/3	0	10.887
ISREC	DWT 5/3	1	11.932
ISREC	POT	1	11.864
ISREC	DPCM	1	11.870
ISREC	R-Haar	1	11.451
Arizona	DWT 5/3	0	9.548
Arizona	DWT 5/3	1	9.804
Arizona	POT	1	9.556
Arizona	DPCM	1	9.575
Arizona	R-Haar	1	9.629

TABLE VI: Average lossless multicomponent compression results in bpp for red/green channel image pairs. Zero spatial DWT decomposition levels are assumed. In every case, using more than zero levels of spatial DWT yields worse performance.

Set	**Spectral transform**	**Spectral levels**	**Bpp**
Yeast	DWT 5/3	0	6.829
Yeast	DWT 5/3	1	6.786
Yeast	POT	1	9.279
Yeast	DPCM	1	6.439
Yeast	R-Haar	1	6.790
ApoA1	DWT 5/3	0	11.524
ApoA1	DWT 5/3	1	11.217
ApoA1	POT	1	10.956
ApoA1	DPCM	1	11.289
ApoA1	R-Haar	1	11.218
ISREC	DWT 5/3	0	10.887
ISREC	DWT 5/3	1	11.451
ISREC	POT	1	11.468
ISREC	DPCM	1	11.203
ISREC	R-Haar	1	11.452
Arizona	DWT 5/3	0	9.548
Arizona	DWT 5/3	1	9.649
Arizona	POT	1	9.439
Arizona	DPCM	1	9.386
Arizona	R-Haar	1	9.649

TABLE VII: Average compression results in bpp for generic image compressors and individual image compression. Results for the best microarray-specific technique (by Battiato) and the best general compressor are also included at the bottom for ease of reference.

Algorithm	**MicroZip**	**Yeast**	**ApoA1**	**ISREC**	**Arizona**
CALIC	9.582	8.502	10.515	10.615	–
JBIG	9.747	6.888	10.852	10.925	8.858
JPEG-LS	9.441	8.580	10.608	11.145	8.646
JPEG2000 (0 DWT 5/3 levels)	10.063	6.863	11.566	10.930	9.582
JPEG2000 (1 DWT 5/3 level)	9.577	9.128	11.134	11.517	9.253
JPEG2000 (5 DWT 5/3 levels)	9.508	9.082	11.052	11.360	9.099
Battiato [11]	8.369	–	9.52	9.49	–
Bzip2	9.841	6.075	11.067	10.921	8.944