Proceedings of the 28th IEEE
EMBS Annual International Conference
New York City, USA, Aug 30-Sept 3, 2006

Extraction of the Number of Peroxisomes in Yeast Cells by Automated Image Analysis

Antti Niemistö*†, Jyrki Selinummi*, Ramsey Saleem†, Ilya Shmulevich†, John Aitchison†, and Olli Yli-Harja*
*Institute of Signal Processing, Tampere University of Technology, Tampere, Finland
†Institute for Systems Biology, Seattle, Washington, USA

***Abstract*— An automated image analysis method for extracting the number of peroxisomes in yeast cells is presented. Two images of the cell population are required for the method: a bright field microscope image from which the yeast cells are detected and the respective fluorescent image from which the number of peroxisomes in each cell is found. The segmentation of the cells is based on clustering the local mean-variance space. The watershed transformation is thereafter employed to separate cells that are clustered together. The peroxisomes are detected by thresholding the fluorescent image. The method is tested with several images of a budding yeast *Saccharomyces cerevisiae* population, and the results are compared with manually obtained results.**

***Index Terms*— Yeast, peroxisome biogenesis, image analysis, quantification, segmentation, watershed transformation**

I. Introduction

Due to its relative simplicity, the budding yeast *Saccharomyces cerevisiae* is often used as a model organism. The sequence of *S. cerevisiae* was completed in 1996 [1] and it is easy to produce knockout strains. This enables the study of the global regulation of different properties of yeast. For example, the regulation of yeast morphology was considered in [2] with the emphasis on developing an image analysis method for automated extraction of yeast cell morphology. Similar methods can also be used in cell cycle studies in which the aim is to invert the population effects [3]–[5]. In [6] knockout mutants and image analysis methods were used to study endocytosis in budding yeast.

It is also possible to study how different genes control the organelles of yeast cells. In this paper we consider peroxisomes in *S. cerevisiae*. Peroxisomes are inducible, single membrane organelles, involved in a variety of processes that include beta-oxidation of fatty acids, cholesterol biosynthesis, and amino acid metabolism. The underlying idea of our study is to relate different gene knockouts to peroxisome biogenesis. For example, deletion of genes involved in the biogenesis of peroxisomes would show a paucity of peroxisomes. The fundamental property that one needs to extract is the number of peroxisomes in each cell. To be able to do this by image analysis, one needs to obtain two microscope images: a bright field image from which the cells are detected and a corresponding fluorescently stained image from which the peroxisomes are detected.

A study may include from several hundred to a few thousands knockout strains [7]. For the results to be statistically meaningful, one needs to have several images for each strain. Thus, there can be tens of thousands images in total, and extracting the number of peroxisomes per cell manually is not feasible. Instead, a fully automated high-throughput image analysis method is needed. Making the analysis fully automated has the further advantage that the quantification results are guaranteed to be objective and reproducible. It also enables one to consider additional features such as peroxisome size and the intensity of the fluorescent marker.

This paper concentrates on developing the automated image analysis methods that are needed for quantification of peroxisomes in yeast. The methods are described with the aid of a representative example image pair in Section II. Section III shows the quantification results for a larger image set, and the validation of the results is also considered. The concluding remarks are given in Section IV.

II. Methods

A. Experimental procedure and image acquisition

A peroxisomal matrix protein (POT1) was C-terminally tagged with an enhanced copy of GFP (EGFP) via homologous recombination in the Y8205 SGA starting strain (mat alpha) [8]. The POT1GFP strain was grown in YEPD (1% yeast extract, 2% peptone, and 2% glucose) overnight, then transferred to SCIM medium (0.7% yeast nitrogen base, 0.5% yeast extract, 0.5% peptone, 0.5% Tween 40, 0.79 g/L of complete synthetic medium (Qbiogene, Morgan Irvine, CA, USA), 0.5% $(NH_4)_2SO_4$, 0.1% glucose, and 0.15% oleic acid) and induced for 4 hours. Post induction, cells were fixed with 3.7% formaldehyde for 30 minutes, then washed three times with phosphate buffered saline solution.

The cells and peroxisomes were visualized using a Leica confocal microscope. Two images were obtained: a bright field image and a fluorescent counterpart that shows the peroxisomes. An example bright field image is shown in Fig. 1 and an example fluorescent image is shown in Fig. 2. Both images are 8-bit grayscale images. Fluorescent images utilized an argon laser at 488 nm at 50% power, while transmission imaging was used for the bright field images. The bright field image is used to detect the boundary of each cell. This defines the region of interest for peroxisome detection from the fluorescent images.

B. Pre-processing

The illumination of the bright field image may not be uniform. We assume the additive nonuniformity model [9]

$$f(m,n) = u(m,n) + \xi(m,n), \tag{1}$$

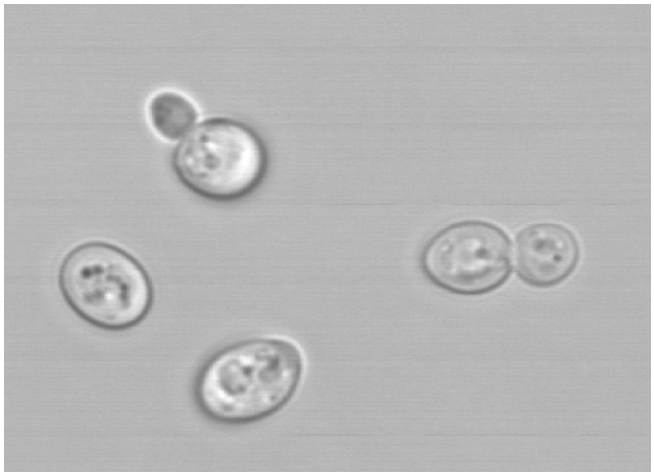

Fig. 1. A bright field microscope image of a budding yeast cell population.

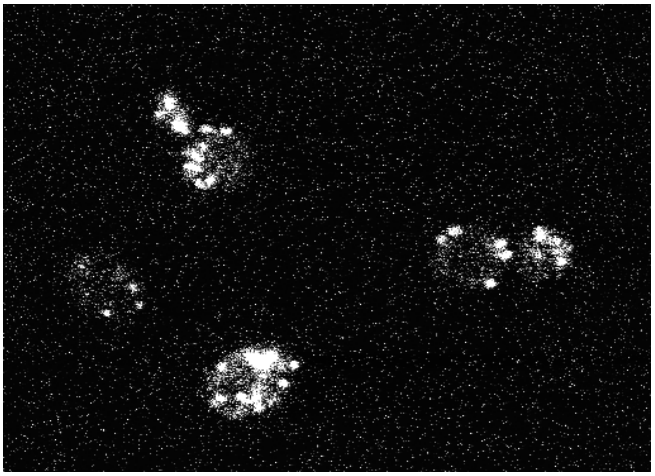

Fig. 2. The fluorescent image corresponding to the image in Fig. 1 showing the peroxisomes marked with a green fluorescent protein (GFP).

where f is the observed image, u is the uniformly illuminated image, and ξ is the nonuniformity component. We estimate ξ by fitting a second degree polynomial surface on the image. The fit is done robustly using M-estimators [10]. The illumination nonuniformity is thereafter corrected by simply solving u from the model (1).

The fluorescent images contain impulsive noise, which is removed with the 3×3 median filter [11]. The result is shown in Fig. 3.

C. Segmentation

Since we want to find the number of peroxisomes in each cell, the first task is to detect the cells from the bright field image. It should be noted that the cells may be clustered together, and we need to detect individual cells from such clusters. A cell may also have a bud. In this paper we treat the buds as separate cells, and separate them from the parent cells just like we separate cells that are otherwise clustered together.

In the bright field image, the cell membranes are clearly visible as elliptic regions that are darker than the background. The cell membrane pixels thus have two useful properties: the local mean is low and the local variance is high. Segmentation can thus be achieved by marking as cell membrane pixels those pixels whose local mean is below a threshold T_μ and whose variance is above a threshold T_σ. Segmentation is thereafter completed by a simple flood fill of the areas inside the detected cells. The threshold T_μ can be determined based on the global mean and variance of the local mean image. Similarly, the threshold T_σ can be determined based on the global mean and variance of the local variance image. This method has been previously employed for segmentation of yeast cells in [3]. After segmentation there may be some objects whose size is so small that the object cannot be a cell, and these small objects are removed. We also remove all objects that touch the edges of the image. The obtained segmentation result is shown in Fig. 4.

In order to enable individual analysis of each cell, cells that are clustered together must be separated from each other. The separation of clustered or overlapping objects has been extensively studied (see, e.g., [12]–[14]), but no method has proven to be superior in all situations. Since yeast cells are approximately convex, we can apply a shape-based watershed separation method [15]. First, the Euclidean distance transformation [16] of the complement of the binary segmentation result is calculated. In this image there should ideally be exactly one local maximum for each cell. However, irregularities in the shapes of the cells tend to result in several local maxima for each cell. Therefore, to prevent over-segmentation, we suppress the irrelevant maxima with the h-maxima transformation [15]. Since the watershed transformation begins to flood the surface from each local minimum, we apply it to the complement of the h-maxima transformed image. This image can be considered as a topographic surface, and watershed lines get placed where ever water from two separate flooding points meet. The watershed lines are then used to separate the clustered cells. The result for our test image is shown in Figure 5. Each cell is shown with a different color.

The second task is to detect the peroxisomes from the fluorescent image. Peroxisomes can be clustered together, and peroxisomes from such clusters should be detected individually. However, peroxisomes may have different shapes. For example, they can be circular or elongated, so it may be difficult to distinguish two partially overlapping peroxisomes from one elongated peroxisome. However, as long as the detection is done consistently using the same methods, small errors in the quantification do not matter, because ultimately the goal is to quantify the relative differences between wild type and mutant strains.

The peroxisomes are represented by pixels with the highest intensities in the fluorescent image. We detect the peroxisomes by thresholding the median filtered image (see Fig. 3). The threshold selection depends on the imaging arrangements, and can be done, for example, with Otsu's method [17]. In our case, the peroxisome pixels are consistently very close to saturation in all of the test images.

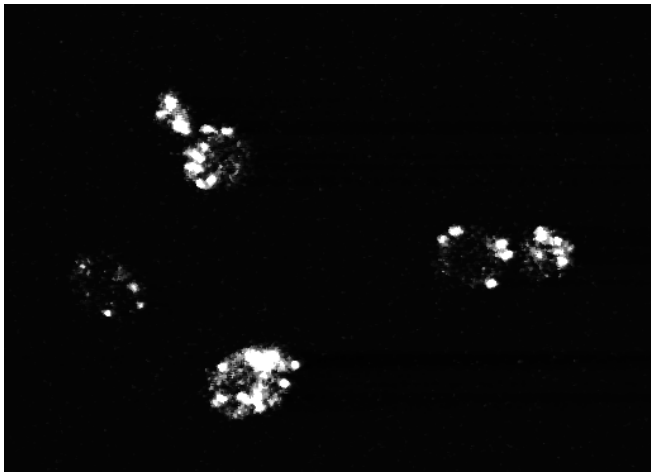

Fig. 3. The fluorescent peroxisome image after filtering with the 3×3 median filter.

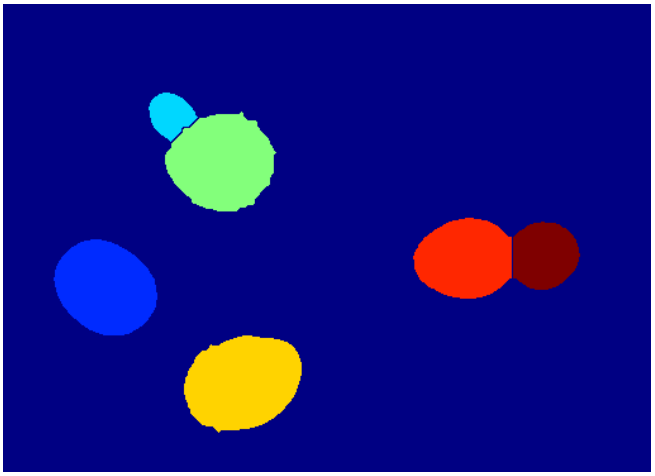

Fig. 5. The segmentation result for the bright field image after separation of clustered cells.

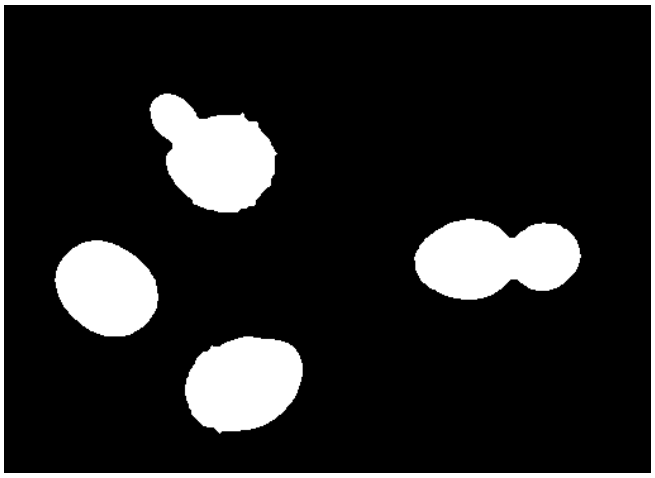

Fig. 4. The segmentation result for the bright field image.

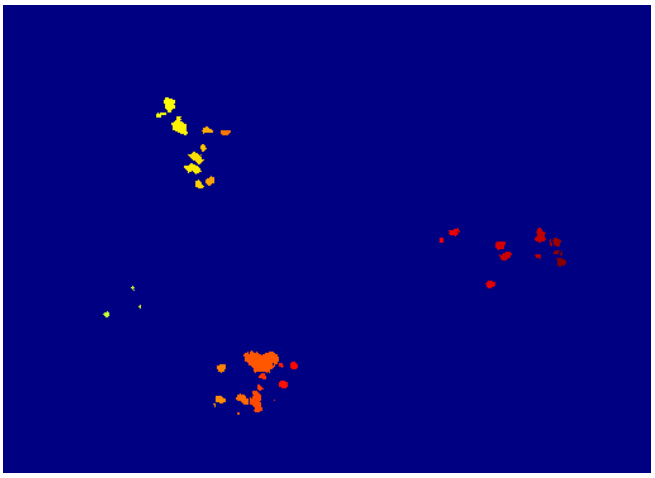

Fig. 6. The segmentation result for the fluorescent image.

Therefore, we use the threshold value 240. The watershed separation method is thereafter applied to the binary image in order to separate peroxisomes that are clustered together. The h-maxima transformation is applied to avoid over-segmentation. The result is shown in Fig. 6. Each peroxisome is shown with a different color.

D. Feature extraction

The number of peroxisomes for each cell can be easily obtained from the labeled images shown in Figs. 5 and 6. Each cell in the segmentation result of the bright field image gives us a region of interest, and we find the number of peroxisomes in that region by looking at the segmentation result of the fluorescent image. If a peroxisome extends into two or more regions of interest, it is counted as belonging to the cell that has the majority of the peroxisome area.

In the example image there are 6 cells and they are detected correctly by our method. The detected numbers of peroxisomes in the cells are 3, 3, 5, 7, 7, and 13. The number of peroxisomes is difficult to assess manually, especially in the case of the cell having the largest number of peroxisomes. However, the detected numbers of peroxisomes seem to be correct. In total there are 38 peroxisomes in the 6 cells, making the mean number of peroxisomes per cell 6.33.

III. Results and Validation

The presented image analysis method was tested with 25 test image pairs taken of the same yeast strain. The automatically obtained number of cells in each image was compared with manual cell counts. The results are illustrated as a scatter plot in Fig. 7. In most cases the automated counts given by the presented image analysis method differ from the manual counts by at most one cell. More significant inaccuracies occur only when there are a large number of cells in the image. In those cases large numbers of cells are clustered together, and the watershed method is unable to separate them all.

Because of the large amount of noise in the fluorescent images, the manual counting of the numbers of peroxisomes in the cells is difficult and tedious. However, we have obtained the manual counts for 30 randomly selected cells. The results are illustrated as a scatter plot in Fig. 8. It can be

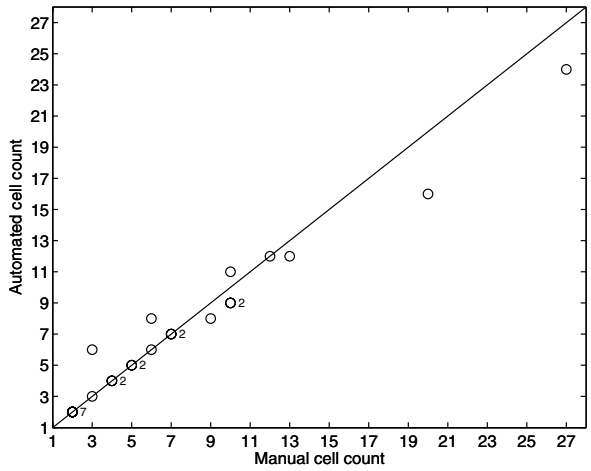

Fig. 7. The scatter plot of manual cell counts vs. automated cell counts for 25 test images.

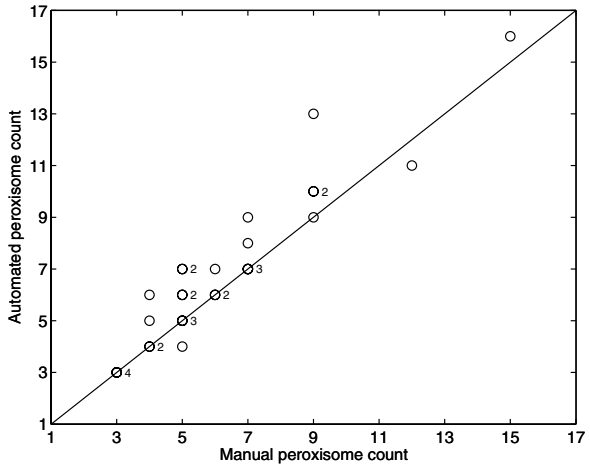

Fig. 8. The scatter plot of manual peroxisome counts vs. automated peroxisome counts for 30 cells.

seen that the automated counts are very close to the manual counts. However, it should be emphasized that the manual counts may contain errors. In fact, it may even be the case that the manual count for a given cell is incorrect and the automated count is correct. The automated method seems to have a tendency to produce higher peroxisome counts than manual counting.

IV. Conclusion

An automated image analysis method for extracting the number of peroxisomes in yeast cells was presented. The comparisons of the results obtained with the proposed method and the manually obtained results indicate that the method is accurate. In the detection of the cells, some problems with the results are observed when a large number of cells are clustered together. However, such heavy clustering can usually be avoided in practice. In the peroxisome counts there seems to be a consistent bias towards higher peroxisome counts. However, such a consistent bias is acceptable in comparative studies in which the interest is to find the relative differences between peroxisome counts in different yeast strains.

Future work will include the extraction of further features of the detected peroxisomes. These include the size and intensity of the peroxisomes as well as their shape. The extraction of these extra features will enable us to use them as inputs for a classifier. In principle the extraction of these features can be easily done based on the work presented here. However, the extraction of these parameters is very sensitive to errors in the segmentation of peroxisomes, and is thus much more difficult than merely obtaining the number of peroxisomes. Another future direction is to utilize the third dimension by obtaining a series of images through the z-axis of each sample.

Acknowledgment

This work was supported by the Academy of Finland, project No. 213462 (Finnish Centre of Excellence programme 2006 – 2011).

References

[1] A. Goffeau, *et al.*, "Life with 6000 genes," *Science*, vol. 274, no. 5287, pp. 563–567, Oct. 1996.

[2] M. Ohtani, A. Saka, F. Sano, Y. Ohya, and S. Morishita, "Development of image processing program for yeast cell morphology," *J. Bioinform. Comput. Biol.*, vol. 1, no. 4, pp. 695–709, Jan. 2004.

[3] A. Niemistö, T. Aho, H. Thesleff, M. Tiainen, K. Marjanen, M.-L. Linne, and O. Yli-Harja, "Estimation of population effects in synchronized budding yeast experiments," in *Image Processing: Algorithms and Systems II*, ser. Proc. SPIE, vol. 5014, 2003, pp. 448–459.

[4] H. Lähdesmäki, T. Aho, H. Huttunen, M.-L. Linne, J. Niemi, J. Kesseli, R. Pearson, and O. Yli-Harja, "Estimation and inversion of the effects of cell population asynchrony in gene expression time-series," *Signal Processing*, vol. 83, no. 4, pp. 835–858, Apr. 2003.

[5] Z. Bar-Joseph, S. Farkash, D. K. Gifford, I. Simon, and R. Rosenfeld, "Deconvolving cell cycle expression data with complementary information," *Bioinformatics*, vol. 20, no. Supplement 1, pp. i23–i30, Aug. 2004.

[6] M. Kaksonen, C. P. Toret, and D. G. Drubin, "A modular design for the clathrin- and actin-mediated endocytosis machinery," *Cell*, vol. 123, no. 2, pp. 305–230, Oct. 2005.

[7] R. Narayanaswamy, W. Niu, A. D. Scouras, G. T. Hart, J. Davies, A. D. Ellington, V. R. Iyer, and E. M. Marcotte, "Systematic profiling of cellular phenotypes with spotted cell microarrays reveals mating-pheromone response genes," *Genome Biol.*, vol. 7, no. 1, Jan. 2006.

[8] A. H. Y. Tong, *et al.*, "Global mapping of the yeast genetic interaction network," *Science*, vol. 303, no. 5659, pp. 808–813, Feb. 2004.

[9] D. Tomazevic, B. Likar, and F. Pernus, "A comparison of retrospective shading correction techniques," in *Proceedings of the 15th International Conference on Pattern Recognition (ICPR'00)*, vol. 3, Barcelona, Spain, Sept. 3–8, 2000, pp. 564–567.

[10] F. R. Hampel, E. M. Ronchetti, P. J. Rousseeuw, and W. A. Stahel, *Robust Statistics: The Approach Based on Influence Functions*. New York: John Wiley & Sons, 1986.

[11] J. Astola and P. Kuosmanen, *Fundamentals of Nonlinear Digital Filtering*. Boca Raton, Florida: CRC Press, 1997.

[12] G. Cong and B. Parvin, "Model-based segmentation of nuclei," *Pattern Recognit.*, vol. 33, no. 8, pp. 1383–1393, Aug. 2000.

[13] Q. Yang and B. Parvin, "Harmonic cut and regularized centroid transform for localization of subcellular structures," *IEEE Trans. Biomed. Eng.*, vol. 50, no. 4, pp. 469–475, Apr. 2003.

[14] L. Vincent, "Morphological grayscale reconstruction in image analysis: applications and efficient algorithms," *IEEE Trans. Image Processing*, vol. 2, no. 2, pp. 176–210, Apr. 1993.

[15] P. Soille, *Morphological Image Analysis: Principles and Applications*, 2nd ed. Heidelberg, Germany: Springer Verlag, 2003.

[16] G. Borgefors, "Distance transformations in digital images," *Comput. Vision Graph. Image Process.*, vol. 34, no. 3, pp. 344–371, June 1986.

[17] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst., Man, Cybern.*, vol. 9, no. 1, pp. 62–66, Jan. 1979.