Image analysis for gene expression based phenotype characterization in yeast cells

Tleis, M.

Summary

Image analysis of objects in the microscope scale requires accuracy so that measurements can be used to differentiate between groups of objects that are being studied. This thesis deals with measurements in yeast biology that are obtained through microscope images. We study the algorithms and workflow of image analysis of yeast cells in order to understand and improve the measurement accuracy. The *Saccharomyces cerevisiae* cell is widely used as a model organism in the life sciences. It is essential to study the gene and protein behaviour within these cells, and consequently making it possible to find treatment and solutions for genetic and hereditary diseases. This is possible since many processes that occurs at the molecular level in this organism are similar to those in human cells.

In the research group Imaging and Bioinformatics, we have developed a framework for analysis of yeast cells. This framework is intended to serve as a support for research in yeast biology. The framework is integrated in one application and presented via a *GUI*. The application integrates modules and algorithms including segmentation, measurement, analysis and visualization.

In **Chapter 1**, we present our research objective: i.e. the problem on how pattern recognition systems can support objective analysis and phenotype characterization of single-cell in image-based gene expression experiments. In addition, we offer a basic background and definitions necessary to follow up in this thesis. Specifically, we introduce the necessary background about cytomic studies and pattern recognition methods followed within our research.

In **Chapter 2**, we present a complete framework for image based experimental read-out in yeast. This framework demonstrates how an automated platform based on a complete image analysis pipeline can assist biologists in their experiments. Moreover, this chapter discusses the individual modules that are integrated in the complete framework: i.e. including the segmentation, measurement, data analysis as well as the GUI that combines these modules together.

In **Chapter 3**, we discuss our novel approach to segmentation based on Hough transform and minimal path algorithms. We show how these can improve the segmentation of ovoid objects, i.e. yeast cells. We start by defining Hough transform and minimal path algorithms. Subsequently we present our general approach to detect ovoid objects in microscope images by detecting circular arcs using a variety of the Hough transform. In addition, we discuss the application of minimal path algorithms to extract the exact contour of detected objects from a polar representation of the image surrounding the object. Furthermore, this chapter presents an additional novel algorithm to expand the extracted contours of ovoid objects. Such expansion is

sometimes necessary settings due to the inherent fuzzy nature of edges and delicate microscope settings. This chapter explains how the polar representation of images is used to expand the initially detected contours by applying circular shortest paths. In addition, it explains the three introduced parameters to control the expansion process. These parameters are *resistance, limit* and *convergence*. Finally, results and comparison with other methods are evaluated using a dataset of *S. cerevisiae* cells.

In **Chapter 4**, we specifically address machine learning where we introduce features to be used in a machine learning approach to automatically identify cell groups cultivated in two different media. We use the same approach to classify cell objects from atrefacts. First we discuss the feature extraction techniques including first-order histogram features, texture measurement, moment invariants, co-occurrence matrix based features and multi-scale wavelet-based texture measurement. Subsequently, various classification methods are evaluated to build a model imported into the yeast analysis platform. This model is trained for the automatic discrimination of cell groups. This discrimination can be used to show that there are different gene expression patterns between cells cultivated under different stress levels. Moreover, the same classification methods are evaluated to build another model for the identification of cell objects. This model is used to discriminate the segmented objects in images into intact cell objects or artefacts such as debris and dead cells.

In **Chapter 5**, the designed image analysis platform is used in a case study to determine the effect of sodium chloride on 14-3-3 genes including *Bmh1* and *Bmh2* in addition to the *Nha1* antiporter. The study also includes a mutant of *BMH1 (Δbmh1)* to study the expression of *Nha1* under different osmotic stress levels. The result obtained from using the yeast analysis software is also validated with that obtained from flow cytometry.

In **Chapter 6**, we present conclusions and lessons learned from this research. Subsequently, we give scope for further research and applications.