TAMPEREEN TEKNILLINEN YLIOPISTO
TAMPERE UNIVERSITY OF TECHNOLOGY

TEPPO ANNILA

DETECTION OF FLUORESCENTLY LABELED PARTICLES IN ESCHERICHIA COLI

Master of Science Thesis

Examiner: Assoc. Prof. Andre Ribeiro
The examiner and topic of the thesis were approved by the Council of the Faculty of Computing and Electrical Engineering on 8th of November 2015

ABSTRACT

ANNILA, TEPPO: Detection of fluorescently labeled particles in *Escherichia coli*
Tampere University of Technology
Master of Science Thesis, 49 pages + 3 Appendix pages
November 2015
Master's Degree Programme in Information Technology
Major: Signal Processing
Examiner: Assoc. Prof. Andre Ribeiro
Keywords: Image analysis, spot detection, scale-space, local thresholding, fluorescence microscopy, *Escherichia coli*

Escherichia coli are one of the most commonly used bacteria to study important biological processes such as transcription and translation. This is due to its simple structure and gene expression system, as well as the easiness to maintain live cultures in a laboratory environment. Due to recent developments in fluorescence microscopy and fluorescence labeling, it is now possible to study such biological processes in live cells at single cell and single molecule level. When analyzing such biological processes, the detection of fluorescent objects and subcellular particles is usually one of the first tasks providing important information for subsequent data analysis.

Although many algorithms have been proposed for the task, it still remains a challenge due to the limitations of image acquisition when imaging live cells. For example, the intensity of the illumination light and the exposure time is usually minimized to prevent damage to the cells, resulting in images with low signal-to-noise ratio. Due to this and the large amount of data typically used for these studies, automated, high quality particle detection algorithms are needed.

In this thesis, we present a novel method for detecting fluorescently labeled subcellular particles in *Escherichia coli*. The proposed method is tested in both synthetic and empirical images and is compared to previous, commonly used methods using standard performance evaluation metrics. The results indicate that the proposed algorithm has a good performance with all image types tested and that it outperforms the previous methods. It is also able to achieve good results with other types of cells than *E. coli*. Moreover, it allows a robust detection of particles from low signal-to-noise ratio images with good accuracy, thus providing accurate and unbiased results for subsequent analysis.

TIIVISTELMÄ

ANNILA, TEPPO: Fluoresoivien partikkeleiden havaitseminen kolibakteereissa
Tampereen teknillinen yliopisto
Diplomityö, 49 sivua + 3 liitesivua
Marraskuu 2015
Tietotekniikan koulutusohjelma
Pääaine: Signaalinkäsittely
Tarkastaja: Assoc. Prof. Andre Ribeiro
Avainsanat: Kuvankäsittely, partikkeleiden havaitseminen, skaala-avaruus, paikallinen kynnystys, fluoresenssimikroskopia, kolibakteeri

Kolibakteeri on yksi käytetyimmistä bakteereista biologisten prosessien, kuten transkription ja translaation tutkimisessa muun muassa sen yksinkertaisen rakenteen ja geeni-ilmentymisjärjestelmän johdosta. Viimeaikainen kehitys niin fluoresenssimikroskopian kuin fluoresoivien proteiinien saralla on tehnyt mahdolliseksi kyseisten prosessien tutkimisen yksittäisten solujen ja molekyylien tasolla. Näissä tutkimuksissa yksi ensimmäisistä tehtävistä on fluoresoivien kappaleiden ja solunsisäisten partikkelien havaitseminen tarjoten tärkeää tietoa datan analysoimiseksi pidemmälle.

Vaikka monia algoritmeja onkin ehdotettu kyseiseen tehtävään, se on yhä haasteellista johtuen elävien solujen kuvantamiseen liittyvistä rajoituksista. Jotta soluja ei mittausten aikana vahingoitettaisi, esimerkiksi herätevalon intensiteetti ja valotusaika pyritään usein minimoimaan, mikä johtaa kohinaisiin kuviin. Kun tämä yhdistetään suureen kuvien määrään, on selvää, että automaattisia ja korkealaatuisia partikkelin havaitsemisalgoritmeja tarvitaan.

Tässä työssä esittelemme uuden tavan solunsisäisten fluoresoivien partikkelien havaitsemiseksi kolibakteereissa. Menetelmä on testattu sekä synteettisillä että oikeilla fluoresenssikuvilla ja vertailtu muiden usein käytettyjen menetelmien kesken käyttäen tavallisimpia suorituskyvyn mittareita. Tulokset osoittavat, että ehdotettu menetelmä toimii hyvin kaikilla testatuilla kuvilla ja se suoriutuu tehtävästä paremmin kuin yksikään muu testattu algoritmi. Kolibakteereiden lisäksi se saavutti hyviä tuloksia myös erityyppisten solujen kanssa. Menetelmän avulla on mahdollisuus saavuttaa tarkkoja ja puolueettomia tuloksia sen hyvän tarkkuuden ja vakaan suorituskyvyn vuoksi, eritoten kohinaisten kuvien yhteydessä.

PREFACE

This work was conducted when working as a research assistant in the Laboratory of Biosystem Dynamics, Department of Signal Processing, Tampere University of Technology.

First, I would like to thank Assoc. Prof. Andre Ribeiro, my thesis supervisor, for giving me this great opportunity and giving me guidance throughout the years. Moreover, I would like to thank him for his valuable comments which improved the quality of the thesis. I also would like to thank all my colleagues whom I have had a pleasure to work with. Especially, I would like to thank Ramakanth Neeli-Venkata for the discussions and comments on biology and Jarno Mäkelä for invaluable discussions regarding the microscopy, imaging and science in general. I am especially grateful for Eero Lihavainen for his comments and relentless guidance in the field of the image analysis.

Finally, I wish to thank my girlfriend Elisa for her continuous support over the years.

Tampere, November 24, 2015

Teppo Annila

CONTENTS

ABBREVIATIONS

CCD	Charged-couple detector
DNA	Deoxyribonucleic Acid
E. coli	*Escherichia coli*
FN	False negative
FP	False positive
FPR	False positive rate
FROC	Free-response receiver operating characteristic
GFP	Green fluorescent protein
HILO	Highly inclined and laminated optical sheet
IPTG	Isopropyl β–D-1-thigalactopyranoside
OD600	Optical density at a wavelength of 600 nm
RBS	Ribosome Binding Site
RNA	Ribonucleic Acid
ROC	Receiver operating characteristic
SNR	Signal-to-noise ratio
TIRF	Total internal reflection fluorescence
TN	True negative
TP	True positive
TPR	True positive rate
YFP	Yellow fluorescent protein

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

Recent developments in fluorescence microscopy and fluorescent labeling techniques have made possible the visualization of subcellular components inside living cells with single-molecule precision and over time. *Escherichia coli* is one model organism where these observations are being conducted at the single-cell level to study, e.g., gene expression and other processes essential for life [1]–[4]. In these studies, the 'objects' of interest are fluorescently labeled so as to emit a detectable signal above the background in the resulting images. To extract information from such data, image analysis is needed, and particle detection is usually one of the first stages of this analysis. Relevantly, the success of most such studies depends on how accurately the subcellular particles are detected.

The developments in data acquisition and microscopy techniques have also increased the amount of data that can be collected. This allows obtaining more solid scientific results but it makes the use of manual detection techniques no longer feasible. Further, the use of automated methods allows more reliable comparison of results between independent measurements and between different studies of the same process. Moreover, since these studies are nowadays conducted mostly using living cells, the amount of light used in the microscopy has to be limited, to minimize damage to the cells. As a consequence, images usually tend to have a low signal-to-noise ratio, thus requiring high quality particle detection algorithms to cope with noisy, heterogeneous backgrounds.

The detection of subcellular objects is important in many applications, varying from particle tracking to basic object detection. In all these applications, the performance of the used method is crucial. For example, the detection accuracy is especially crucial in object tracking algorithms consisting of separate detection and linking stages [5], [6]. Poor detection results might lead to nonsensible tracks if true objects are linked with the false positives or the tracks are terminated due to objects that are not detected. The detection also affects the following data analysis: if the detection rate of the algorithm is poor, the results are biased towards objects that are clearly distinguishable. On the other hand, if the algorithm has a good detection rate but it includes large number of false positives, the results are again biased, due to the presence of false objects.

There are already multiple algorithms proposed for detecting fluorescently labeled particles in living cells, e.g. see [7]–[12]. As reported in [7], all those algorithms are able to perform well with high signal-to-noise ratio images. However, when the quality of the

images decreases the performance of the algorithms rapidly decreases as well, hindering the data analysis. Also, it was shown that all the algorithms were sensitive to data, i.e. none of the proposed detectors worked well with all data.

In this thesis, we propose a novel spot detection method which outperforms several previous methods as described in [8], [13], [10]. Moreover, the method's good performance with low signal-to-noise ratio systems, such as Tsr-Venus, makes it useful for studying the dynamics of gene expression from time-lapse images where the detection accuracy is crucial in order to not bias the results. To evaluate the performance of the various algorithms, we first use simulated images where the background and noise parameters are known and can be controlled. To evaluate the performance in realistic situations, we use empirical image data acquired by fluorescent microscopy. Namely, we have imaged two different fluorescently labeled particles in live cells. These have different characteristics in that one consists of an RNA molecule bound by multiple MS2 coat proteins, each fused with a green fluorescent protein (GFP), while the other consists of aggregates of a yellow fluorescent protein (YFP) variant, Venus, each fused with a Tsr membrane protein.

The results presented in this thesis are partly from a project done in collaboration with fellow research group members of the Laboratory of Biosystem Dynamics of the Department of Signal Processing. We have now submitted part of the results, concerning the new methods used, in the 7th International Conference on Bioinformatics Models, Methods and Algorithms (BIOSTEC 2016). In addition, we are submitting another part of the results, concerning the biological findings, in 10th International Conference on Practical Application of Computational Biology & Bioinformatics (PACBB 2016). Finally, we are currently finishing a work, to be submitted in top scientific journal in the field of Single Cell Biology, which we expect to be accepted for publication by March 2016.

We begin by explaining the biological background and framework of the thesis. In the third chapter, the traditional framework for detecting fluorescently labeled particles with basic image analysis operations is presented. The framework is further illustrated with examples of previous methods. Also, various techniques for validation of the detectors are covered. In the fourth chapter, the generation of simulated images and the necessary protocol for acquiring the empirical images are explained together with the proposed algorithm for subcellular spot detection. In chapter 5, the performance of the proposed algorithm is then analyzed and compared to several other spot detection algorithms by using both synthetic and empirical images. Finally, the conclusions and final discussion are presented in chapter 6.

2. BIOLOGICAL BACKGROUND

In this chapter, we describe the biological background of the work. First, the organism used in the study, *Escherichia coli*, is described. Then, the basic principles of the process of gene expression in bacteria are explained. Finally, the basics of fluorescence microscopy are explained together with fluorescent labeling, which are then combined so as to form a visualization technique of intracellular components in *E. coli*.

2.1 *Escherichia coli*

Escherichia coli is a rod-shaped bacterium with a typical length of 2µm, commonly found in the environment and in the intestine of humans, animals and other warm-blooded organisms. It is a prokaryote and, thus, has no membrane enclosed compartment to house DNA (Deoxyribonucleic acid), as opposite to eukaryotes. It is protected by a tough outer cell wall followed by a periplasmic space and a plasma membrane that encapsulates all the material (DNA, RNA (Ribonucleic acid), proteins, ribosomes and other molecules) in a single compartment consisting mainly of cytoplasm. The genetic material is localized in a specific structure within the cytoplasm called nucleoid. Similarly to having a simpler inner structure than eukaryotes, their *gene expression* system, which "converts" the genetic information in the DNA into functional proteins, is also simpler when compared to the complex process occurring in eukaryotes. [14] The ability of these cells to reproduce repeatedly, while under microscope observation, by elongating, forming a wall at midcell along the major cell axis, and then dividing in two symmetrical daughter cells allows studying subsequent generations in a short-time period. For the above reasons, *E. coli* is considered as a model organism to study cellular processes, and has been widely used in several studies, including of synthetic genetic circuits, aging and gene expression [1], [15]–[17].

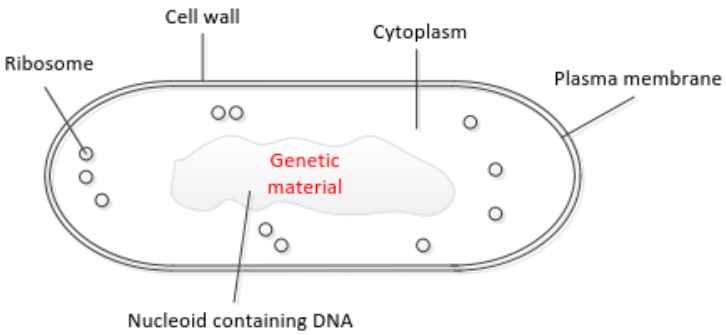


Figure 1: *The structure of Escherichia coli.*

2.2 Gene expression

Advances in fluorescence microscopy, fluorescent labeling techniques and image analysis have enabled studying the dynamics of gene expression in live *E. coli* cells with single molecule precision. Gene expression can be described as a process where the genetic information encoded by the DNA is first transcribed into RNA and then translated into functional proteins. It consists of two main steps, transcription and translation, which together form the Central Dogma of Molecular Biology (Figure 2). The detection of fluorescent particles inside cells was an important step for studying these processes since, for example, it allows the estimation of the number of RNAs or proteins from the spots intensities [2], [18], which gives us information on transcription and translation.

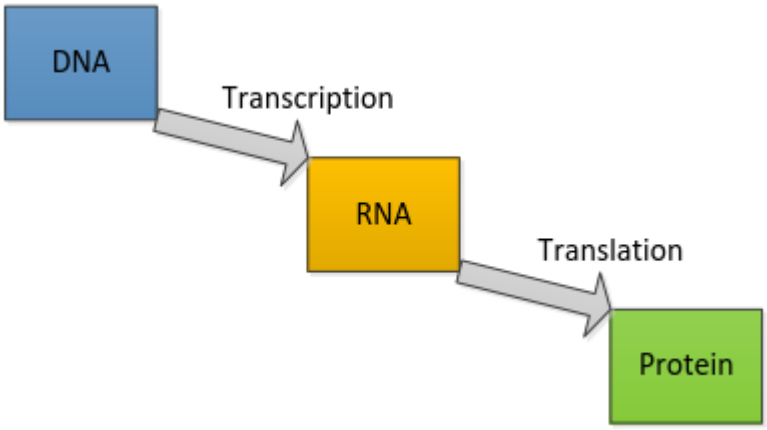


Figure 2: *The Central Dogma of Molecular Biology: the genetic information in DNA is transcribed into RNA which is then translated into proteins.*

The genetic information in DNA is encoded in the sequence of nucleotides consisting of four different bases, namely adenine (A), thymine (T), guanine (G) and cytosine (C). The bases always pair together according to base pairing rules (A with T and G with C) to form a double stranded DNA where the order of the bases specifies the effect of a gene (particular sequence in DNA). [14] This information is then used when creating proteins through the processes of transcription and translation.

In the first step of the gene expression, DNA is transcribed into RNA. Transcription consists of three main steps: initiation, elongation and termination. In transcription initiation, an RNA polymerase binds to a particular region in a gene called the promoter region, which indicates the starting site of the transcription for a specific gene. During the elongation phase, the RNA polymerase unwinds a part of DNA and copies the nucleotides from a strand of DNA to complementary RNA molecule using the base pairing rules with exception that RNA uses ribonucleotides instead of nucleotides and the base thymine (T) is replaced with uracil (U). The resulting one stranded molecule is called a Messenger RNA (mRNA). In the final termination stage, synthesis is stopped and the transcribed mRNA is released together with the RNA polymerase. [14]

The second step of gene expression is translation where the information in the transcribed mRNA is translated into functional proteins by ribosomes. Similarly to transcription, the process of translation also has three main steps (initiation, elongation and termination). In the initiation step, a ribosome binds to a particular starting region of the mRNA, called the ribosome binding site (RBS), starting the synthesis of polypeptides. The synthesis can be started as soon as the RBS is transcribed, and thus, in bacteria, transcription and translation are coupled. Then, in the elongation step, the nucleotides of mRNA are read in sequences of three, called codons, each specifying a particular amino acid that is added to the growing polypeptide chain. This process is repeated until the ribosome finds a specific stop-codon which informs that the translation needs to be terminated. The new polypeptide is then released and folded to a functional protein where it is able to carry its functions. The ribosome is also released and it is ready to start the translation process again. [14]

2.3 Fluorescent labeling

Intracellular objects can be tracked within living cells by fluorescently labeling the objects to be detected. Fluorescent labeling is based on fluorescent tagging or staining. More specifically, it is based on an attachment of a fluorescent molecule, *fluorophore*, to a target molecule. For example, the attachment can be done by genetically encoding a molecule of interest and a fluorophore as a gene fusion in DNA and enabling cells to produce these molecules by themselves. Another example of fluorescent tagging could be chemical labeling which, on the other hand, relies on interaction between a fluorophore and a specific target sequence of the target molecule. [19]

One of the most commonly used fluorescent molecules, the green fluorescent protein (GFP), obtained from a bioluminescent jellyfish *Aequorea Victoria*, can be used for such purposes [20]. It is a protein that exhibits green fluorescence when irradiated by light in the ultraviolet to blue range. Apart from that, it is also possible to use different derivatives of GFP (such as yellow fluorescent protein, YFP) having different emission and excitation wavelengths. This makes it possible to study different structures in a single cell, at the same time, by using multicolor imaging. [21]

In live cell imaging, there is a tradeoff between the quality of images and the healthiness of the cells. For example, a strong excitation light needed to generate a sufficient emission signal might cause *photodamage* to the cells. Moreover, light itself can cause photodamage, e.g. ultraviolet light is known to cause mutations in DNA [22]. Another problem is so called *photobleaching* in which the fluorophores become non-functional upon the illumination. This also generates chemically reactive free radicals which might damage the cell. [21] Photobleaching is also problematic in time-lapse fluorescence microscopy since the fluorescent intensity of the cells decreases over time. This might affect

the image analysis making it more difficult or even impossible since the cells become less visible and the signal-to-noise ratio decreases.

To minimize the amount of damage taken by the cells and to achieve reliable results when imaging living cells, it is vital to consider both the environment of the cells as well as the microscopy aspect. For example, the cellular environment should be kept as constant as possible by controlling i.e. temperature, humidity and CO_2 levels. Furthermore, damage taken by the cells should be limited as much as possible by minimizing light intensity and the exposure while doing the microscopy measurements. The sensitivity of used camera is also crucial; with highly sensitive detector the illumination intensity can be lowered. [21]

2.4 Fluorescence microscopy

By using a fluorescent labeling technique together with fluorescence microscopy, intracellular organisms and their dynamics can be studied in live cells. It is based on a phenomenon in which the irradiation of a fluorescently labeled specimen with a specific *excitation* wavelength of light causes the specimen to emit light at the longer *emission* wavelength which can then be detected in the microscope. [21]

In Figure 3a, we have illustrated the main components of the fluorescence microscope. Typically, a system consists of a light source, excitation filters, a dichroic mirror, an objective lens, emission filters and a detector. The sample to be observed or *specimen* is illuminated by a light source having specific wavelengths capable to excite the sample. The excitation filter is designed so that only the specific wavelengths of light are allowed to go through while all other wavelengths are filtered. Then, the dichroic mirror reflects the illumination beam to the specimen through the objective lens. In the specimen, fluorophores are excited and they emit light at the specific emission wavelengths. The dichroic mirror, which only reflects the light of certain wavelengths, passes the emitted signal which is then detected by a detector. Before the detector, signal is filtered again so that only the emission wavelengths are captured.

2.4.1 Total internal reflection fluorescence microscopy

In total internal reflection fluorescence (TIRF) microscopy, the idea is to illuminate and excite fluorophores immediately adjacent to the surface glass using evanescent waves [21], [23]. The evanescent waves are generated when light is totally internally reflected at the glass-water surface as shown in Figure 3. The energy of the evanescent waves decays exponentially with distance so only fluorophores approximately at depth of ~100nm or closer to the surface glass are excited (the range of the evanescent waves are shown in yellow) [23]. This is a major advantage of TIRF microscopy; the fluorophores

from other than focus area are not excited reducing the amount of background fluorescence.

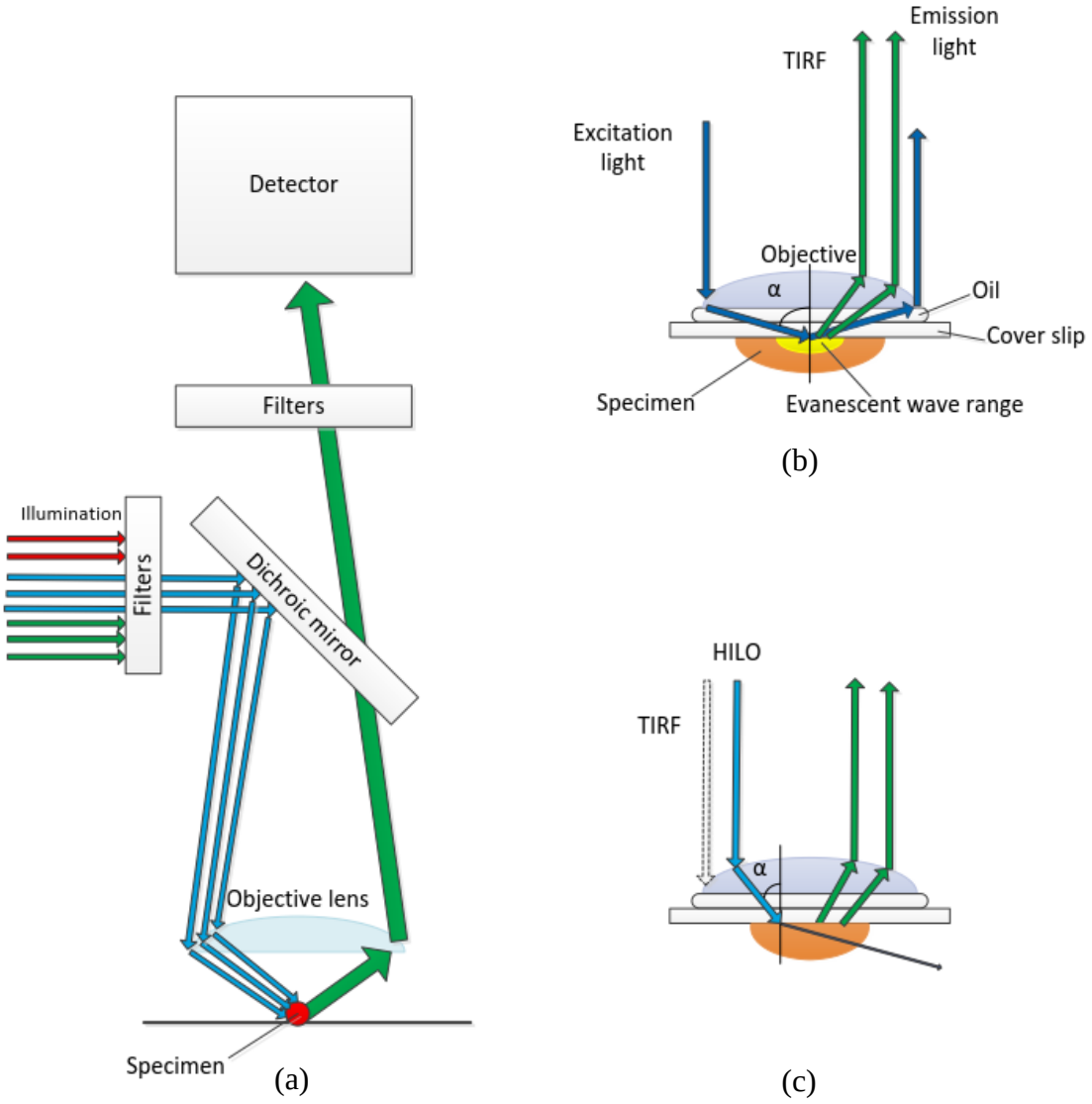


Figure 3: *(a) Main components of a fluorescence microscope. The excitation light (blue) passes through the excitation filters and is reflected to the specimen by the dichroic mirror. Emitted light passes the dichroic and is filtered before the detector. (a) TIRF microscopy. The excitation light is totally internally reflected at glass-water surface when the angle α is big enough. This creates evanescent waves at the boundary (yellow) exciting the fluorophores of the specimen. (c) HILO microscopy. In comparison to TIRF microscopy, different angle α is used. As a result, the illumination beam is highly inclined and laminated as a thin optical sheet at the specimen side exciting the fluorophores in its path.*

2.4.2 Highly inclined and laminated optical sheet microscopy

The main limitation of TIRF microscopy is that it illuminates only the surface of the specimen [23]. Since the average diameter of *E. coli* is approximately 0.5-1.0µm [24], it

is not enough to illuminate the subcellular components properly. Highly inclined and laminated optical sheet (HILO) microscopy overcomes this limitation by highly inclining the illumination beam causing the beam to be laminated as a thin optical sheet at the specimen side [25]. By changing the inclination of the illumination beam and minimizing the illumination area, the background fluorescence can be minimized increasing the signal-to-background ratio. As a result, the intensity of the illumination beam can be decreased while still resulting in high quality images. [25] This is important when imaging the living cells since we want to minimize the amount of damage taken by the cells.

In this thesis, we have used HILO microscopy to achieve a good signal-to-noise ratio (SNR) allowing us to decrease the intensity of light and minimizing the damage caused to cells. It is also faster than the traditional confocal microscopy since HILO is able to capture the whole field at once while the traditional confocal microscope requires scanning of the whole area since each pixel is illuminated separately.

3. IMAGE ANALYSIS BACKGROUND

In this chapter, the fundamental concepts related to detection of intracellular particles from microscopy images are presented. In the first section, the basic framework behind the detection algorithms is presented together with basic image analysis operations needed in these steps. Within this framework, some of the commonly used methods [7], [8] for the spot detection are presented. These methods are then compared with the proposed method later in this thesis. The second section describes the Laplacian of Gaussian blob detection method which the proposed algorithm is also based on. Finally, the third section presents the performance evaluation metrics used to compare the algorithms.

3.1 General framework for detecting subcellular objects

Detection framework usually consists of three main stages (Figure 4), namely, noise reduction, signal enhancement and signal thresholding [7]. All detectors usually include these steps in one form or another with signal enhancement being the most distinguishable part. In the first step, noise is reduced via some denoising algorithm producing a denoised image $J(x,y)$. Techniques can vary from basic filtering techniques to more sophisticated algorithms such as patch based denoising, e.g. [26]. In the second step, the denoised image is further enhanced so that the objects of interest are highlighted and background and other objects are suppressed. The output is often called grayscale classification map $C(x,y)$ since it might not represent the original data in the same way anymore but instead it represents, for example, the likelihood of the subcellular objects. In the final step, a threshold is applied to the classification map in order to separate the objects of interests from the background and other structures. This step produces a binary image $B(x,y)$ where the objects above the threshold have been marked with one (spots) and everything below the threshold is marked with zero (background). [7] From the binary image, several measurements can then be computed, for example, we can determine the area and shape of the objects and we can also extract intensities from the fluorescent image using the mask.

Before any of these steps, images have to be captured with microscope. Although image acquisition is not included in the pipeline, this step is important since the chosen microscopy technique and chosen parameters affect the image analysis step. For example, the limitations when imaging living cells must be taken into account. Badly done imaging can damage cells producing irrelevant data for the study or it can degrade the quality of images making the following image analysis hard or even impossible. As such, it can

be though as part of the framework. In order to select the best parameters, close inter-communication between the image analyzer and the biologist taking the images is usually recommended.

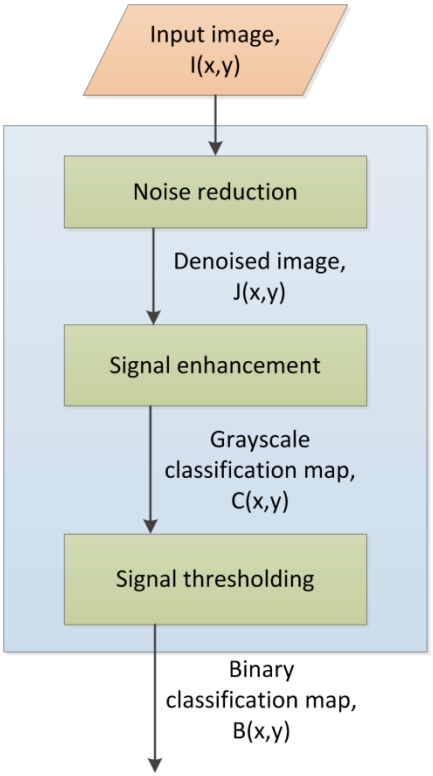


Figure 4: *Framework for detecting subcellular objects.*

3.1.1 Noise reduction

In image acquisition, the photons are detected by the imaging system and converted to intensity values based on the number of detected photons. Due to the quantum nature of light and uncertainty of measurements of such stochastic events, there is variation, or *noise*, in the measured intensity values. This noise is signal-dependent photon shot noise, also called Poisson noise, and can be modelled as

$$g(x,y) = \eta\big(f(x,y)\big), \tag{1}$$

where $g(x,y)$ is the corrupted image, $f(x,y)$ is the noiseless image to be measured and $\eta(x,y)$ is the noise model. The signal-dependent noise follows a Poisson distribution and can be thus modeled as

$$\eta\big(f(x,y)\big) \sim Poisson\big(f(x,y)\big), \tag{2}$$

where each pixel of the image at the coordinates (x,y) is Poisson-distributed random variable with mean $f(x,y)$, i.e. the intensity of the pixel at the point (x,y). [27]

In fluorescence microscopy, the measured signal contains also additive noise components which might come from a variety of sources. For example, by reading the signal from the CCD sensor, additive Gaussian distributed noise signal is included. Moreover, when we present the measured intensities with finite number of discrete values, another additive noise component, quantization noise, is introduced. [27] Also, the natural emission of light by the biological specimen, *autofluorescence*, is one important additive noise source [28]. The additive noise component η_a can be modelled as

$$I(x,y) = g(x,y) + \eta_a(x,y), \tag{3}$$

where *I(x,y)* is the acquired noisy image and *g(x,y)* is the photon-shot noise corrupted image.

In the noise reduction stage, we try to minimize the noise signal of the system. This can be done, for example, by correlating the image with smoothing filter where the value of each pixel is replaced by a weighted sum of its neighboring pixels. Note that, in the case of the symmetric filter, this equals to convolution. More specifically, let us assume a smoothing kernel *W* of size *M x N*, where *M* and *N* are odd nonnegative integers and the center of kernel is indexed as $w_{0,0}$. By making this assumption we do not actually lose anything since every filter can be padded with zeros to have an odd size. Then the filtering operation at position *(x,y)* can be defined as

$$J(x,y) = \sum_{j=-\lfloor M/2 \rfloor}^{\lfloor M/2 \rfloor} \sum_{i=-\lfloor N/2 \rfloor}^{\lfloor N/2 \rfloor} W(i,j) I(x+i, y+j), \tag{4}$$

where *I* is the noisy image, *J(x,y)* is the resulting denoised image and $\lfloor x \rfloor$ is the flooring operation returning the largest integer smaller or equal to *x*. The filter *W* is also called *linear filter* if it preserves the linearity, i.e. if we have two images, A and B, then filtering them with kernel *W* preserves the relationship $F_w(A + \lambda B) = F_w(A) + \lambda F_w(B)$, where λ is a scalar. [29] Filtering operation with kernel of size *5 x 5* is further illustrated in Figure 5.

The blue area illustrates the pixels underneath the kernel and the coefficients of the kernel are marked relatively to the point which we are filtering, i.e. the center of the mask (marked in red). In the special case, if $\omega_{i,j} = \frac{1}{MN}$ for all *i* and *j*, we have a uniform kernel, also called mean or average filter. If the coefficients are defined in such a way that they follow a 2-dimensional Gaussian function, we obtain a Gaussian filter. These filters are also called low-pass filters since they blur the image in the process by attenuating high frequencies while passing the low frequencies. [29]

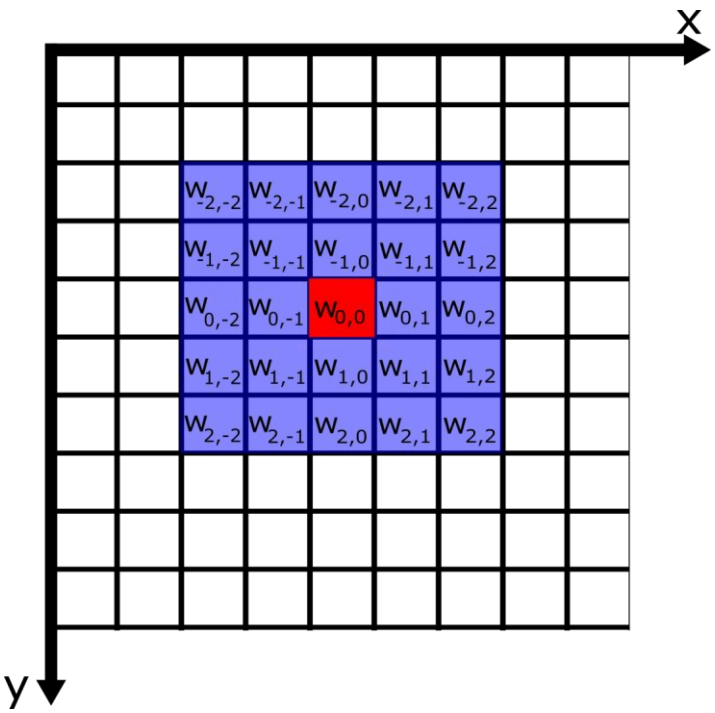
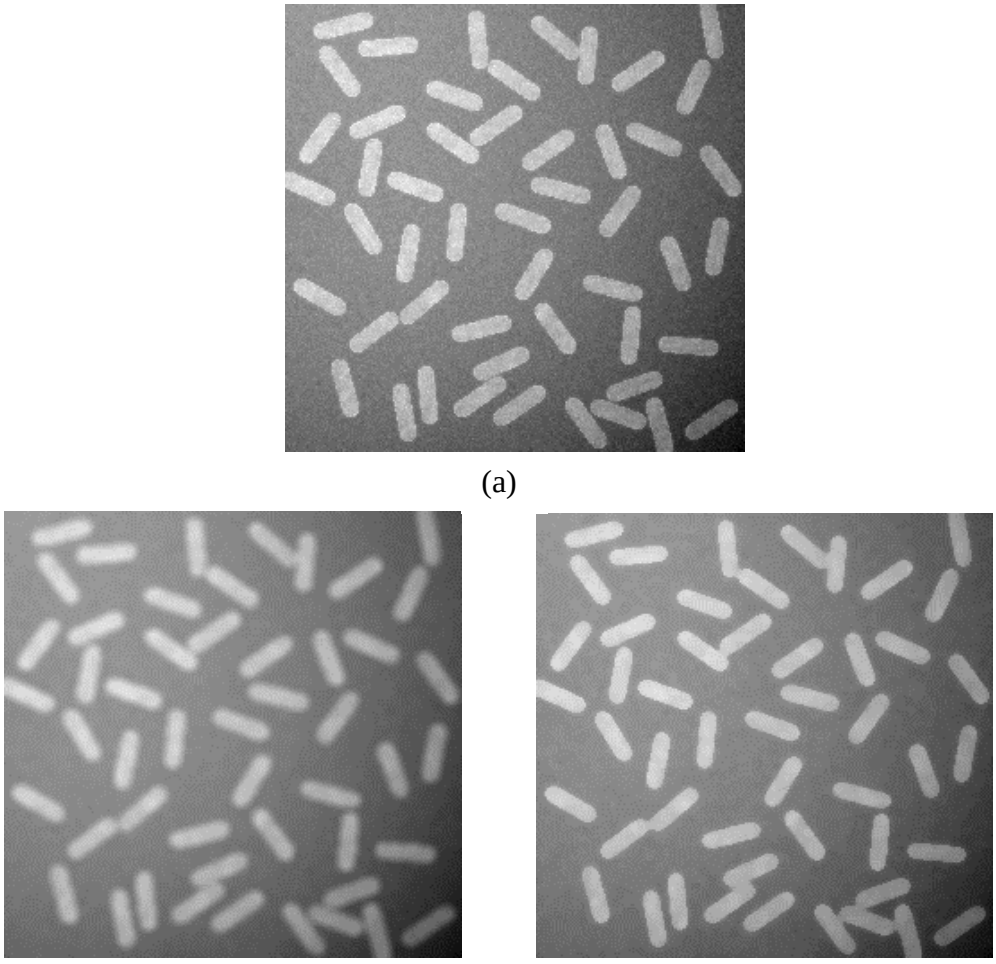


Figure 5: *5x5 filter kernel with coefficients. The kernel is shown in blue and the pixel that is being operated is shown in red.*

The blurring effect can be also a disadvantage. If some fine details, for example edges, need to be preserved, *nonlinear filtering* techniques can be used instead. One such method is a median filter where a target pixel is replaced with the median of neighboring pixels. It preserves the details and edges better than linear filters while still removing noise. [29] Average filtering and median filtering are illustrated in Figure 6. The blurring effect caused by the average filter can be clearly seen as well as the preservation of edges in case of the median filter where the edges are sharper.

However, median filters also have disadvantages. For example, they cannot distinguish between fine detail and noise since anything relatively small compared to the size of the filter will have minimal effect on the median value and will be thus removed. Thus, more complex techniques have been developed, e.g. introducing more accurate models for the noise in the image. For example, a filtering technique based on block matching and 3D filtering (BM3D) has been reported to outperform the previous state-of-the-art methods [26]. However, the denoising technique depends always on the application and what is the goal of the filtering and, thus, should be chosen accordingly.



(a) (b) (c)

Figure 6: *Comparison of noise reduction methods. (a) Original image, (b) original image filtered with a 5x5 average filter. Notice the blurring effect, especially on the edges. (c) Original image filtered with a 5x5 median filter which preserves the edges better.*

3.1.2 Signal enhancement

After the noise reduction step, the signal is enhanced. The goal is to highlight the objects of interest while suppressing the background signal. [7] Usually the signal enhancement part is the most characteristic feature of the spot detection algorithm. Here, we present examples of such techniques that are also used in this thesis in comparison to the proposed method. These methods were chosen because they have been shown to work well with fluorescent images [8].

The method based on kernel density estimation (KDE) [8], filters the image with a kernel function *K* as follows

$$C(i,j) = \frac{1}{card(N)} \sum_{m,n \in N} K\left(\frac{I(i,j) - I(i+m, j+n)}{\alpha}\right), \tag{5}$$

where *N* is the set of neighboring pixels inside the kernel, card is the cardinality (number of elements) of the set, α is the smoothing parameter, also known as bandwidth, and *I(i,j)* is the intensity value of the image at coordinates *(i,j)*. As seen from Equation (5), the output depends only on the image intensities. For example, in case of the Gaussian kernel, uniform-like areas (difference *I(i,j)-I(i+m,j+n)* is close to zero) result in high value in the grayscale map *C(i,j)*. On the other hand, spots that usually have an abrupt change on the boundary and consist of distinct intensity values above the background, result in low value in the grayscale classification map and can be then thresholded accordingly. Typical kernel choice for spot detection is e.g. circular Gaussian kernel as in [4], [30].

Another method, h-dome detector (HD) [7] uses a morphological h-maxima transform [31] as a signal enhancement technique. Figure 7 shows example of h-maxima transformation in case of 1D-signal. All the local maxima that are below the user-defined threshold *h* will be suppressed and the others structures are cut by height *h* from the top of the local maxima. The image on the left shows the blue signal with several different peaks above the background intensity which was set to 50. The image on the right shows the result of the h-transformed signal when the parameter *h* was set to 20. All the spots that are above background more than that are retained, while the others are suppressed. Note that, at this stage, objects are not detected and it is still up to the observer to decide what parts of the enhanced signal represent spots.

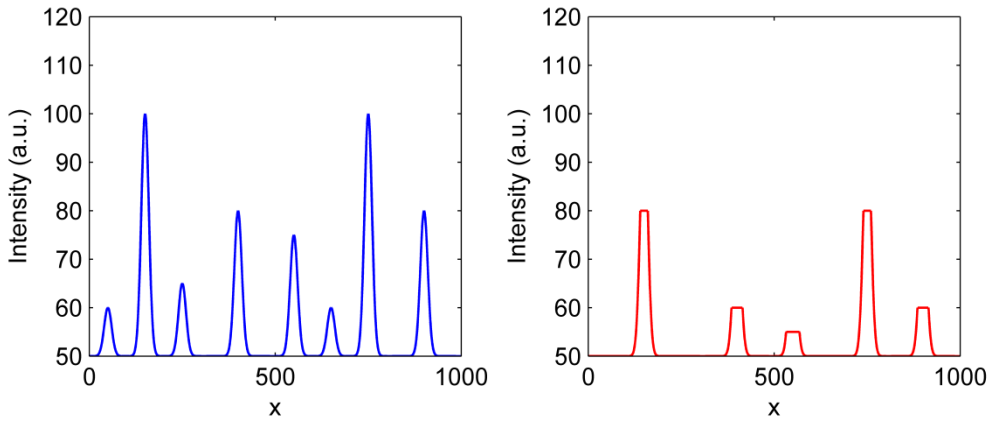


Figure 7: *Example of signal enhancement via h-maxima transform. The original signal has peaks with intensities 10,15,25,30 and 50 above the background. H-transformation suppresses all local maxima that are below the threshold h (in this case, h = 20).*

Grayscale morphological operations [29] have also been introduced for spot detection. One technique is based on a grayscale morphological top-hat-filtering (THE) [29], [32] which suppress the background and enhances spot like structures at the same time. The algorithm performs a grayscale opening [29] with a disk-shaped structuring element and subtracts the output from the original image resulting in an enhanced image where objects corresponding roughly to a size of structuring element are enhanced while the oth-

er structures are suppressed. The resulting image is then thresholded in the next stage of the spot detection framework.

Local enhancement filters (LEF) [8] rely on matched filters for the signal enhancement step. It utilizes the concept that the known signal (filter, or template) can be detected from an unknown signal when the two signals are correlated. The filter is defined in two parts, as illustrated in Figure 8.

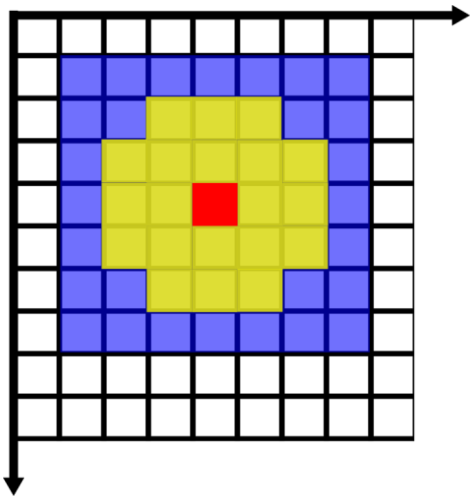


Figure 8: *Local enhancement filter of size 7x7. Kernel is divided into two parts: circular support area A (shown in yellow) and outside part B (shown in blue). The pixel to be filtered is shown in red.*

First, the circular inner part (defined as set A, shown in red) of the filter enhances the local intensity peaks while the outside part of the kernel (defined as set B, shown in blue) is used to suppress the background by division. The filtering operation can be thus defined as

$$C(i,j) = \frac{\sum_{i,j \in A} I(i,j)}{\sum_{k,l \in B} I(k,l)}, \tag{6}$$

where the (i,j) are the coordinates in set A, and (k,l) are the coordinates in set B. Output is a grayscale classification map representing a spot likelihood that is then thresholded in the next step of the framework. [8]

Another detection method for subcellular objects is the local comparison and selection (LC) method, described in detail in [8]. It uses directional filters and compares the outputs of the filters locally to make decision if object is present or not. Circular filter (similar to the support area A in Figure 8) works as a base template for the directional filters: it is separated into four quarters h_{NE}, h_{SE}, h_{SW} and h_{NW} where the three other quarters are set to zero (Figure 9). The original image is then filtered with these four filters and the original pixel is replaced with the maximum pixel value representing the likelihood of

the spot. In the signal thresholding step, local thresholding technique is then applied producing a binary image with the detected spots.

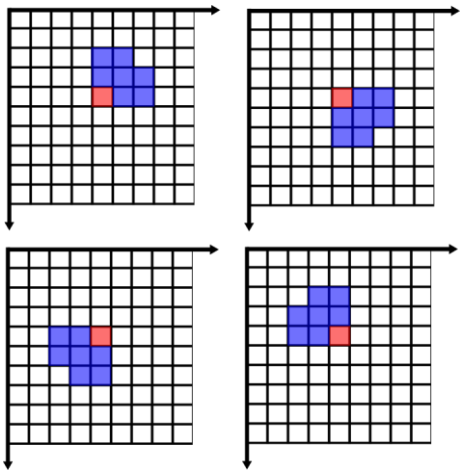


Figure 9: *Directional filters of the Local Comparison and Selection (LC) method. The red shows the pixel to be filtered and the kernel is shown in blue. The name of the filters from left to right and from top to bottom:* h_{NE}, h_{SE}, h_{SW} *and* h_{NW}.

Wavelets have been also proposed for the detection of subcellular objects in [9] where a method based on the multiscale product of wavelet coefficients (MW) was introduced. The method is based on wavelet decomposition where the image $A_0(x,y)$ is convolved with a [1/16, 1/4, 3/8, 1/4, 1/16] kernel row by row and column by column resulting in a smoothed image $A_1(x,y)$. To obtain the smoothed image at scale *i*, kernel is modified by inserting 2^{i-1}-1 zeroes resulting in $A_i(x,y)$. The so called wavelet plane *W* at level *i* is then computed by subtracting the smoothed image at level *i* from the smoothed image of previous level *i-1*, $W_i(x,y)=A_{i-1}(x,y)-A_i(x,y)$, $0 < i \leq J$. As the name of the method suggests, grayscale classification map *C(x,y)* is obtained by multiscale product of the *J* layers:

$$C(x,y) = \prod_{i=1}^{J} W_i(x,y). \tag{7}$$

Intuitively, spots will be present in the final product since they are correlated across the different levels whereas response from the noise will be significantly lower.

Source extractor (SE), originally presented in [33], was recently presented for subcellular spot detection [8]. It is a local method as it estimates and removes the background in small blocks and applies a threshold based on the local blocks. The blockwise background estimates are calculated based on histogram clipping where the histogram is clipped at both ends iteratively until convergence at three standard deviations around its median. If the standard deviation of the clipped histogram is changed less than 20% during the process, the mean is taken as background intensity, otherwise it is estimated

to be BG = 2.5 x Median - 1.5 x Mean. Finally, by interpolating the blockwise background estimates, the estimates for each pixel are obtained. The resulting background corrected image contains the enhanced signal which is then thresholded based on the local background to get the initial spot candidates.

3.1.3 Signal thresholding

In image analysis, *segmentation* is a process where the image is subdivided into regions or objects having a certain characteristics [29]. In the spot detection framework, the goal of the segmentation is to separate the objects of interests (spots) from other structures and the background. This is often done with thresholding, which is a procedure where the segmentation is done by comparing the intensity values of the image to a certain threshold value.

More specifically, a threshold value is applied to an enhanced grayscale image *C(x,y)* producing a binary image *B(x,y)* of the same size, where the objects of interest are marked. In *global thresholding*, a single gray-level value *T* is selected and used as a threshold for the whole image. A thresholded binary image *B* is then defined as

$$B(x,y) = \begin{cases} 0, if\ I(x,y) \leq T \\ 1, if\ I(x,y) > T \end{cases} \quad (8)$$

where the *I(x,y)* is the pixel value of the input image at coordinates *(x,y)*. Thus, all the pixels above the selected threshold *T* (corresponding to objects of interests) are labeled with 1 whereas other pixels (corresponding to background) are labeled as 0. [29]

One of the most commonly used automatic global thresholding procedures is Otsu´s method [34]. It assumes that the histogram of the image is bimodal, i.e. pixels can be divided into two separate classes (foreground and background) based on their intensities. Then, the idea is to find a threshold that best separates these classes, i.e. a threshold value that maximize the inter-class variance $\sigma_b^2(t) = \omega_1(t)\omega_2(t)[\mu_1(t) - \mu_2(t)]^2$, where $\omega_1(t)$ and $\omega_2(t)$ are the probabilities of the two classes and $\mu_1(t)$ and $\mu_2(t)$ are the class means. If the intensity values of image with N pixels are represented with values [1,2,…,L] with corresponding number of pixels [n_1,n_2,…,n_L] both the class probabilities and the means can then be calculated from the normalized gray-level histogram *p* for a given threshold *t* with

$$\omega_1(t) = \sum_{i=1}^{t} p(i) \qquad \mu_1(t) = \sum_{i=1}^{t} \frac{ip(i)}{\omega_1}$$
$$\omega_2(t) = \sum_{i=t+1}^{L} p(i) \qquad \mu_2(t) = \sum_{i=t+1}^{L} \frac{ip(i)}{\omega_2} \quad . \tag{9}$$

where $p(i)=n_i/N$. Finding of the threshold is exemplified in Figure 10 where the selected threshold is shown with green and the separated classes are marked with different colors.

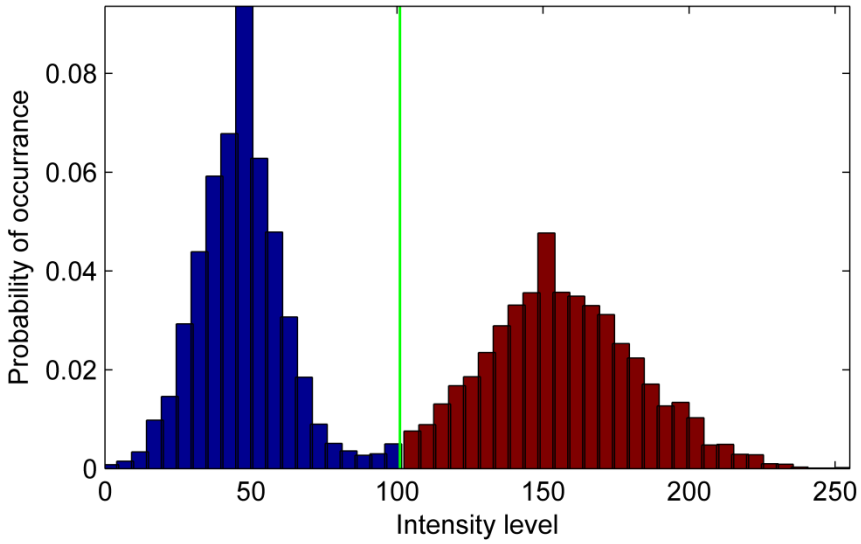


Figure 10: *Otsu´s thresholding. The threshold (green line) separates the intensity histogram into two classes so that the inter-class variance is maximized.*

If the threshold depends on both image intensities and local image characteristics, it is called *local thresholding* [29]. It takes the local image variations into account by choosing a threshold based on the local neighborhood - for instance, the average intensity of *m x n* neighborhood can be used as a threshold. An obvious advantage is that the local thresholding methods are able to cope with changing background much better than the global threshold. Illustration of global and local image thresholding in case of nonuniform background is shown in Figure 11.

(a) (b) (c)

Figure 11: *(a) Simulated image (256x256) of cells with non-uniform noisy background. (b) The segmentation result with a global threshold. (b) The segmentation result with a local thresholding procedure.*

A global threshold was computed using Otsu´s method as described above using Equation (9), whereas for the local thresholding, Otsu´s method was used in local 64x64 neighborhoods. It is clear that the local thresholding method is able to deal with the changing background much more effectively. For example, the global threshold fails to segment the cells due to the nonuniform background. However, in local neighborhoods, the assumption of bimodal distribution of pixel intensities holds true. The block size has to be chosen so that both classes (background and foreground objects) are present in the window, but are small enough so that the background does not vary too much.

In fluorescence microscopy, the assumption of bimodal intensity distribution does not necessarily hold true. Instead, one can make use of the fact that the pixel intensity distribution follows e.g. a Gaussian distribution. If so, it might be beneficial to choose a threshold based on the assumed statistical model in order to get better estimate of the object. One such method is MDE [35], which uses the local cell intensities to get a good estimate for the threshold. It assumes that background pixels follow a Gaussian distribution with the same median $q_{0.5}$ and upper quartile $q_{0.75}$ as the pixels inside the cell. The threshold *T* is then defined as

$$T = q_{0.5} + k(q_{0.75} - q_{0.5}), \tag{10}$$

where the threshold multiplier *k* can be chosen so that the probability of mislabeling a pixel from the assumed distribution is less than some user defined probability *p*.

3.2 Spot detection by Laplacian of Gaussian

One method to detect the subcellular particles from fluorescence microscopy images is based on the Laplacian of Gaussian (LoG) filter. The filter has strong responses for *blobs*, that is, bright, connected regions having intensity above the background making

it suitable for detecting fluorescent particles. Since the response of the filter with a certain size depends on the scale of the blob we can construct a so called *scale-space representation* where image structures are represented at different scales. This is simply obtained by filtering the image with kernels of different sizes. Together with automatic scale-selection procedure, where the appropriate scale for each blob is detected individually, we can automatically detect blobs with varying sizes. [11], [12]

We first present the LoG filter. A Gaussian smoothing kernel is defined as

$$G(x, y, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x^2+y^2)}{2\sigma^2}}, \tag{11}$$

where the *x* and *y* are the coordinates and σ is the standard deviation of the filter. The Laplacian of Gaussian (LoG) operator is then defined as

$$L(x, y, \sigma) = \nabla^2(G_\sigma(x, y) * I(x, y)), \tag{12}$$

where an input image *I(x,y)* is first convolved with the Gaussian smoothing kernel *G* and then the Laplacian operator $\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$ is applied. [29] The Laplacian and Gaussian operators can be further combined to a single Laplacian of Gaussian filter (Figure 12) defined as

$$LoG(x, y, \sigma) = -\frac{1}{\pi\sigma^4}\left[1 - \frac{x^2 + y^2}{2\sigma^2}\right] e^{-\frac{(x^2+y^2)}{2\sigma^2}}, \tag{13}$$

resulting in a simpler form

$$L(x, y, \sigma) = LoG(x, y, \sigma) * I(x, y), \tag{14}$$

where we have only convolution between the LoG filter and the input image.

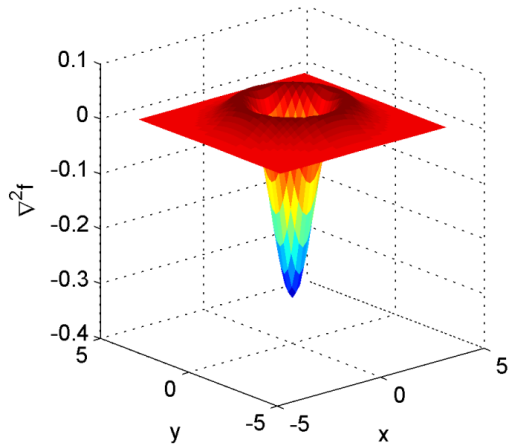


Figure 12: *Laplacian of Gaussian kernel with standard deviation 2 and kernel size 9x9.*

As seen from Figure 12, LoG filter is a circularly symmetric operator making it suitable for blob detection. As defined in Equation (11), it is controlled by the standard deviation σ and the size of the filter (x, y) coordinates). Usually, odd kernel size is used to prevent artifacts caused by the non-symmetric filtering. Moreover, the size of the kernel can be related to the standard deviation, e.g. we can set the kernel size to be $s = 4 * \sigma + 1$ which is motivated by the fact that over 99% of the energy of the Gaussian bell is concentrated within four standard deviations of its mean.

One specific property of the Laplacian of Gaussian filter is that the magnitude of the response is maximized if the standard deviation of the Gaussian is matched with the scale of the blob. More specifically, for a binary circle of radius *r*, the Laplacian achieves a maximum at the center of the blob at scale [11], [12]

$$\sigma = \frac{r}{\sqrt{2}}. \tag{15}$$

However, the magnitude of the Laplacian response depends on a scale σ as it can be seen from Equation (13), i.e. the response decreases with increasing scale. To make the LoG operator invariant to scales, LoG filter has to multiplied with σ^2 [12].

To further illustrate the behavior of the filter, we have plotted a one dimensional signal containing a blob (bright structure above the background) and the response of the LoG filter to that signal (Figure 13). As seen from the image, the filter with σ = 2 does not match very well with the radius of the blob and the response of the filter is small. On the other hand, the response of the filter with σ = 3.5 fits to the scale of the blob well, resulting in a high response. Since with real images the scales of the blobs are unknown, a single scale level is not enough for stable extraction of blob-like structures. For example, closely located blobs might be detected as one unit if the scale is too large or they might not be detected at all. Two closely located blobs are separated only if the fine scale of the blobs is found. Thus, it is important to construct a so called scale-space representation, where the input image is convolved with the LoG filters of different scales and the output contains the responses at each scale.

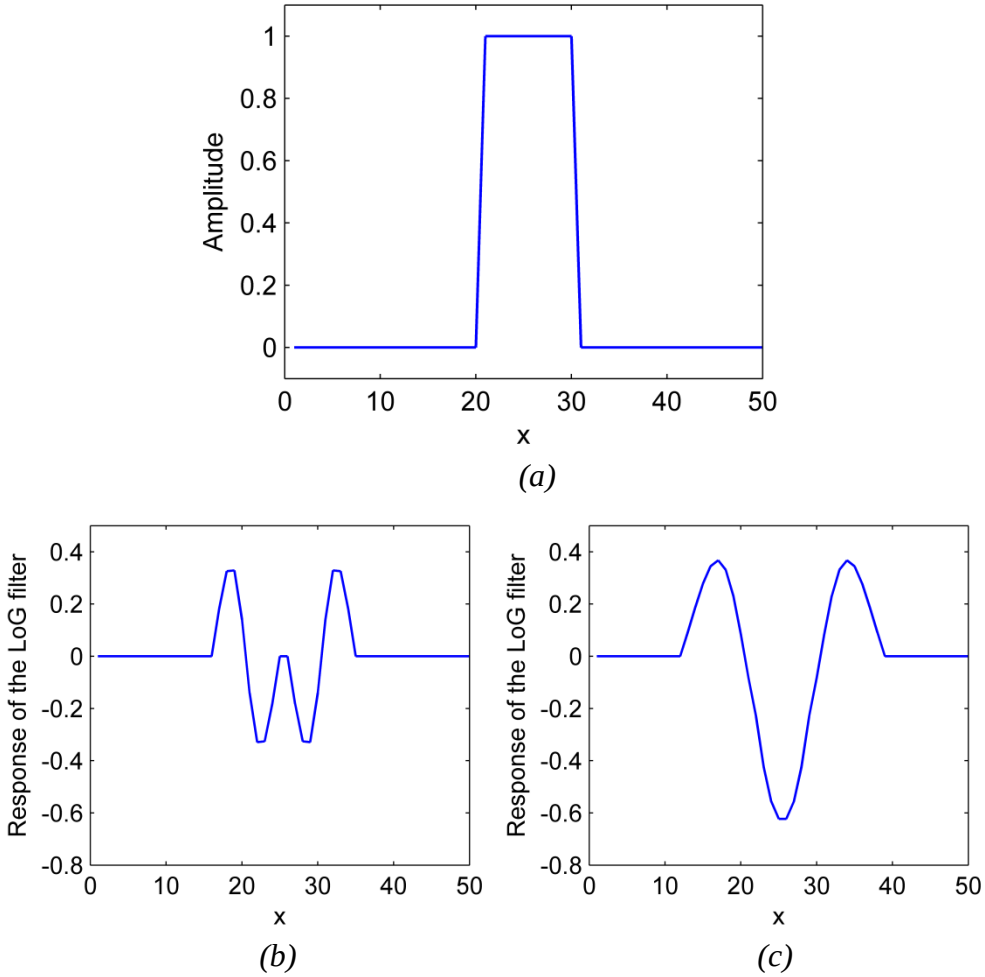


Figure 13: *(a) 1-D signal of 50 pixels containing a blob with radius 5. (b) Response of the normalized LoG-filter to the blob with σ = 2 and filter size 9. (c) Response of the normalized LoG-filter to the step edge with σ = 3.5 and filter size 15.*

After obtaining the scale space, the idea is to search the optimal scale for each pixel which corresponds to the maximum response of the filter. This is done by searching a local extremum for each pixel in a 3D-space (3x3x3 neighborhood). If the value of the pixel at certain scale σ is greater than the neighboring pixels, it is chosen as candidate blob. [11], [12] Scale-space representation and the search of local extrema are depicted in Figure 14. Scale-space representation was obtained by filtering the image with different LoG filters, namely with σ=1, σ=3 and σ=5, where the size of the kernel was chosen as explained above. Then, the maximum for each pixel is search in 3x3x3 neighborhood. In this case, the search was done at scale *i*.

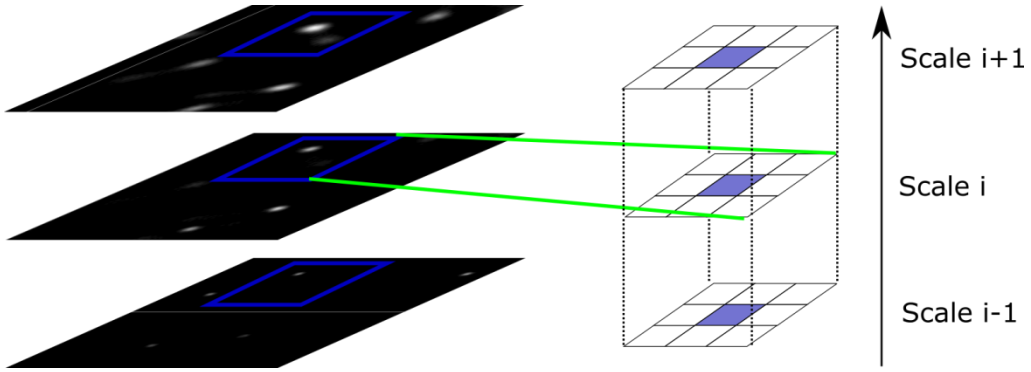


Figure 14: *Illustration of scale-space representation and the local extrema searching. Local extreme is searched in 3x3x3 space across the different scales. If the value of the pixel is above the neighborhood, it is chosen as a candidate.*

As a result of LoG filtering and the local extrema searching, we obtain a gray-scale classification map with set of candidates and their responses. Also, corresponding scales for each candidate are saved for object reconstruction. Then, as in the final step of the spot detection framework, a threshold is applied in order to extract the location of detected objects. For this, a global threshold, as described in Equation (8), can be applied to filter out the weak responses. Finally, the objects of right size can be constructed by using the scale information of the remaining candidates and the relationship between the scale σ and the blob radius r as described in Equation (15).

3.3 Performance evaluation metrics for classification

In order to do an objective analysis of the detectors, we use standard classification performance evaluation metrics, namely, precision, recall and f-score. We also use another common measure used for comparisons of spot detection algorithms [7], the true positive rate when the false positive rate is 0.01, which shows the performance when only low number of false positives is allowed. Moreover, we explain briefly the principles behind the free-response receiver operating characteristic curve which is used to show the effect of varying threshold parameters.

3.3.1 Precision, recall and F-score

We start by defining the 'ground truth image' as an image containing the true labels of objects. The classification map obtained after the signal thresholding is then compared to the ground truth. Object is a true positive (TP) if it is correctly matched to the ground truth and it is a false positive (FP) if there is no match in the ground truth image. If there is a missing object in the detection result, it is called a false negative (FN). True negative (TN) is defined as an accurate rejection of an object. In our case, the number of true negatives is undefined since the negative labels are not available. We denote the number of true positives as N_{TP}, the number of false positives as N_{FP} and the number of false positives as N_{FN}. The total number of objects in the ground truth is defined as N.

Since the false positive rate is defined as $FPR=N_{FP}/(N_{FP}+N_{TN})$ and N_{TN} is not known, we use a modified version of FPR [7], given by

$$FPR^* = \frac{N_{FP}}{N} \tag{16}$$

As in [36], we define the true positive rate, also noted as recall *r*, as

$$TPR = r = \frac{N_{TP}}{N_{TP} + N_{FN}} = \frac{N_{TP}}{N}, \tag{17}$$

and the precision *p* as

$$p = \frac{N_{TP}}{N_{TP} + N_{FP}}. \tag{18}$$

Intuitively, recall emphasizes the number of true objects found, in other words, the failure to detect the true objects is penalized. Precision emphasizes the ratio of N_{TP} over all the detections; the measure is thus penalized by falsely detected objects.

Furthermore, the F-score is then defined as a harmonic mean of the precision and recall [36] as

$$f = \frac{2p * r}{p + r}. \tag{19}$$

The main advantage of the F-score is that it gives a single measure of the detection accuracy by combining and equally weighting precision and recall. In this work, we use a best reachable F-score denoted as f*.

3.3.2 Free-response receiver operating characteristic curve

The receiver operating characteristic (ROC) curve is widely used method to evaluate binary classifiers. In this, TPR is plotted as a function of FPR as the parameters of the detection method are varied. However, since TN is not known, the maximum possible number of FPs is not bounded since there can be any number of false detections per image. As a result, the standard ROC curves cannot be used and the modified version of ROC curve, free-response operating characteristic curve (FROC), is used instead where FPR is replaced with FPR* [37], [38]. FROC-curves are especially useful when studying the sensitivity of the algorithm to parameters. The FROC space with five different discrete classifiers is demonstrated in Figure 15.

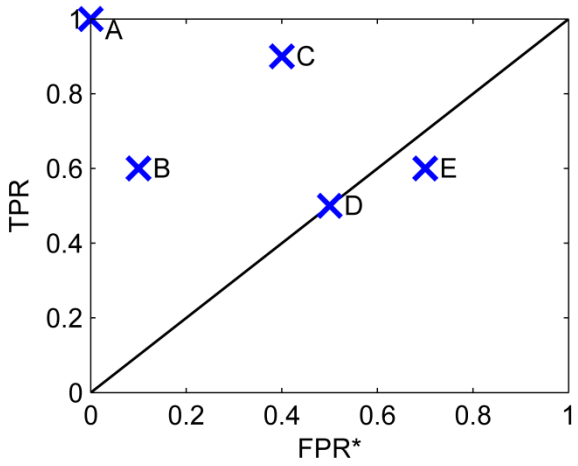


Figure 15: *The performance of five different discrete classifiers in FROC-space.*

Note that FPR* is not necessarily bounded to 1, even though the x-axis was set between (0,1). The point (0,0) represents a point where no classification is done, i.e. no true positives are found but neither any false detections are made. On the contrary, the point (1,1) represents a strategy where all the objects are detected but with cost of same number of false positives. Classifier A at point (1,0) is an example of perfect classifier with maximum TPR and minimum FPR*. The diagonal line between (0,0) and (1,1) represents the strategy of randomly guessing a class. All the classifiers above that diagonal line are working better than a random classifier. For example, classifiers B and C perform worse than A, but their results are still better than random selection. More generally, the classifier performs better than the other classifier if the point in the FROC space is located to the northwest of the other.

In addition, as defined in [36], the classifiers with low FPR are thought as "conservative" classifiers. The classification is done only when there is strong evidence that the object is true, resulting in low FPR*. However, this is shown usually in low TPR as well. On the other hand, classifiers on the north-east side are thought as "liberal" classifiers. They make classifications also with weaker evidence, which lead to high TPR but also to higher FPR*. In Figure 15, classifier B represents the conservative classifier whereas classifier C represents a more liberal classifier. Classifier D, on the other hand, represents a random classifier since it lies on the diagonal line at point (0.5, 0.5), i.e. half of the true objects are found but also the same number of false objects is detected. Classifier E is the worst classifier, as FPR* exceeds the TPR, indicating that even randomly selecting a class could yield better results.

Figure 15 shows only five discrete classifiers but usually it is useful to produce a FROC-curve for each classifier. Each parameter set of the algorithm produces one point in FROC-space so by varying the parameters we can construct a FROC-curve which demonstrates the sensitivity of the algorithm to parameter changes.

4. DETECTION OF FLUORESCENT PARTICLES IN *ESCHERICHIA COLI*

We present a novel framework for detecting the fluorescently labeled subcellular objects in *Escherichia coli*. The pipeline consists of two major stages, a global processing block and a local processing block. In this chapter, we present image analysis steps in these blocks which are required for detecting the fluorescently labeled objects inside the cells. In addition, the performance evaluation metrics are presented for validating the methods. The methods were implemented in MATLAB version 8.3.0.532 (R2014a).

4.1 Fluorescent labeling of MS2 and Tsr proteins

One of the current single-molecule detection techniques to study RNA production in living *E. coli* cells is based on MS2-GFP tagging of RNA. The system is based on the coat protein of bacteriophage MS2, which is able to recognize and bind to specific sequences of RNA. This protein has been fused with the GFP, thus allowing detection and tracking of individual mRNA molecules containing the specific sequences. The information used to produce these MS2-GFP molecules is usually encoded into a multi-copy plasmid that is inserted into the cells, allowing production of these molecules when induced. The target RNA molecules are engineered so as to have an array of MS2 binding sites and their transcription is controlled by the promoter of interest. [16] When binding to the RNA, the concentration of these molecules increases which allows seeing the tagged RNAs as bright spots with fluorescence microscope. Together with time-lapse imaging this can be used to, for example, study the dynamics of transcription [1], [4], [16].

Similarly, we can detect the production of single proteins and consequently, study the dynamics of translation. One construct for such task is Tsr-Venus which is based on the fusion of the membrane protein Tsr and the yellow fluorescent protein (YFP) variant, Venus. This membrane protein has been chosen since its slow diffusion rate makes the detection of individual molecules possible, contrary to the fast diffusing molecules where the diffused signal is overwhelmed by the cellular autofluorescence, making the detection impossible. [2] This construct enables the study of single proteins giving insights to gene expression and other fundamental biological processes such as thermosensing [39] and aerotaxis [40] in which Tsr proteins are known to be involved as chemoreceptors of *E. coli*.

4.2 Cell culturing and image acquisition

For Tsr-Venus, we used *E. coli* K-12 strain SX4, harboring the Tsr-Venus gene construct under the control of the lac promoter (P_{lac}) [2]. Liquid cultures were grown in M9 glucose (0.4%) media overnight, supplemented with amino acids at 37° C with shaking (250 RPM). Then, cells were inoculated into a fresh media and grown until a final optical density of OD_{600} of 0.3 was reached. The induction of Tsr-Venus was then performed by adding 200 µM IPTG to the culture and then cells were incubated at 37 °C for 5 min before imaging.

For MS2-GFP, we used *E. coli* strain DH5α-PRO, provided by I. Golding (Baylor College of Medicine, Houston, TX). From overnight cultures, precultures were diluted in fresh LB media supplemented with appropriate antibiotics to an OD_{600} of 0.05. Then, they were kept at 37 °C in a shaker (250 RPM) for approximately 1.5 hours until OD_{600} of 0.3 was reached. Cells were then induced with 100ng/ml of aTc for 45 min until the final OD_{600} of 0.5 was reached. To complete the induction of the target RNA production, 1mM of IPTG was added and cells were incubated for 5 min before imaging.

In both cases, images of the cells were acquired by fluorescence microscopy, using a Nikon Eclipse (Ti-E, Nikon, Japan) inverted microscope with a C2 confocal laser scanning system using a 100x apochromat TIRF (1.49 NA, oil) objective. The GFP fluorescence was measured using a 488nm argon ion laser (Melles-Griot) and a 515/30 nm detection filter. Phase-contrast images were acquired using a CCD camera (DS-Fi2, Nikon).

4.3 Generation of simulated images

In the case of the empirical images, the ground truth is not usually available and the labeling must be done through a subjective and error-prone manual annotation by an expert biologist. Moreover, the labeling process is very time-consuming making it difficult to evaluate large amount of data. Simulated images have great advantages compared to them: the ground truth is known exactly allowing objective evaluation of the algorithms and the labels are generated automatically during the simulation process allowing easy and fast way to evaluate large amount of data [41].

We used a modified version of SIMCEP cell image simulation framework [41], [42] to generate synthetic images for validation of the algorithms. The original program which can be downloaded at *http://www.cs.tut.fi/sgn/csb/simcep/tool.html*, is able to generate synthetic cells consisting of different components, such as cytoplasm, nuclei and subcellular components. The original SIMCEP framework consists of two main stages. In the first step, ideal images are generated based on the cell and population characteristics.

Then, the ideal image is degraded progressively by simulating the imaging system and image acquisition.

The simulated microscopy images consist of simulated single cells that together form the cell population. As in fluorescent labeling, different objects in the cell can be simulated using a multichannel representation of the cell allowing simulation of e.g. nucleus and subcellular components. In SIMCEP, cells and their shapes are simulated using a parametric model where the randomness and size of the shape can be controlled allowing creation of varying shapes. The texture of the simulated cells are then generated by using a procedural texture model [43], where the texture can be controlled with a variety of parameters allowing creation of different textures. At the population level, it is then important to simulate the locations of the cells realistically. SIMCEP organizes the simulated cells randomly throughout the image, since the locations of cells rarely follow uniform distributions in real life as well. The program is also capable of clustering the cells with some probability *p* in order to simulate cell populations realistically. Furthermore, the overlap between the cells can be controlled in order to simulate spatially tightly packed cells. [41]

In the second step, ideal images are degraded progressively to simulate the effects of real microscopy system. For this, several sources degrading the quality are introduced, including nonuniform illumination, autofluorescence, detector noise and optical aberrations such as blurring. [41] Briefly, nonuniform illumination is modelled by adding a second degree polynomial surface in the image where the energy and the center of the illumination source can be controlled. Autofluorescence is simulated by using the texture model [43] and blurring of the image is modeled by a linear space-variant blur model using a Gaussian kernel [44]. Finally, the detector noise is modeled by adding a zero-mean Gaussian noise with variance σ^2. [41] The background parameters used for our test set are shown in Table 1.

To modify the program to our needs, an additional model of rod-like shaped bacteria as described in [45], e.g. *E. coli*, was added to the program. Rod-like bacteria was generated by first inserting a center point of the cells at random locations in the image. Then, morphological dilation [29] with line-shaped structuring element of length *l* and angle α was performed to get the lines. The length of the *E. coli* was set to 25 for our test set and the angles were drawn from the uniform distribution on the interval (0,180). Final shape was then obtained by dilating the binary line with another structuring element, a disk with radius r. The radius was set to 5 for our test set.

Also, additional texture model for spots was added. Instead of using the default model [43] for the spots, we generated Gaussian spots of radius *r* and sigma σ. For the *E. coli* test set, we used $r = 3$. Standard deviation σ was then automatically computed from the radius so that over 99% of the Gaussian was included as described in Equation (15).

Table 1: *Parameters of the synthetic image generator*

Parameters	Values
Number of cells	200
Number of objects per cell	1
Background energy	0,25
Autofluorescence energy	0,25
CCD-noise	0,005

4.4 Proposed spot detection method

4.4.1 Cell segmentation

Although the cell segmentation is out of the scope of this work, it is usually an important step since one usually wants to relate the fluorescent spots to the cells that produced them. The cell segmentation was performed with a custom-made software that integrates MAMLE [46] for cell segmentation and CellAging [47] for alignment of the fluorescence and phase contrast images and for lineage construction.

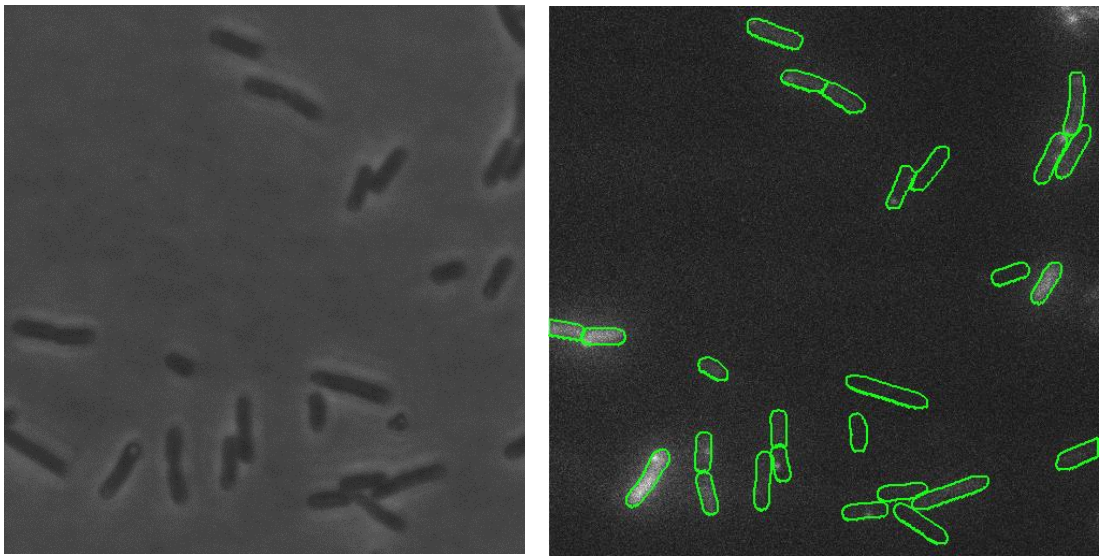
The cell segmentation method, MAMLE (Multi-Resolution Analysis and Maximum Likelihood Estimation) [46], have also the three main stages as described in Section 3.1, i.e., noise reduction, signal enhancement and signal thresholding. For noise reduction, it uses efficient image denoising technique called Block-Matching and 3D filtering (BM3D) [26] which divides the image in 8x8 blocks grouping similar 2D blocks together forming a 3D array. Special collaborative filtering is then applied for this group and finally the filtered blocks are returned to their original positions. In the second stage, the signal is enhanced by applying multi-resolution edge detection to the image. Initial segmentation is then acquired by threshold decomposition which utilizes several thresholds in order to get the initial estimate. Then, morphological features of the cells are acquired from this initial estimate and used as a prior in the next step in which the segmentation result is corrected by using a maximum likelihood strategy for splitting and merging the cells. The output of the algorithm is a classification map where the area occupied by each cell is marked with a different label. [46]

Alignment of the images and the construction of lineages of the cells are then done via CellAging [47]. The alignment is needed since the fluorescence and the phase contrast images are acquired with different cameras with different resolution. The alignment procedure consists of a global and a local alignment stage. First, the global alignment is done using scaling and translation operations. The best combination of these is chosen by maximizing the total fluorescence under the masks of segmented cells. Then, the local alignment is performed since the phase contrast and fluorescence images may not be acquired exactly at the same time and the cells might have moved. Local alignment is

done by computing a local translation separately for each cluster of cells using cross-correlation. [47]

For time-series, lineages are constructed by establishing the relationship between the cells of consecutive frames. For this, the amount of overlap between the cells is used to associate the cells between frames. At this point, cell divisions are taken into account, namely, if one cell was associated with multiple cells in the next frame, a division is expected to have occurred [47]. This step, while performed automatically, is then verified by a human, to determine whether manual corrections are needed.

As a result of cell segmentation and alignment, the aligned labels for the cells are obtained as illustrated in Figure 16. With time-series, also lineage information is extracted, which allows us to track individual cells and their descendants over time.



(a) (b)

Figure 16: *(a) Cropped phase-contrast image of the cells. (b) Aligned fluorescence image with boundaries of the cells.*

4.4.2 Spot detection procedure

The proposed algorithm can be divided into two main stages as depicted in Figure 17. First, candidate spots are searched globally using the Laplacian of Gaussian blob detector described in Section 3.2 and in [11], [12].

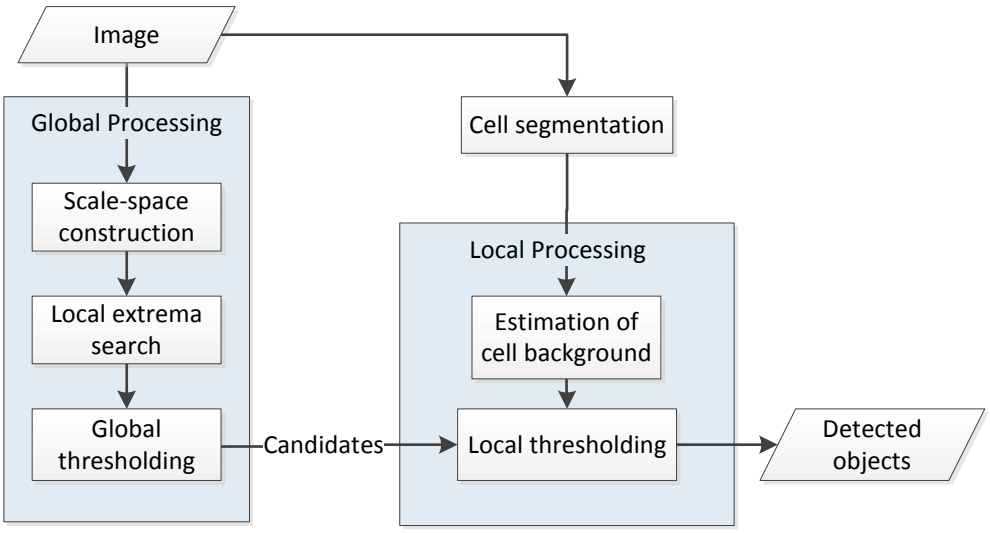
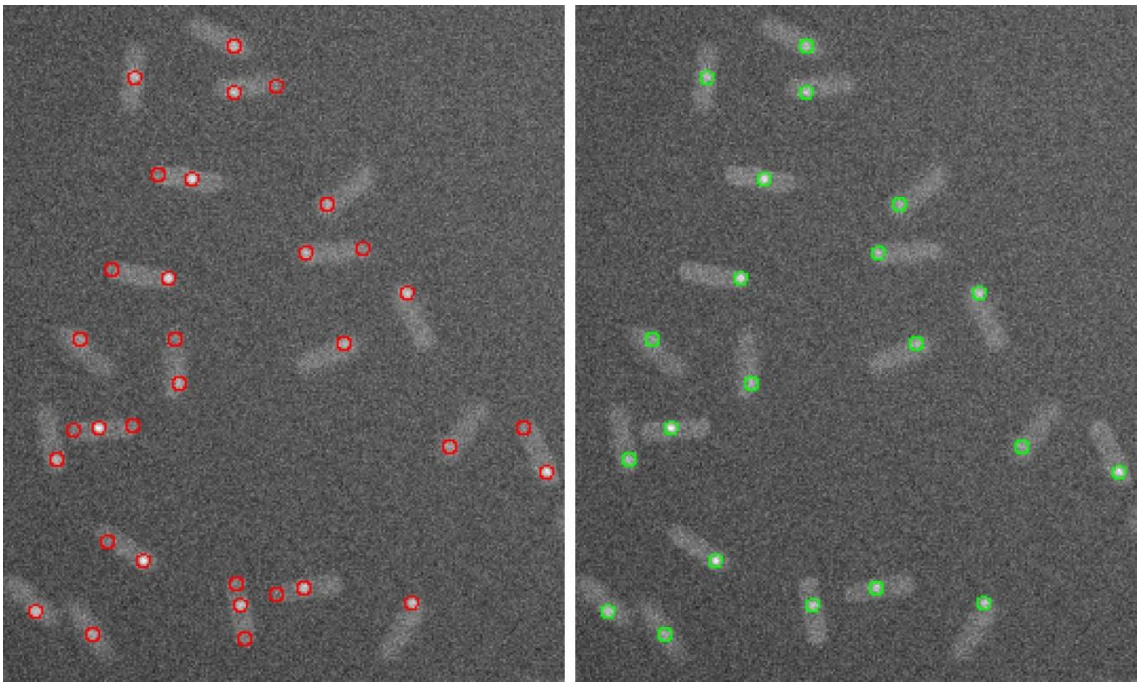


Figure 17: *Overview of the image analysis pipeline.*

Second, true spots are selected among the candidates using a locally adaptive threshold selection where the threshold is chosen based on the intensities of each individual cell. This allows us to utilize the local image characteristics but at the same time, candidates will have blob-like properties detected by the Laplacian of Gaussian detector.

The global procedure starts with the construction of the scale-space. Prior to this operation and as a preprocessing step, the image is normalized to the interval [0,1]. Then, the image is filtered with the Laplacian of Gaussian filter with different standard deviations and sizes to obtain the different scales. We used 20 scales spaced between the minimum and the maximum scale. Minimum scale was set to $\sigma = 2$, corresponding roughly to a spot with radius three. The maximum scale of the spots was inferred from the cell segmentation results: the mean minor axis length of cells divided by the factor 2 was used as the maximum radius of the blob and the appropriate scale was then calculated as in Equation (15). After filtering, only negative responses are taken since we are only interested in bright spots in dark background. The maximum negative response at the optimal scale is achieved at the center of blob as depicted in Figure 13.

After the construction of the scale-space, the local extrema is searched in 3x3x3 neighborhood. This can be done efficiently utilizing mathematical morphology since with grayscale images the dilation operation corresponds to maximum filter [29]. The filtered image in each scale is dilated with 3x3 structuring element and then the maximum value for each pixel is searched between neighboring scales. If the value of a pixel at a certain scale is greater than all its neighbors the pixel is picked out and the location and the scale are stored. Finally, a low global threshold is applied to all detected pixels to get the initial candidates for the local thresholding procedure. Initial candidates after global thresholding are shown in Figure 18 in red color.



(a) (b)

Figure 18: (a) *Simulated fluorescence image with **s**pot candidates (marked as red) after global thresholding. (b) Detected spots (green) after removal of false candidates via adaptive local threshold.*

It is worth to note that, usually, it is not necessary to tune the global threshold but it is enough to set it low enough to filter some of the candidates out. For example, for all our different datasets, this initial threshold was set to 0.002.

After getting the initial candidates, cell segmentation result is used to filter out any fake candidates outside of the cells, i.e. all detections outside the masks of the segmented cells are removed. Moreover, since the location of the detected blob might not be exactly the same than the location of the true center, there can be multiple responses to the same blob. If two candidate blobs with center coordinates and radii (x_1, y_1, r_1) and (x_2, y_2, r_2) are too close to each other as defined in

$$\sqrt{(x_2 - x_1) + (y_2 - y_1)} < \max(r_1, r_2), \tag{20}$$

the smaller blob is removed as a fake candidate since we want to be sure that the whole spot is captured.

Last, a local thresholding is applied to determine if the candidate blob is an actual object. This threshold is adaptively calculated for each cell using the following procedure. The local threshold is determined from the cell intensities by assuming that the pixel intensities of the cell follow approximately a normal distribution. Before fitting the normal distribution, noise is reduced by filtering the cell intensities with Gaussian low-pass filter with standard deviation 0.5. Then, the parameters of the normal distribution,

the mean $\hat{\mu}$ and the variance $\hat{\sigma}^2$, are estimated using the maximum likelihood estimation [48]. Briefly, the estimates can be expressed as:

$$\hat{\mu} = \frac{1}{n}\sum_{j=1}^{n} x_j$$
$$\hat{\sigma}^2 = \frac{1}{n}\sum_{j=1}^{n} (x_j - \hat{\mu})^2 \quad (21)$$

where x_j is the jth pixel inside cell with total number of n pixels.

After obtaining the maximum likelihood estimates, a threshold p is chosen based on the fitted distribution. The threshold is defined as a probability of mislabeling the candidate given the null hypothesis that it is a background pixel following the estimated Gaussian distribution. Figure 19 illustrates the fitting and the searching of the local threshold. Red line shows the fitted normal distribution and the green line shows the chosen threshold when p was set to 0.01. If the candidate blob intensity is above the intensity determined by the chosen threshold, it is accepted as a true spot.

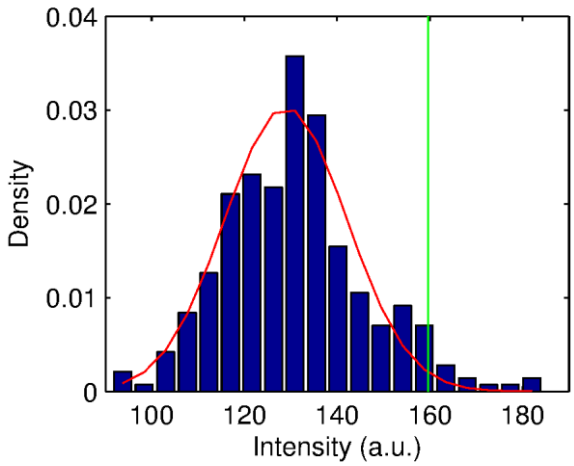


Figure 19: *Probability density of the intensities inside the cell. Red solid line is the fitted Gaussian distribution and the green line is the selected local threshold with p=0.01.*

Finally, the final binary classification output is constructed by setting the pixels at the locations of the accepted candidates to 1 and then utilizing morphological dilation with disk structure element. The radius of the disk is determined by the scale in which the blob was detected.

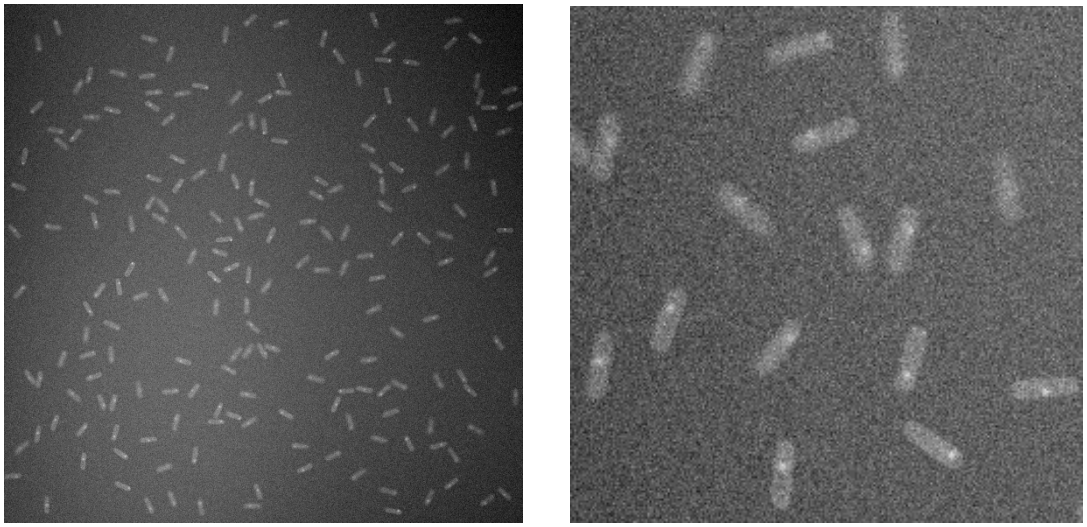
5. RESULTS & DISCUSSION

The performance of the proposed algorithm was evaluated on both synthetic (for whom ground truth is known) and real image data. The results were compared with 9 unsupervised techniques, described briefly in Sections 3.1 and 3.2 and in detail in [7], [8], [12], [35], and 1 supervised spot detection technique [13]. The implementations of unsupervised techniques presented in [8] are available at *http://www.cs.tut.fi/sgn/csb/subcell/*. We also included one recent technique, ATLAS, proposed in [10] and publicly available for use at their website: *http://serpico.rennes.inria.fr/doku.php?id=software:atlas:atlas*.

The best parameters in terms of evaluation metrics (f*, TPR at FPR*=0.01) were searched by a grid-search for each algorithm and for each dataset, similarly to [8]. The found parameters for the different datasets are listed in Appendix 1. Detailed explanation about the parameters can be found in [8]. We note that for objective comparison, the cell segmentation results were also used for these algorithms to filter out all false positives outside the cells similarly to the proposed method.

5.1 Evaluation on Synthetic Image Data

Synthetic images provide an objective approach for validation since the procedure is easily repeatable and does not rely on the expert biologist manually annotating the images. Also, the parameters are easily tunable allowing experiments with, e.g., specific noise levels and background models. To evaluate the detectors objectively, we used two different sets of simulated images. First, simulated *E. coli* images with non-uniform background were generated where the CCD noise was set to 0.005 to simulate the low signal-to-noise ratio peculiar to real-life images (Figure 20). Second, we demonstrated applicability of our algorithm to other types of cells as well. For that, we used an independent dataset provided in [45].



(a) (b)

Figure 20: *Example image of the simulated Escherichia coli dataset (a) and a close-up image (b).*

5.1.1 Results for simulated *Escherichia coli* images

The first data set consisted of 5 images with 200 cells in each. The number of subcellular objects inside a cell was set to 1, so the total number of spots to be detected in the whole set was 1000. Performance of the detectors was evaluated by the best reachable F-score (f*) and TPR at FPR*=0.01. Also, the precision and recall of the detectors in which the f* was achieved are shown (Table 2).

Table 2: *Performance of the detectors.*

Method	p	r	f*	TPR at FPR* = 0.01
Proposed method	0,977	0,976	**0,976**	**0,962**
MED	0,945	0,965	0,955	0,860
HD	0,939	0,877	0,907	0,684
KDE	0,895	0,952	0,922	0,806
LC	0,866	0,861	0,864	0,579
LEF	0,809	0,731	0,768	0,479
LoG	0,865	0,849	0,857	0,702
MW	0,929	0,800	0,860	0,735
SE	0,929	0,887	0,907	0,730
THE	0,863	0,938	0,899	0,773
ATLAS	-	-	-	-

The proposed method shows superior performance in terms of f* and TPR at FPR*=0.01. Next, MED and KDE methods were the most satisfactory, reaching high f* over 0.9 and TPR>0.8 with FPR*=0.01. While almost all algorithms achieve reasonable

f* over 0.85 (except LEF), ATLAS failed to detect the subcellular components. This might be due to the fact that it automatically tries to select one optimal scale for the spots, instead of finding a scale for each pixel individually. Due to the noisy spots and small size of the bacteria, the scale chosen was the level of cells which prohibited the detection of subcellular components.

Although the top three detectors were able to achieve similar F-score, there is lot more variation when low FPR* is required. Only the proposed method, MED and KDE were able to achieve TPR over 0.8. Moreover, the proposed method was the only method to achieve TPR over 0.9 with FPR* = 0.01, with increase of over 11% compared to second best result. This illustrates the main advantage of the method; it is able to achieve high TPR with a small number of false positives. Note that the improvement from the traditional LoG detector (without the local detection step) is approximately 13% in terms of f* and 37% in terms of TPR at FPR*=0.01, which shows the problems that global thresholding can cause when studying varying and/or noisy environments.

To further demonstrate the sensitivity of the three best performing methods to parameter changes, we varied the parameters of the detectors and studied their influence on TPR and FPR*. We also included traditional LoG detector to show the power of the additional local detection procedure. The resulting FROC-curves are shown in Figure 21.

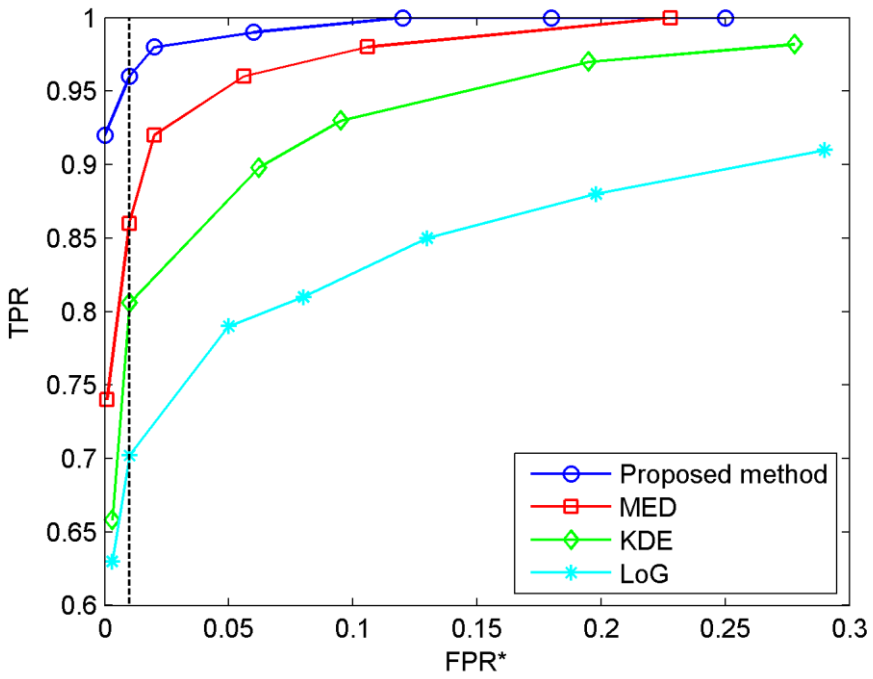


Figure 21: *FROC-curves of the three best performing methods and traditional LoG detector with simulated E. coli dataset. False positive rate of 0.01 is marked with the vertical black dashed line.*

Figure 21 shows that the proposed method is the best performing detector overall. It also appears to be very robust to parameter changes since the TPR only decreased 4% when FPR* decreased from 0.25 to 0.01. Meanwhile, MED, KDE and LoG seem to be more sensitive to parameter changes (TPR decreased approximately 14%, 16% and 29% within the same range, respectively). MED method, which uses the cell intensities and local thresholding as described in Section 3.1.2, is able to achieve decent performance while KDE and especially the traditional LoG detector seem to have more problems with their global thresholding. Notably, the proposed method is the only method able to achieve TPR over 0.9 within the whole FPR* range.

5.1.2 Results for 'Subcell' dataset

Recently, it was reported that the supervised method LR-MRF [13] outperformed all the 11 unsupervised methods described in [8] (and mostly used in this thesis as well) on the dataset named *Subcell* [45]. This dataset consists of 20 low quality synthetic images containing cells with nuclei and subcellular objects. One example image of the dataset, which is available at *http:\\www.cs.tut.fi\sgn\csb\simcep\benchmark\subcell_low.zip*, is shown in Figure 22. The dataset utilizes multichannel representation, i.e. cytoplasms are labeled with red, nuclei are labeled with blue and, subcellular objects are labeled with green. Several sources of errors are present, such as nonuniform illumination, blurring, detector noise and intensity variations in cell texture. Detailed parameters are presented in [45]. Before the analysis, the standard conversion of 0.2989 * R + 0.5870 * G + 0.1140 * B was used to convert the images to grayscale.

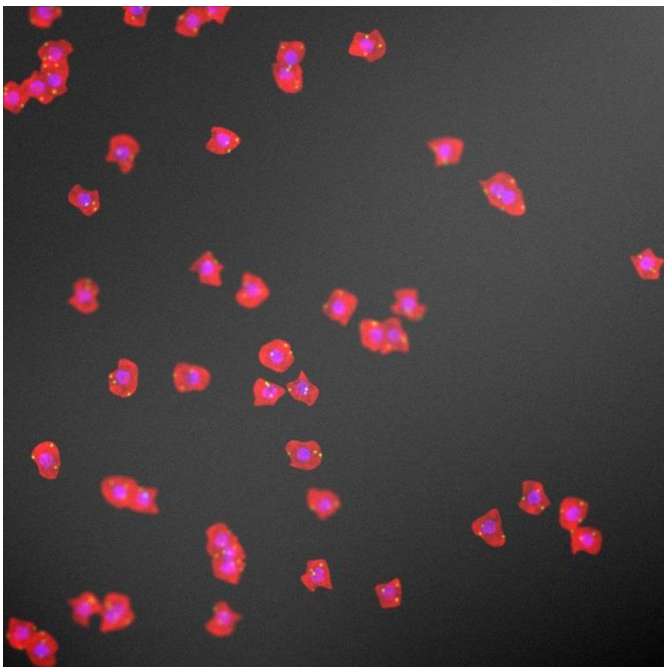


Figure 22: *Example image of the low quality Subcell dataset. Spots are shown in yellow, cytoplasm of the cell is shown in red and nuclei are shown in purple.*

Since the cell segmentation procedure described in the Section 4.4.1 is currently only for rod-like bacteria, the labels for the cells were obtained with following procedure. First, background was removed using a median filter of size 50x50. Then, the Otsu´s method was applied to find the optimal threshold automatically and the threshold was applied as described in Section 3.1.3. Finally, small objects with less than 10 pixels were removed and the morphological closing operation [29] was used to fill any holes in the detected cells. With this amount of overlapping, automated separation of the cells is almost impossible so we used simply the whole detected segment as a local neighborhood. With more accurate cell segmentation, we hypothesize that the result would be even higher. As in [8], [13], the best reachable F-score was searched in order to compare between the methods. Note that TPR at FPR*=0.01 was not calculated since it was not reported in the original article [13] where the LR-MRF was presented. The results of these three algorithms are shown in Table 3.

Table 3: *Comparison with LR-MRF and ATLAS on the low quality Subcell dataset.*

Method	p	r	f*
Proposed method	0,957	0,858	**0,905**
LR-MRF	0,876	0,858	0,867
ATLAS	0,881	0,635	0,734

The results demonstrate that our method achieves the best f*, outperforming ATLAS [10], the supervised method LR-MRF [13] and therefore all the 11 unsupervised methods presented in [8]. The increase in f* compared to LR-MRF is 4.4%. As seen from the recall and precision values, the precision of the proposed method is well above the others indicating that the algorithm rarely does a wrong decision in classification. Notably, this result demonstrates, although the proposed method was designed for *E. coli*, that it is easily applicable to other cell types as well.

5.2 Evaluation on Empirical Image Data

Next, we tested the algorithms with empirical image data in order to validate the performance in real environments. We acquired two different datasets using the procedure described in Section 4.2. Both sets consist of images of *E. coli* but with different proteins. The first set uses MS2-GFP protein which can be used to measure e.g. RNA production in bacteria [1]. The second set uses a fast maturating YFP variant, Venus, fused with membrane protein Tsr [2], which can be used e.g. to study translation in single cells. These sets have very different image characteristics, for example, MS2-GFP images have high dynamic range, i.e. the difference between the smallest and highest value is large and the intensity of spots to be detected can vary widely even with a single cell. On the other hand, low signal-to-noise ratio is peculiar to Tsr-Venus images, requiring a good performance from algorithms also when there is a high amount of noise present.

With real images, the ground truth is not available, unlike synthetic images. Therefore, the spots of interest have been manually marked by an expert and these labels then served as ground truth in the validation.

5.2.1 MS2-GFP dataset

The first set consists of 8 images of *E. coli* cells with MS2-GFP proteins tagged to RNA molecules (identified by the marked labels), and is illustrated in Figure 23. The total number of cells in the set was 180 and the total number of spots to be detected was 76.

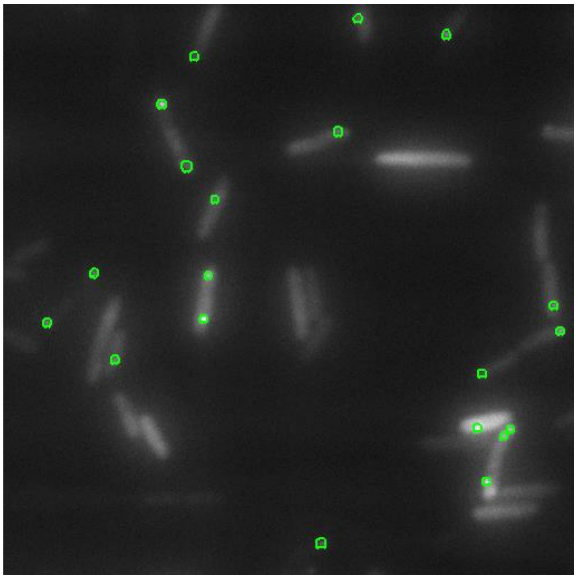


Figure 23: *Example image of the MS2-GFP dataset. Spots are labeled with green circles.*

The best reachable F-score and TPR at FPR*=0.01 for each algorithm were searched and the results are reported in Table 4 together with the precision and recall values. From the data, we can observe that several algorithms including the proposed method, KDE, LC, LEF and SE behave decently regarding their F-score (over 0.85). According to the precision and recall values, some of the algorithms were able to achieve very similar (LC, LEF) or even higher (THE) recall value but at the cost of weaker precision. Notably, if the algorithms are allowed to have only small numbers of false positives, the performance of the algorithms decreases quickly. Only two detectors are able to achieve a TPR over 0.80 and the proposed method is the only one achieving a TPR over 0.85. The increase from the second best algorithm, LC with value 0.816, is over 6%. The traditional LoG detector works decently in terms of f* but its TPR is quickly decreased when FPR*=0.01. This is most likely due to the global thresholding of the algorithm; it fails to detect spots from images which have high dynamic range.

Table 4: *Performance of the detectors with MS2-GFP data*

Method	p	r	f*	TPR at FPR* = 0.01
Proposed method	0,910	0,934	**0,922**	**0,868**
MDE	0,838	0,810	0,827	0,513
HD	0,812	0,737	0,772	0,434
KDE	0,887	0,829	0,857	0,539
LC	0,897	0,921	0,909	0,816
LEF	0,816	0,934	0,871	0,658
LoG	0,833	0,789	0,811	0,500
MW	0,865	0,842	0,853	0,684
SE	0,889	0,842	0,865	0,645
THE	0,605	0,987	0,750	0,237
ATLAS	-	-	-	-

All in all, the proposed algorithm outperformed all the other methods with both evaluation metrics. Note that ATLAS did not manage to segment any of the spots due the problems with automatic scale selection, similarly to what occurred when applied to the simulated *E. coli* data.

We also plotted the FROC-curves for the three best performing methods in order to evaluate the effect of parameter changes on TPR and FPR*. The curves are shown in Figure 24. As observed from the figure, proposed method together with LC performed reasonably well. The best overall performance was achieved by the proposed method showing that it is robust to parameter changes (TPR decreased only 9% when FPR* decreased from 0.25 to 0.01, whereas the decrease was approximately 12%, 31% and 40% in case of LC, LEF and LoG, respectively. Especially, LEF and LoG performed notably worse with low FPR*.

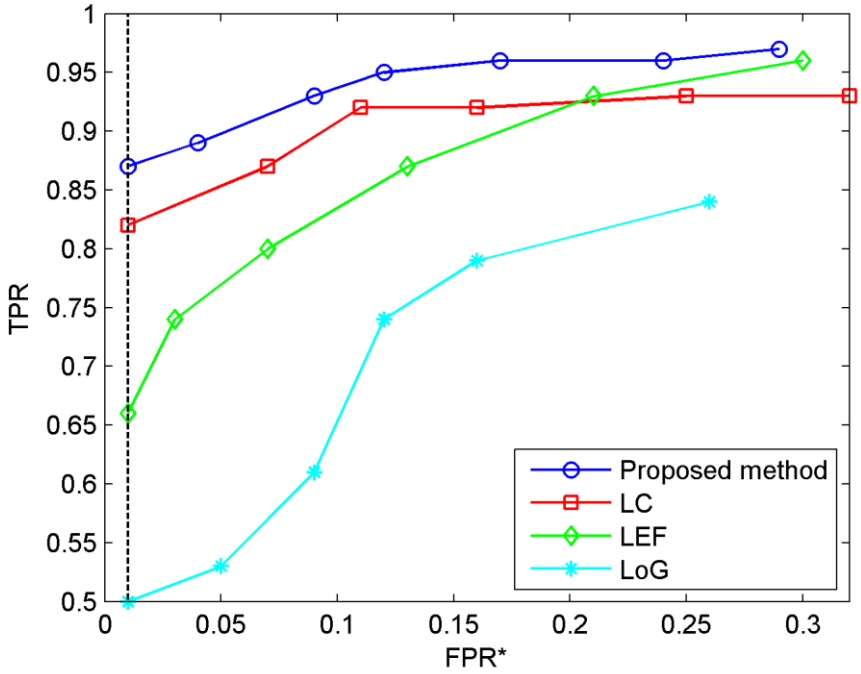


Figure 24: *FROC-curves of the three best performing methods and the traditional LoG detector with MS2-GFP dataset. False positive rate of 0.01 is marked with the black dashed line.*

5.2.2 Tsr-Venus dataset

Next, we used the Tsr-Venus protein dataset, whose images are often characterized by noisy background and low signal-to-noise ratio. The set consists of 13 images containing a total of 177 cells and 145 spots. One example image of this dataset is shown in Figure 25, together with labels of the spots.

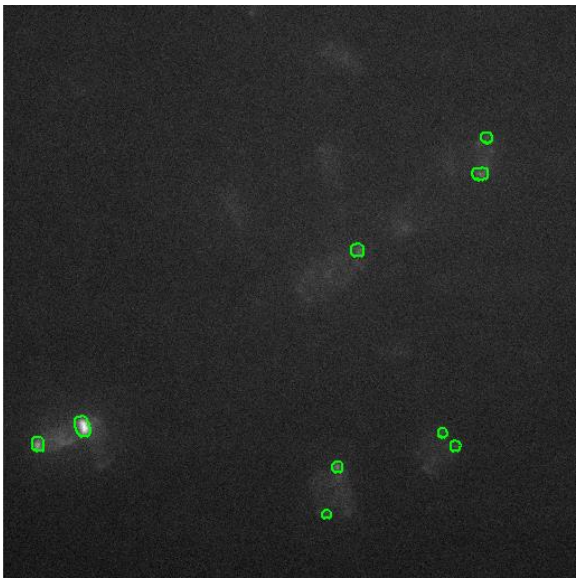


Figure 25: *Example image of Tsr-Venus dataset. Spots are labeled with green circles.*

Note that the labeling of the spots was done from the time-series and the visual comparison between neighboring frames was used as a means to decide if the object was present.

Similarly to previous analysis, the maximum F-score* and TPR at FPR*=0.01 were measured together with the precision and recall values. The results are shown in Table 5.

Table 5: *Performance of the detectors with Tsr-Venus data*

Method	p	r	f*	TPR at FPR* = 0.01
Proposed method	0,950	0,924	**0,937**	**0,835**
MDE	0,710	0,862	0,779	0,466
HD	0,567	0,931	0,705	0,386
KDE	0,729	0,855	0,787	0,138
LC	0,949	0,897	0,922	0,522
LEF	0,920	0,559	0,695	0,269
LoG	0,886	0,910	0,898	0,745
MW	0,709	0,807	0,755	0,366
SE	0,914	0,662	0,768	0,331
THE	0,842	0,959	0,897	0,655
ATLAS	-	-	-	-

Again, the proposed method outperformed all the other methods in terms of best achievable F-score* and TPR at FPR*=0.01. Some of the methods, namely, LC, SE, LoG and THE have a reasonable f* but the performance decreased quite rapidly when low false positive rate was required. This is most likely due to the characteristic noisy background of Tsr-Venus images, which usually hampers the detection of spots with high precision. The proposed detector was the only method having both precision and recall over 0.9 showing again superior performance. Some of the methods had better recall value but at the cost of decreased precision. Only the proposed algorithm was able to achieve TPR > 0.8 being over 12% better than the second best performing algorithm which is interestingly the traditional LoG detector in this case. This shows the advantage of LoG-detector, it was able to reduce the noise efficiently while still preserving the spots. Also, the global thresholding was now more satisfactory since the images did not have as large contrast variations than in MS2-GFP case.

The FROC curves were calculated for the three best performing algorithms and for the traditional LoG detector as in above. The curves are shown in Figure 26. We can observe that when FPR* is over 0.1, all detectors work reasonably well achieving TPR over 0.9. However, the performance of algorithms with FPR < 0.1 is more important since the false positives can bias the results by including false spots to data. When con-

sidering the FPR in range 0-0.1, the proposed method outperforms the other methods. It is also the only method able to achieve TPR over 0.8 at FPR*=0.01 level. Also, the decrease in TPR is only ~14% within FPR* range 0.01-0.25 when the corresponding values for LC, THE and LoG are roughly 43%, 34% and 22% showing that the proposed method is the most robust method to parameter changes.

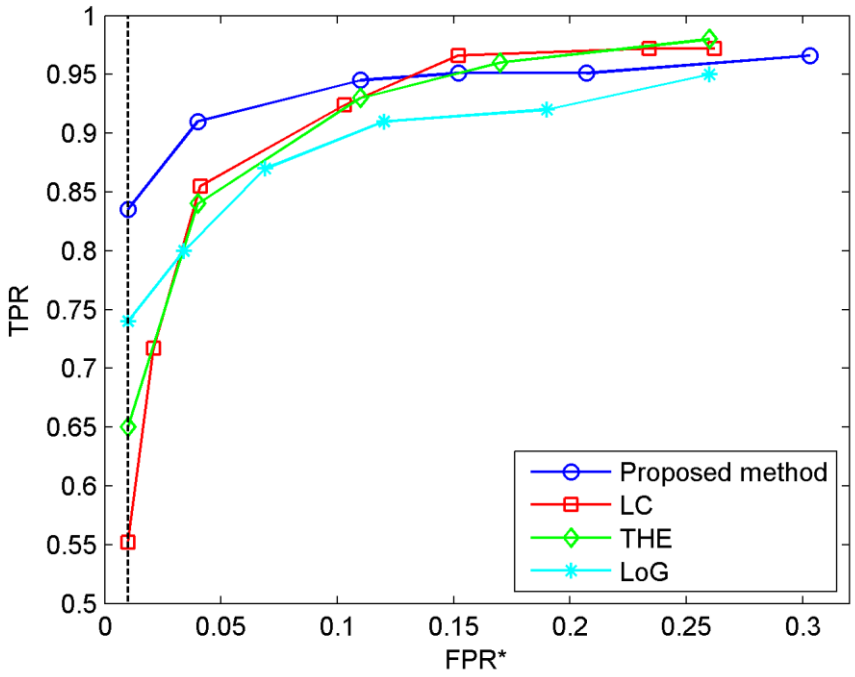


Figure 26: *FROC-curves of the three best performing methods and the traditional LoG-detector with Tsr-Venus dataset. False positive rate of 0.01 is marked with the black dashed line.*

All in all, the evaluation on the empirical image data demonstrates the power of the proposed method. It was able to achieve good performance with both MS2-GFP and Tsr-Venus data, especially when low FPR* was required. It showed superior performance and lowest sensitivity to all image data (simulated *E. coli*, Subcell dataset, MS2-GFP and Tsr-Venus). Among the best performing methods, it also showed the lowest sensitivity to parameter changes. Apart from the proposed detector, not a single method was able to reach good performance with all datasets. In fact, LC is the only method which is in the top three with both MS2-GFP and Tsr-Venus datasets together with proposed algorithm. However, it seems that the performance of the LC detector is still very sensitive to image data, as shown with simulated *E. coli* and Subcell [8] datasets. Notably, the MED detector which uses cell intensities as a prior for finding the local threshold was not able to achieve good results with the empirical image data when the low number of false positives was required. This further demonstrates the power of the proposed detector; the global blob detection procedure ensures that the candidates have blob-like features, while the local threshold is then able to cope with varying backgrounds and textures.

6. CONCLUSION

Escherichia coli has an important role as a model organism in molecular biology since it is well studied bacterium with a great number of potential applications, varying from studies of gene expression to synthetic genetic circuits. In these, the detection of subcellular particles has recently become one of the first 'standard' steps during the analysis and, therefore, the performance of image analysis algorithms in detection of objects is crucial, since it determines the success of following steps. Since the studies are often conducted with live cells, illumination intensity and exposure time is usually minimized to prevent photobleaching and photodamage to the cells resulting in low SNR images. The goal of this thesis was to develop a novel spot detection able to perform well with such low SNR images.

There are multiple algorithms proposed for detecting the subcellular objects in fluorescent images [7], [8]. As shown in [7], the difference in performance of the algorithms becomes negligible when SNR is sufficiently high. However, when SNR is low, the performance of the algorithms quickly decreases and the sensitivity to different types of data increases. Therefore, the subsequent analysis might be biased by too many false particles or by not detecting enough particles biasing the analysis towards more clearly distinguishable objects.

To overcome this, we have proposed a novel spot detection method for fluorescently labeled particles in *E. coli* which achieves a good performance with the low SNR images. It combines a Laplacian of Gaussian blob detector together with an adaptive local thresholding in order to get a better estimate for the threshold. In the first stage, a scale-space is constructed by filtering the image with the Laplacian of Gaussian kernel at different scales and a right scale for each particle is found via an automatic scale-selection procedure as in [11], [12]. Then, a low global threshold is applied to acquire the spot candidates. In the second stage, a local thresholding procedure is applied to the spot candidates based on the local cell characteristics allowing more accurate selection of the threshold. This is the major advantage of the method: the local thresholding procedure is only applied on the blob-like candidates detected by the LoG-detector.

We evaluated the proposed method by using both synthetic and empirical images. The algorithm was compared to previous methods [8], [13], [10] in order to validate the performance objectively. The results indicate that the proposed detector outperforms previous methods with both synthetic and empirical images, showing the robustness of the algorithm to different types of data. Furthermore, we have shown that the method is also

robust to parameter changes; the level of TPR remains high even when only low number of false positives is allowed. Thus, the method is very suitable for situations where low FPR is required. In addition, we also tested the algorithm with the low quality Subcell dataset [41], [42], which consists of round cells with cytoplasm, nuclei and subcellular objects. Again, the propose method outperformed the others, indicating its applicability to cells with a morphology that differs widely from that of *E. coli*.

In the future, the scope of applicability of the proposed method could be further expanded by generalizing the concept of local neighborhood, which would allow spot detection in situations where cell segmentation information is not available. However, we note that this is not a typical scenario in single-cell studies. Also, depending on application, the background model could be modified based on the prior information on the cells' properties. This would allow more accurate models, thus improving the detection accuracy even further. With time-series data, the method could be also further improved by utilizing information from adjacent frames.

REFERENCES

[1] I. Golding, J. Paulsson, S. M. Zawilski, and E. C. Cox, “Real-time kinetics of gene activity in individual bacteria,” *Cell*, vol. 123, no. 6, pp. 1025–1036, 2005.

[2] J. Yu, J. Xiao, X. Ren, K. Lao, and X. S. Xie, “Probing gene expression in live cells, one protein molecule at a time,” *Science*, vol. 311, no. 5767, pp. 1600–3, Mar. 2006.

[3] A.-B. Muthukrishnan, M. Kandhavelu, J. Lloyd-Price, F. Kudasov, S. Chowdhury, O. Yli-Harja, and a. S. Ribeiro, “Dynamics of transcription driven by the tetA promoter, one event at a time, in live Escherichia coli cells,” *Nucleic Acids Res.*, vol. 40, no. 17, pp. 8472–8483, 2012.

[4] M. Kandhavelu, A. Häkkinen, O. Yli-Harja, and a S. Ribeiro, “Single-molecule dynamics of transcription of the lar promoter,” *Phys. Biol.*, vol. 9, no. 2, p. 026004, 2012.

[5] E. Meijering, I. Smal, and G. Danuser, “Tracking in molecular bioimaging,” *IEEE Signal Process. Mag.*, vol. 23, no. 3, pp. 46–53, 2006.

[6] E. Meijering, I. Smal, O. Dzyubachyk, and J. C. Olivo-Marin, “Time-Lapse Imaging,” in *Microscope Image Processing*, Q. Wu, F. A. Merchant, and K. R. Castleman, Eds. Burlington: Elsevier, 2008, pp. 401–440.

[7] I. Smal, M. Loog, W. Niessen, and E. Meijering, “Quantitative comparison of spot detection methods in fluorescence microscopy,” *IEEE Trans. Med. Imaging*, vol. 29, no. 2, pp. 282–301, 2010.

[8] P. Ruusuvuori, T. Äijö, S. Chowdhury, C. Garmendia-Torres, J. Selinummi, M. Birbaumer, A. M. Dudley, L. Pelkmans, and O. Yli-Harja, “Evaluation of methods for detection of fluorescence labeled subcellular objects in microscope images,” *BMC Bioinformatics*, vol. 11, p. 248, 2010.

[9] J.-C. Olivo-Marin, “Extraction of spots in biological images using multiscale products,” *Pattern Recognit.*, vol. 35, no. 9, pp. 1989–1996, Sep. 2002.

[10] A. Basset, J. Boulanger, J. Salamero, P. Bouthemy, and C. Kervrann, “Adaptive spot detection with optimal scale selection in fluorescence microscopy images,” *IEEE Trans. image Process.*, vol. 24, no. 11, pp. 4512–4527, 2015.

[11] T. Lindeberg, “Detecting Salient Blob-Like Image Structures and Their Scales with Scale-Space Primal Sketch: A method for Focus-of-Attention,” *Int. J. Comput. Vis.*, vol. 11, no. 3, pp. 283–318, 1993.

[12] T. Lindeberg, “Feature Detection with Automatic Scale Selection,” *Int. J. Comput. Vis.*, vol. 30, pp. 79 – 116, 1998.

[13] P. Ruusuvuori, T. Manninen, and H. Huttunen, “Image segmentation using sparse logistic regression with spatial prior,” in *20th European Signal Processing Conference*, 2012, pp. 2253–2257.

[14] B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter, *Molecular Biology of the Cell*, 4th ed. New York: Garland Science, 2002.

[15] A. B. Lindner, R. Madden, A. Demarez, E. J. Stewart, and F. Taddei, “Asymmetric segregation of protein aggregates is associated with cellular aging and rejuvenation,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 105, no. 8, pp. 3076–81, Feb. 2008.

[16] I. Golding and E. C. Cox, “RNA dynamics in live Escherichia coli cells,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 101, no. 31, pp. 11310–5, 2004.

[17] M. B. Elowitz and S. Leibler, “A synthetic oscillatory network of transcriptional regulators,” *Nature*, vol. 403, no. 6767, pp. 335–338, 2000.

[18] A. Häkkinen, M. Kandhavelu, S. Garasto, and A. S. Ribeiro, “Estimation of fluorescence-tagged RNA numbers from spot intensities,” *Bioinformatics*, vol. 30, no. 8, pp. 1146–1153, 2014.

[19] L. W. Miller and V. W. Cornish, “Selective chemical labeling of proteins in living cells,” *Curr. Opin. Chem. Biol.*, vol. 9, pp. 56–61, 2005.

[20] M. Chalfie, Y. Tu, G. Euskirchen, W. W. Ward, and D. C. Prasher, “Green fluorescent protein as a marker for gene expression,” *Science (80-.).*, vol. 263, pp. 802–805, 1994.

[21] D. J. Stephens and V. J. Allan, “Light Microscopy Techniques for Live Cell Imaging,” *Science (80-.).*, vol. 300, no. 5616, pp. 82–87, 2003.

[22] A. J. Ridley, J. R. Whiteside, T. J. McMillan, and S. L. Allinson, “Cellular and sub-cellular responses to UVA in relation to carcinogenesis,” *Int. J. Radiat. Biol.*, vol. 85, no. 3, pp. 177–195, 2009.

[23] D. Axelrod, “Cell-substrate contacts illuminated by total internal reflection fluorescence,” *J. Cell Biol.*, vol. 89, pp. 141–145, 1981.

[24] N. Grossman, E. Z. Ron, and C. L. Woldringh, “Changes in cell dimensions during amino acid starvation of Escherichia coli,” *J. Bacteriol.*, vol. 152, no. 1, pp. 35–41, 1982.

[25] M. Tokunaga, N. Imamoto, and K. Sakata-sogawa, “Highly inclined thin illumination enables clear single-molecule imaging in cells,” *Nat. Methods*, vol. 5, no. 2, pp. 159–161, 2008.

[26] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, “Image denoising by sparse 3-D transform-domain collaborative filtering,” *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080–2095, 2007.

[27] L. Van Vliet, D. Sudar, and I. Young, “Digital fluorescence imaging using cooled charge-coupled device array cameras,” in *Cell Biology*, 2nd ed., vol. III, J. E. Celis, Ed. New York: Academic Press, USA, 1998, pp. 109–120.

[28] M. Monici, “Cell and tissue autofluorescence research and diagnostic applications,” *Biotechnol. Annu. Rev.*, vol. 11, pp. 227–256, 2005.

[29] R. C. Gonzales and R. E. Woods, *Digital Image Processing*, 2nd ed. New Jersey: Prentice Hall, USA, 2002.

[30] A. Gupta, J. Lloyd-Price, R. Neeli-Venkata, S. M. D. Oliveira, and A. S. Ribeiro, “In Vivo Kinetics of Segregation and Polar Retention of MS2-GFP-RNA Complexes in Escherichia coli.,” *Biophys. J.*, vol. 106, no. 9, pp. 1928–37, May 2014.

[31] L. Vincent, “Morphological Gray scale Reconstruction in Image Analysis : Applications and Efficient Algorithms,” *IEEE Trans. image Process.*, vol. 2, no. 2, pp. 176–201, 1993.

[32] P. Soille, *Morphological Image Analysis: Principles and Applications*. Berlin: Springer-Verlag, Germany, 2004.

[33] E. Bertin and S. Arnouts, “SExtractor: Software for source extraction,” *Astron. Astrophys. Suppl. Ser.*, vol. 117, pp. 393–404, 1996.

[34] N. Otsu, “A Threshold Selection Method from Gray-Level Histograms,” *IEEE Trans. Syst. Man. Cybern.*, vol. 9, no. 1, pp. 62–66, 1979.

[35] A. Martikainen, “Segregation of protein aggregates in Escherichia coli,” M.S. thesis, Department of Signal Processing, Tampere University of Technology, Finland, 2015.

[36] T. Fawcett, “An introduction to ROC analysis,” *Pattern Recognit. Lett.*, vol. 27, no. 8, pp. 861–874, 2006.

[37] D. P. Chakraborty, “Maximum likelihood analysis of free-response receiver operating characteristic (FROC) data,” *Med. Phys.*, vol. 16, no. 4, pp. 561–568, 1989.

[38] D. P. Chakraborty and L. H. Winter, “Free-response methodology: alternate analysis and a new observer-performance experiment,” *Radiology*, vol. 174, pp. 873–881, 1990.

[39] L. Lee, T. Mizuno, and Y. Imae, “Thermosensing properties of Escherichia coli tsr mutants defective in serine chemoreception,” *J. Bacteriol.*, vol. 170, no. 10, pp. 4769–74, Oct. 1988.

[40] A. Rebbapragada, M. S. Johnson, G. P. Harding, A. J. Zuccarelli, H. M. Fletcher, I. B. Zhulin, and B. L. Taylor, “The Aer protein and the serine chemoreceptor Tsr independently sense intracellular energy levels and transduce oxygen, redox, and

energy signals for Escherichia coli behavior," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 94, no. 20, pp. 10541–6, Sep. 1997.

[41] A. Lehmussola, P. Ruusuvuori, J. Selinummi, H. Huttunen, and O. Yli-Harja, "Computational framework for simulating fluorescence microscope images with cell populations," *IEEE Trans. Med. Imaging*, vol. 26, no. 7, pp. 1010–1016, 2007.

[42] A. Lehmussola, P. Ruusuvuori, J. Selinummi, T. Rajala, and O. Yli-Harja, "Synthetic images of high-throughput microscopy for validation of image analysis methods," *Proc. IEEE*, vol. 96, no. 8, pp. 1348–1360, 2008.

[43] K. Perlin, "An image synthesizer," *Comput. Graph. (ACM).*, vol. 19, no. 3, pp. 287–296, 1985.

[44] H. Kaufman and M. Tekalp, "Survey of estimation techniques in image restoration," *IEEE Control Syst.*, vol. 11, no. 1, pp. 16–24, 1991.

[45] P. Ruusuvuori, A. Lehmussola, J. Selinummi, T. Rajala, H. Huttunen, and O. Yli-harja, "Benchmark Set of Synthetic Images for Validating Cell Image Analysis Algorithms," in *Proceedings of the 16th European Signal Processing Conference*, 2008, no. EUSIPCO.

[46] S. Chowdhury, M. Kandhavelu, O. Yli-Harja, and A. S. Ribeiro, "Cell segmentation by multi-resolution analysis and maximum likelihood estimation (MAMLE)," *BMC Bioinformatics*, vol. 14 Suppl 1, no. Suppl 10, p. S8, Jan. 2013.

[47] A. Häkkinen, A.-B. Muthukrishnan, A. Mora, J. M. Fonseca, and A. S. Ribeiro, "CellAging: a tool to study segregation and partitioning in division in cell lineages of Escherichia coli," *Bioinformatics*, vol. 29, no. 13, pp. 1708–9, Jul. 2013.

[48] A. R. Webb and K. D. Copsey, *Statistical Pattern Recognition*, 3rd ed. John Wiley and Sons, United Kingdom, 2011.

APPENDIX 1: PARAMETER LISTS

Table A1: *Parameter list for simulated E. coli dataset. Detailed explanation for the parameters of the previous methods can be found in* [8].

Method	**Parameter**	**Parameter value (for f*)**	**Parameter value (for TPR at FPR*=0.01)**
Proposed method	th_local th_global	0.07 0.002 (fixed)	0.06 0.002 (fixed)
LoG	th	0.006	0.0085
MED	th	0.02	0.005
HD	s h nb_size radius win_size σ_M	5 0.3 7 3 15 5	5 0.4 7 3 15 5
KDE	R h	14 0.06	13 0.1
LC	R α	7 0.87	5 0.88
LEF	th_s	2.5	3.29
MW	J l_d	3 650	3 380
SE	block_size th_{BG} th_{detect}	25 0.2 1.8	25 0.26 1.85
THE	r σ (Gaussian filter)	5 1.4	5 0.8
ATLAS	-	-	-

Table A2: *Parameter list for MS2-GFP dataset. Detailed explanation for the parameters of the previous methods can be found in* [8].

Method	**Parameter**	**Parameter value (for f*)**	**Parameter value (for TPR at FPR*=0.01)**
Proposed method	th_local	0.002	0.001
	th_global	0.002 (fixed)	0.002 (fixed)
LoG	th	0.01	0.02
MED	th	0.005	0.001
HD	s	7	7
	h	0.35	0.5
	nb_size	7	12
	radius	3	3
	win_size	8	10
	σ_M	5	5
KDE	R	10	10
	h	0.08	0.14
LC	R	8	5
	α	0.8	0.76
LEF	th_s	1.7	2.5
MW	J	3	3
	l_d	430	1900
SE	block_size	32	32
	th_{BG}	0.5	0.6
	th_{detect}	1.5	1.5
THE	r	3	2
	σ (Gaussian filter)	1.25	0.5
ATLAS	-	-	-

Table A3: *Parameter list for Tsr-Venus dataset. Detailed explanation for the parameters of the previous methods can be found in* [8].

Method	**Parameter**	**Parameter value (for f*)**	**Parameter value (for TPR at FPR*=0.01)**
Proposed method	th_local th_global	0.02 0.002 (fixed)	0.008 0.002 (fixed)
LoG	th	0.004	0.008
MED	th	0.01	0.0001
HD	s h nb_size radius win_size σ_M	7 0.2 7 3 8 5	7 0.5 9 3 10 5
KDE	R h	8 0.04	4 0.15
LC	R α	8 0.9	5 0.9
LEF	th_s	2.1	3.1
MW	J l_d	3 350	3 2060
SE	block_size th_{BG} th_{detect}	25 0.3 0.8	32 0.6 0.3
THE	r σ (Gaussian filter)	4 1.4	4 0.6
ATLAS	-	-	-

Table A4: *Parameter list for low quality Subcell dataset. Parameter descriptions for ATLAS can be found in* [10].

Method	**Parameter**	**Parameter value (for f*)**
Proposed method	th_local th_global	0.12 0.002 (fixed)
ATLAS	th win_size	0.006 15