Cytometry
PART A
Journal of the
International Society for
Advancement of Cytometry

Comparison of Parameter-Adapted Segmentation Methods for Fluorescence Micrographs

Christian Held,[1,2] Ralf Palmisano,[2] Lothar Häberle,[3] Michael Hensel,[2,4] Thomas Wittenberg[1]*

[1]Department of Image Processisng and Biomedical Engineering, Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany

[2]Institute of Microbiology, University Hospital, Erlangen, Germany

[3]Department of Obstetrics and Gynecology, Erlangen University Hospital, Germany

[4]Division of Microbiology, University of Osnabrück, Osnabrück, Germany

*Correspondence to: Thomas Wittenberg, Department of Image Processing and Medical Engineering, Fraunhofer Institute for Integrated Circuits IIS, Am Wolfsmantel 33, 91058 Erlangen, Germany

Email: thomas.wittenberg@iis.fraunhofer.de

ISAC
International Society for Advancement of Cytometry

- **Abstract**
Interpreting images from fluorescence microscopy is often a time-consuming task with poor reproducibility. Various image processing routines that can help investigators evaluate the images are therefore useful. The critical aspect for a reliable automatic image analysis system is a robust segmentation algorithm that can perform accurate segmentation for different cell types. In this study, several image segmentation methods were therefore compared and evaluated in order to identify the most appropriate segmentation schemes that are usable with little new parameterization and robustly with different types of fluorescence-stained cells for various biological and biomedical tasks. The study investigated, compared, and enhanced four different methods for segmentation of cultured epithelial cells. The maximum-intensity linking (MIL) method, an improved MIL, a watershed method, and an improved watershed method based on morphological reconstruction were used. Three manually annotated datasets consisting of 261, 817, and 1,333 HeLa or L929 cells were used to compare the different algorithms. The comparisons and evaluations showed that the segmentation performance of methods based on the watershed transform was significantly superior to the performance of the MIL method. The results also indicate that using morphological opening by reconstruction can improve the segmentation of cells stained with a marker that exhibits the dotted surface of cells.   © 2011 International Society for Advancement of Cytometry

- **Key terms**
image analysis; cell segmentation; optimization; genetic algorithm; watershed; fluorescence imaging; evaluation

THE interpretation of fluorescence micrographs plays an important role in the analysis of cellular events such as cell growth and division, cell death, apoptosis, intercellular communication, and microbial interactions with host cells. For statistical reasons, typical experiments in this field usually require detection and visual assessment of at least 100 individual specimens, which are most often prepared using several dyes and antibodies.

The tasks involved in analyzing such images typically include detection, selection, segmentation, measuring, and counting cells and cellular interactions. Manual analysis of experimental samples of this type is possible when there are only a few cells and images. When larger numbers of images, cells, or events are to be analyzed, however, the task becomes extremely tedious, repetitive, and time-consuming (1). Specifically, as stated by Vidal et al. (2): "Segmentation is often the major bottleneck in clinical applications—it takes a long time and the results are often hard to reproduce because of the user involvement." This is most often due to a lack of attention, and biased interpretation is also possible. Experimental results of this type are therefore usually hard to reproduce in intralaboratory and interlaboratory experiments. There is a strong need for objective image analysis tools in order to increase the reproducibility and effectiveness of this type of image-based evaluation. Unfortunately, hardly any generally applicable sets of image-processing methods or tools are available to support leading-edge medical and biomedical research—sets of tools that
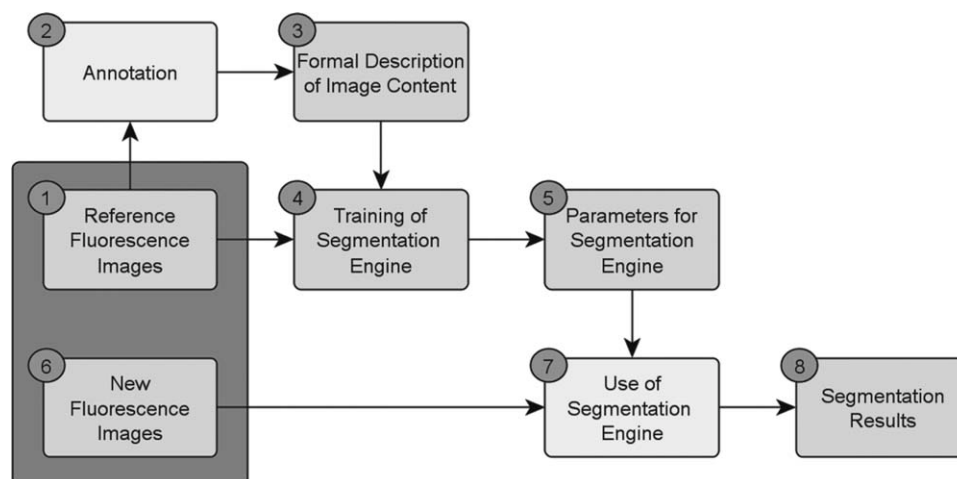
Figure 1. The building blocks of an adaptive segmentation scheme (see text for explanation). Reproduced with permission from Wittenberg et al. (3).

could be adopted on the fly for frequently changing analytical tasks for fully or partly automated analysis of micrographs (3). As Srinivasa noted (4), "While there has been an increasing effort to automate analysis of biological images, tools to meet the various challenges posed by specific applications in this area are still in their infancy." Most image analysis and interpretation in this field is therefore still being done manually.

On the other hand, a strong background in automated image processing and image interpretation is needed in order to adapt available image analysis tools to the desired task, or to write new scripts in image-processing toolboxes. A survey of recent publications in the field shows that the state of the art for fully automated or interactive micrograph image analysis involves two different approaches, from opposite directions. On the one hand, there are knowledge-driven, top-down methods that are dedicated to very specific and narrow applications in the field of microscopic image analysis (as detailed in the following section here). This observation was recently supported by Wang et al. (5), who state that "depending on the cells being segmented . . . various existing algorithms are available. A universal solution to cell segmentation . . . applicable across cell types has yet to be described." Specifically, due to the differing characteristics of the cell types to be measured and assessed, image-processing methods applicable to one set of images and cells most often perform poorly on another set. On the other hand, there are also several data-driven, procedure-oriented image-processing frameworks and toolboxes available, which can be applied to the analysis of micrographs but are usually not dedicated to any specific application.

This observed mismatch—the known and still unsolved difficulty of describing the high-level semantics of cells and cellular interactions in the field of medicine and biology using adequate image analysis methods or low-level image-based feature analysis—is also known as the "semantic gap" (6). In the field of computer vision research, Garbay, for example, has also noted this gap between the symbolic apprehension of

high-level concepts (such as cells or cellular interactions) and their concrete instantiation in images (7). To bridge this semantic gap between analytical issues in the interpretation of micrographs by users, on the one hand, and the consequent (frequently changing) sets of image-processing procedures required on the other, we recently suggested using an image analysis approach (3) (Fig. 1).

Using fluorescence-stained HeLa cells as a typical example of the analysis of microbial effectors in host cells, the steps involved in the proposed adaptive image analysis scheme are as follows:

- On the basis of a small but representative set of reference images (step 1),
- An interactive segmentation (step 2) is made by a user, delineating and labeling the objects of interest, in this case, cell plasma (Fig. 2).
- The manually segmented cells are then described as so-called "ground truth" in a machine-readable manner (step 3).
- Using the annotated reference image data (step 1) and the formal image description (step 3) as input data, a so-called "segmentation engine" is applied and optimized on a set of parametric segmentation methods (step 4).
- As a result, the optimal segmentation parameters for the training set are obtained (step 5).
- Based on these parameters (step 5), further fluorescence micrographs (step 6) of the same type and image quality can then be segmented using the trained segmentation engine (step 7), yielding a set of segmentation objects (step 8) as a result.

Objectives

Against this background, the aim of the present study was to compare and evaluate different image segmentation methods for a variety of fluorescence-stained cells that could be
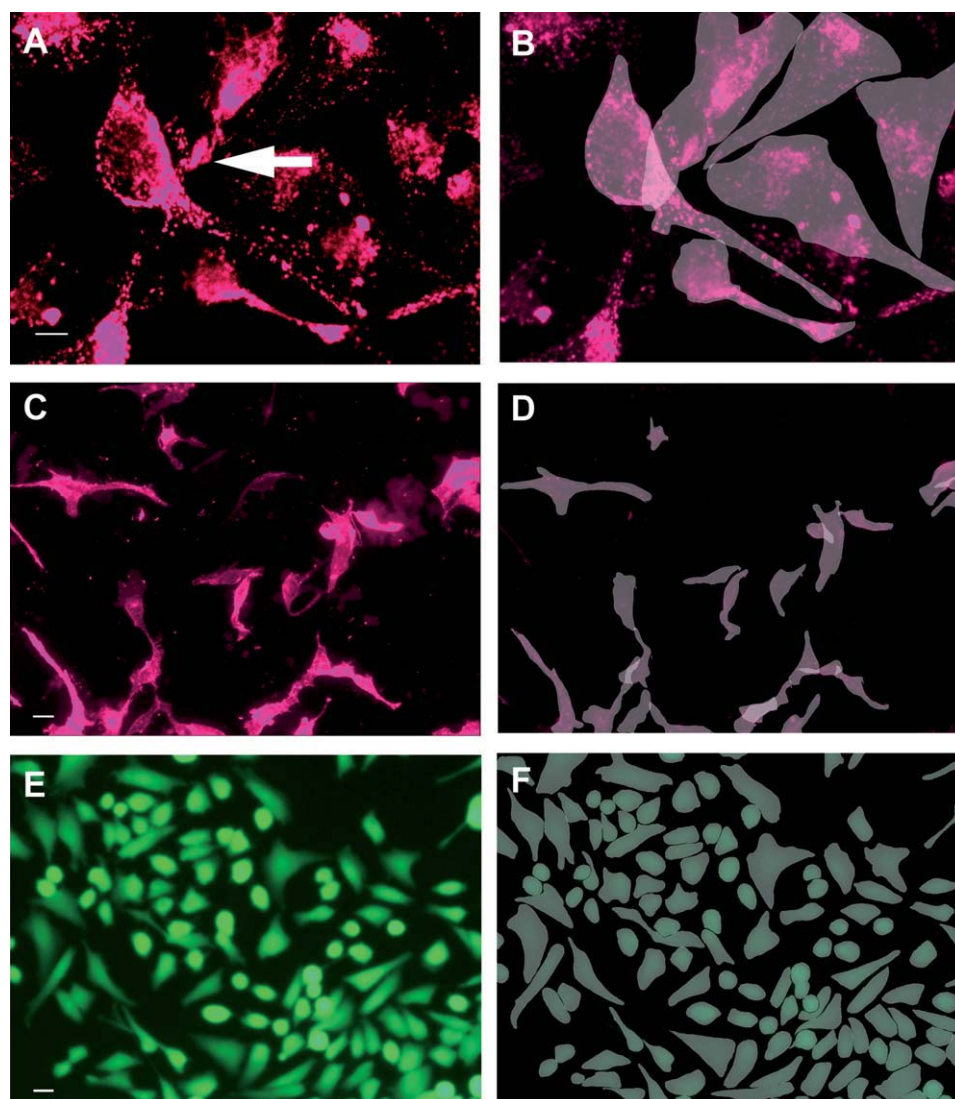
Figure 2. Example images of the various original datasets (**A**, **C**, **E**) and the corresponding manually annotated ground truth (**B**, **D**, **F**) obtained from an experienced user. The actin cytoskeleton of the HeLa cells was labeled with phalloidin Alexa 568 (A, B), the cytoplasmic membrane of the HeLa cells was labeled with DiD^{TM} (C, D), and the L929 cells were stained with the viability marker fluorescein diacetate (FDA) (E, F). The scale bars correspond to 10 μm in A and B (original magnification ×63) and 20 μm in C—F (original magnification ×20). From the point of view of automated image processing, these example images raise several challenges: segmentation of the thin extensions of the cells overlapping with other cells (e.g., the cell indicated by an arrow in A); the dotted structure of the phalloidin Alexa 568-stained cytoskeleton; and the dark areas resulting from the absence of the fluorescent marker in the cell nuclei (A). In addition, the cells differ widely in terms of size, morphology, and intensity.

incorporated into the segmentation engine described above, as core methods from which a selection can be made. Specifically, HeLa cells stained intracellularly with Alexa Fluor® 568 phalloidin (Invitrogen, Darmstadt, Germany), HeLa cells surface-stained with DiD^{TM} (Vybrant®, Invitrogen), and L929 cells surface-stained with fluorescein diacetate (FDA) were used as three typical example applications to evaluate adaptable image segmentation and analysis tools selected from our previous research. Image segmentation methods that are known to work with fluorescent cells, with parameters that can be used to steer and control the segmentation result as described above, were therefore selected and partly extended.

From the image-processing point of view, selected segmentation tasks for phalloidin Alexa 568-stained and DiD-stained HeLa cells are particularly challenging. Depending on the confluence, HeLa cells have different morphological types. Tested samples that are highly confluent may result in intercellular overlapping among adjacent cells. In these conditions, even experienced users are unable to distinguish some of the cells unambiguously and cannot outline the boundaries due to the overlaps. Indistinguishable cells of this type were excluded from the ground truth. Overlapping cells with boundaries that could be determined unambiguously were annotated as overlapping cells (Fig. 2).

A further challenge is created by the phalloidin Alexa 568 staining used in this experiment, which is not equally distributed inside the cytoplasm and shows dotted structures (Fig. 2A). The segmentation of L929 cells also poses a challenge, as the cells depicted in this dataset have very large morphological differences, with some showing a compact circular shape and others a triangular or bipolar morphology (Fig. 2C). Finally, there are always a certain number of apoptotic cells, which generally show a small, circular morphology with relatively increased intensity signals (Fig. 2).

RELATED RESEARCH AND STATE OF THE ART

As mentioned above, many image-processing methods have been proposed during the last 10 years for segmentation of single cells in fluorescence images, but most of these are dedicated to very specific and narrow applications in the field of microscopic image analysis. A broad overview of segmentation techniques for fluorescence images is presented by Restif (8).

Many different approaches have been proposed in the area of single-cell image segmentation in fluorescence micrographs. The most basic methods include local and adaptive thresholding (1,9–12), based on Otsu's thresholding approach (13), as well as region-growing methods. Due to the relative homogeneity of the statistical properties of the foreground and background in such images, these approaches are suitable for delineating single cells and cell groups in fluorescence micrographs. However, they are not able to split and separate adjacent and overlapping cells. For robust detection, segmentation, and separation of clustered cell nuclei stained with 4,6-diamino-2-phenylindole (DAPI), Nandy et al. (14) proposed an algorithm based on calculating the gradient magnitude and direction, k-means clustering, weighted distance transform (DT), and dynamic programming (DP). Du and Dua (15) compared different clustering approaches, including the expectation-maximization (EM) algorithm, k-means, thresholding, and global minimization for active contours, for real and synthetic fluorescent images. To separate clustered cell nuclei in peripheral blood and bone-marrow preparations in fluorescence in situ hybridization (FISH) images, Malpica et al. (16) used the watershed transform (WT) in combination with a level set (LS) approach. A similar chain of methods (known as "BlobFinder") for delineating a variety of cells has recently been proposed by Allalou and Wählby (1); the method consists of thresholding, DT, and WT and also allows user interaction at any time. Wählby et al. have also described a multiple-step algorithm for segmentation of Chinese hamster ovary (CHO) cells stained with calcein in fluorescence images (17). After initial segmentation using a WT, small regions are merged or deleted. Based on a statistical classification, larger image regions denoting possible cell aggregates are tested for splitting into smaller cell-like regions. On the basis of these studies (17,18), Wang et al. (5) suggested a delineation scheme for HeLa cells based on binarization, cell detection using gradient vector fields (GVFs), and seeded WT-based segmentation. Zhu et al. (19) used an automatic quantification method applicable to fluorescence imagery using local maximums to identify labeling targets and watershed segmentation to define their boundaries.

Another approach often used to delineate and separate cells in fluorescence micrographs is mathematical morphology (20), as in the studies by Metzler et al. (21,22), who used a morphological multiple-scale approach to separate mouse fibroblasts. Zhang et al. (23) used multiple-stage morphological operations to extract cell boundaries. Wang et al. (5) recently extended the morphological approach for the analysis of bacterial, yeast, and human cells using what are known as nonlinear or hybrid range filters (HRFs) with circular structuring elements.

For micrograph segmentation, active contours (snakes) and level set approaches are increasingly being regarded as the state of the art, and they have recently been used for various applications. Segmentation of fluorescence-stained HeLa cells using a stochastic active contour scheme (STACS), for example, has been suggested by Srinivasa et al. (24,25). Specifically, as HeLa cells do not show any real edges and each cell has a different shape, only an external, region-based and an internal, curvature-based force are applied to develop the contour. Contour initialization is obtained from the DNA information depicted in a neighboring channel. The STACS method was recently enhanced using combined region-growing and multiple-scale approaches (24). Möller at al. used a snake approach with an extended-energy function for segmentation of monoverlapping cells stemming from the human hepatoma cell line (HUH7 cells) (26), while Ersoy et al. (27) used level set-based multiple-phase fast graph partitioning active contours (FastGPAC), a method that is an extension of graph partitioning active contours (GPAC).

Yu et al. (28) have suggested an approach based on enhanced level set (also known as geometric active contours) segmentation, in which detected cell nuclei are used to initialize the level set function. A dynamic watershed approach was also used to prevent merging and splitting of cell segments. This approach was recently enhanced (29) by employing what is known as an evolving generalized Voronoi diagram algorithm, incorporating image intensity and geometric information. Cheng and Rajapakse (30) proposed a segmentation method for neuronal as well as *Drosophila* cell fluorescence micrographs with a level set function in which the image energy is defined using intensity variances inside and outside of the contour. The contour is initialized using a watershed approach. With regard to live cell imaging, in which the aim is to track vital cells, Dzyubachyk et al. (31,32) have proposed a level set-based cell segmentation (and tracking) method based on a model evolution approach. According to the authors, this approach ensures a high quality of segmentation with widely varying object intensities.

An approach based on artificial neural networks (ANNs) for automatic detection, localization, and segmentation of fluorescence micrographs has been presented by Nattkemper et al. (33–35). This method applies image patches of 15 × 15 pixels for the training cycles. Approximately a quarter of the training patches consist of preclassified and hand-labeled image patches depicting fluorescent cells, while the remaining

Table 1. Overview of the image datasets used for the experiments

	CELL TYPE		
	HELA CELLS	HELA CELLS	L929 CELLS
Staining	Phalloidin Alexa 568	DiD^{TM}	FDA
Magnification	63 × oil immersion	20 × oil immersion	20 × air
	1.3 NA	0.8	
	0.1 μm/pixel	0.32 μm/pixel	
Microscope	Zeiss Axiovert	Zeiss Axiovert	Olympus IX71 inversion
Image size	1,388 × 1,040 pixels	1,388 × 1,040 pixels	1,376 × 1,032 pixels
Images	60	30	10
Manually annotated cells	261	817	1,333

FDA, fluorescein diacetate.

training regions are randomly chosen background image regions without cells. For the architecture of the ANN, both a local linear map (LLM) as a variant of the self-organizing map (SOM) and also a multilayer perceptron (MLP) with a back-propagation learning scheme have been described. Both approaches are effective in learning object recognition tasks from small training sets. The two architectures yield what are known as confidence maps, describing the probability of the occurrence of a cell. To eliminate false-positive cells, these confidence maps are then further processed in successive image-processing steps.

The aim with almost all of these approaches is to provide solutions for only a single problem, such as analysis of mitotic phenotypes of human cells (9,10), human hepatoma cells (26), characterization of protein-protein interactions (11), agonist-induced translocation of green fluorescent protein (GFP) Rac1 to cellular membranes (18), or delineation of HeLa cells (4,27,36), blood and bone-marrow nuclei (16), CHO cells (17), mouse neuroblastoma neural cells (28), mouse fibroblasts (21,22), DAPI-stained cervical cell nuclei (12,14), and synthetically computed cells (15). A universal solution for cell segmentation and tracking that would be applicable across all types of cells and stains has therefore yet to be described. Algorithms that work well on one set of images often perform poorly on another set, due to differences in the features that are exploited. Among the studies mentioned above, only the approaches by Allalou and Wählby (1), Wang et al. (5), and Zhu et al. (19) attempt to cover more than one problem with the segmentation approach investigated. In addition, studies comparing the performance of different segmentation approaches for fluorescence micrographs have the common drawback that the segmentation ground truth is seldom available—as in the study by Coelho et al., for example (37). It is therefore difficult to compare the performance of the various approaches with the state of the art in this field.

CONTRIBUTION OF THE PRESENT STUDY

As mentioned above, almost all of these segmentation methods are strongly dependent on specific analytic applications, so that they cannot be directly reused for other applications without the effort involved in manual reprogramming.

In addition, when the adaptive segmentation approach described initially here is used—in which the best segmentation scheme applicable to a new set of fluorescence images is selected and its parameters are optimized automatically relative to previous manual segmentation of a representative training dataset—not every named segmentation scheme can be used in this context. In particular, methods that use additional information (such as initialization schemes) from other modalities are (not yet) applicable to the method proposed here. However, as noted by Srinivasa (4) and Wang et al. (5) in studies that investigated automated segmentation of fluorescence micrographs, the watershed algorithm is regarded as one of the most accurate methods. To evaluate the proposed self-adapting image segmentation concept, we therefore used the watershed algorithm (38) and a variant of it known as the maximum-intensity linking (MIL) approach (39), originally developed for segmentation of fluorescence-stained stem cells. All of the segmentation approaches were evaluated using threefold cross-validation on three reference image datasets depicting 261, 817, and 1,333 fluorescence-stained cells, with corresponding ground-truth data.

MATERIALS AND METHODS

Image Data

Although several image reference datasets of fluorescence images are available for public research purposes, such as the Yeast Protein Localization database (40,41), the Yeast Resource Center Public Image Repository (42), the Distributed Database for BioMolecular Images (43), and the database for dynamics and localizations of endogenous fluorescence-tagged proteins in living human cells (44), none of these repositories serves our purposes. Specifically, to the best of our knowledge none of these databases has a reference annotation for the cells depicted, and none of the prepared and depicted cells are related to the scope of our ongoing project. In order to obtain representative fluorescence image datasets applicable for the present study, therefore, three datasets were created. Table 1 presents a detailed overview of them.

- The first dataset contained 261 HeLa cells that were stained with phalloidin Alexa 568 (excitation wavelength 568 nm)
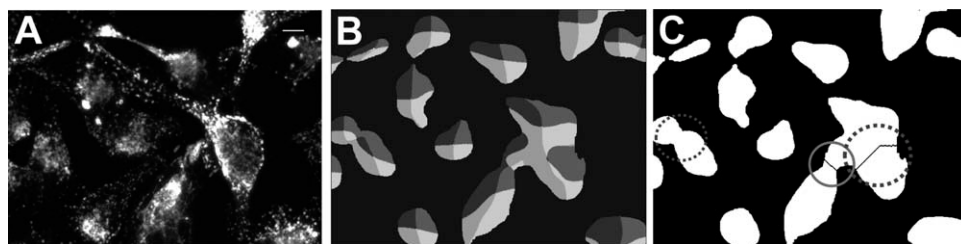
Figure 3. Example of image analysis using maximum-intensity linking (MIL). (**A**) A representative image with fluorescence-stained HeLa cells from the first dataset. (**B**) Intensity-coded visualization of the graph directions. (**C**) The MIL result. In this example, some large cells were correctly separated (C, solid circle), but there were also oversegmented and undersegmented cells (C, dotted circles). The scale bar corresponds to 10 μm (original magnification $\times$63).

to detect the F-actin cytoskeleton. For image acquisition, a Zeiss Axiovert microscope was used equipped with a 63$\times$ objective, an AxioCam, and AxioVision Capture software (Carl Zeiss MicroImaging, Jena, Germany).

- The second dataset consisted of 817 HeLa cells that were stained with DiD (excitation wavelength 644 nm), an intermittent cell membrane marker. To capture these images, a Zeiss Axiovert microscope was used in combination with a 20$\times$ objective, an AxioCam, and AxioVision software (Carl Zeiss MicroImaging, Jena, Germany).

- The third dataset consisted of mouse-derived L929 cells stained with fluorescein diacetate (FDA; excitation wavelength 488 nm). This dataset contained 1,333 cells. The data were acquired using an Olympus IX71 inverted microscope (Olympus Germany, Hamburg, Germany) equipped with a 20$\times$ air objective, and the images were captured and stored with the AnalySIS$^{\text{TM}}$ software package.

To obtain the necessary ground-truth data for training and evaluation of the proposed adaptive segmentation scheme, cells were manually annotated by an experienced user. It was extremely important that only those cells were annotated that could clearly be distinguished and outlined by the operator and were fully visible in the field of view. To allow assessment of intraobserver and interobserver variability, 10% of all three datasets were randomly selected in order to reduce the amount of data. Each subset was then annotated a second time by a second user.

To provide optimal ground-truth segmentation data, a Wacom Cintiq 21UX$^{\text{TM}}$ digital drawing board (Wacom Europe, Krefeld, Germany) was used for the annotation process. This choice of input device for manual ground-truth annotation was based on a previous internal study in which the precision of interactive segmentation devices such as a Wacom board was compared with a conventional mouse and touchscreen device.

All three datasets will be made publicly available on the publication of the present study and can be obtained from the authors for comparative studies.

Methods Overview

Based on the observations by Srinivasa (4) and Wang et al. (5), several variations of watershed segmentation were used in the present study and were automatically adapted to the segmentation of fluorescence micrographs of different cell types. The first method is known as maximum-intensity linking (39) (see below); it is a graph-based variant of the watershed approach. The second method is an extension of MIL using an improved image preprocessing chain that is capable of handling the dotted structure of phalloidin Alexa 568 staining for F-actin in particular. In addition, the proposed preprocessing chain combines flexibility and a low dimensional parameter space (see the section on improved MIL below). The third and fourth algorithms investigated use different versions of an efficient preprocessing scheme that consists of a noise reduction step, mathematical morphology, and a threshold operation and is capable of separating touching or overlapping cells using the watershed transform (see below).

Maximum-Intensity Linking (MIL)

The MIL algorithm (39), which was originally developed for segmentation and separation of stem cells, exploits the fact that the intensity of fluorescence-stained cells usually decreases from the core to the boundary. The method can be subdivided into three steps. In order to remove tiny artifacts and noisy background pixels, a preprocessing step smoothes the image using a Gaussian filter kernel. Background and foreground pixels are separated by a global threshold obtained using Otsu's threshold approach. In a second step, the individual cells are segmented and separated by interpreting the image as a directed graph structure, in which pixels represent nodes linked to the brightest pixel in their immediate (8-connected) neighborhood. This results in a set of trees whose roots correspond to local intensity maximums in the image, in which each local image maximum relates to a tree. A color-coded example of this type of tree structure is shown in Figure 3B. Segmentation of cells can be obtained by adding corresponding successors to the source node associated with a specific cell. In the third step, cell regions are merged to reduce oversegmentation. For a more detailed description of the MIL method and the merging step, see Elter et al. (39). As can be seen in Figure 3C, some touching cells can be separated using this approach, but some oversegmentation and undersegmentation artifacts still remain.

Improved Maximum-Intensity Linking

When applied to the segmentation of HeLa cells with phalloidin Alexa 568 staining as used in the present study, there are two major drawbacks with MIL. The first is caused by intracellular staining, in which the dotted structure of the cytoplasm leads to strong oversegmentation artifacts, as each local maximum in the graph structure is segmented as a new cell. This effect can partially be fixed by merging regions in the postprocessing step. To reduce this type of oversegmentation, local maximums can be removed by an additional strong Gaussian smoothing in the preprocessing step. However, this smoothing also involves a loss of segmentation accuracy, as the boundaries between adjacent cells are also blurred. The second drawback is caused by the absence of dye in the nucleus, which leads to decreased intensities in comparison with the intensity of the surrounding cell (Fig. 2A). To reduce these problems, a chain of preprocessing steps was developed.

The dotted structure in the micrographs resulting from the phalloidin Alexa 568 marker can be reduced by carrying out morphological opening operations (45) with a flat circular structuring element of radius r. To separate individual cells from the image background, k-means clustering (46) is applied, where k denotes the number of clusters. Instead of performing clustering on the image intensity value, the histogram is used to accelerate this algorithm. After the clustering has been performed, a threshold limit is set by regarding the darkest cluster as background and everything else as foreground. In comparison with competing thresholding methods, this provides a very flexible method, as the number of clusters may vary while a small parameter space is preserved (usually $2 \leq k < 10$). As the MIL builds a tree for each local maximum, the procedure works best on cells with intensity values that decrease from the core to the boundary. A distance transform is therefore used on the binary image obtained from clustering, yielding an input image for the MIL. The intermediate steps in the improved MIL are illustrated in Figure 4.

Improved Watershed and Watershed by Reconstruction

The third algorithm in the present study uses the watershed transform (38), which is more widely established in the literature than the MIL described above for segmentation of cells. The previously described combination of smoothing, mathematical morphology (45) and k-means clustering (46) is used for preprocessing and thresholding. An improved watershed algorithm is then applied for splitting of the cells. The watershed algorithm interprets the gradient strength as relief in the image. Successive flooding of the basins is then performed on this relief. During this flooding process, watersheds are formed between adjacent catchment basins. In addition, knowledge about the typical size of the displayed cells is incorporated by merging adjacent catchment basins. This makes it possible to define a minimum size for each cell. This improved watershed method is thus capable of reducing oversegmentation artifacts.
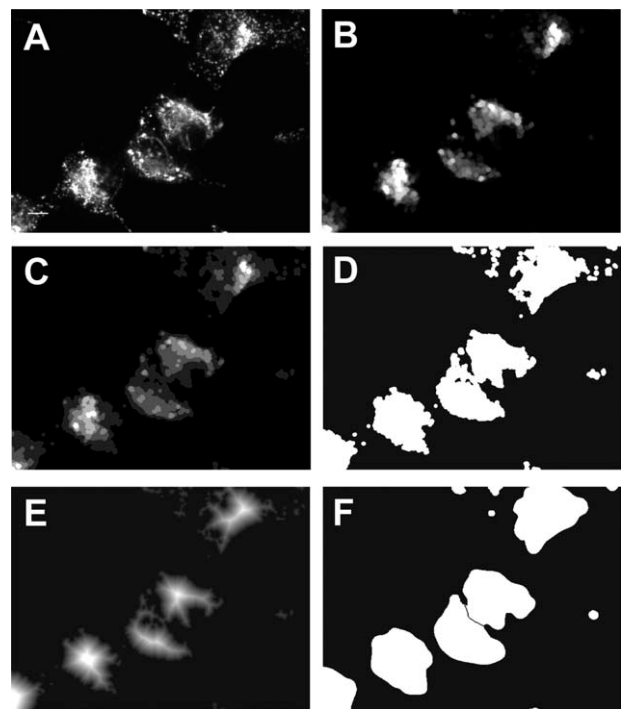

Figure 4. Example of the workflow with the improved maximum-intensity linking (MIL) method. (**A**) The original image, showing HeLa cells stained with phalloidin Alexa 568. (**B**) The image after morphological opening with a circular flat structuring element. (**C**) The method of k-means clustering was used to assign the pixels into k clusters ($k = 5$). (**D**) The darkest cluster was assumed to represent the background, using a threshold of $\tau = 1$. (**E**) A distance transform is applied to transform the image into a more suitable representation for the MIL approach. (**F**) The final segmentation result achieved with the improved MIL method is shown. The scale bar corresponds to 10 μm (original magnification $\times$63).

In principle, arbitrary images can be used as input for the watershed algorithm. In the literature, the gradient image is used, as well as other preprocessed variants of the fluorescence image. The present implementation uses three different input images:

- The fluorescence image, which is blurred with a Gaussian filter using standard deviation σ and enhanced by a morphological opening with radius r.
- A gradient-filtered version of the fluorescence image. To estimate the gradient, the differential of a Gaussian filter with standard deviation σ is used.
- A distance-transformed fluorescence image that has been preprocessed with morphological opening and binarized by k-means clustering.

In the current implementation, determination of the best input image for the watershed is part of the optimization process. This implies usage of a further parameter, denoted as m.

The watershed-by-reconstruction procedure is implemented in the same way as the improved watershed algorithm.

Table 2. Free parameters and range of the different segmentation algorithms that are automatically optimized by the genetic algorithm

METHOD	PARAMETER	RANGE
MIL	α	$[1, 2, \ldots, 20]$
	τ	$[1, 2, \ldots, 255]$
Improved MIL	k	$[2, 3, \ldots, 10]$
	r	$[5, 6, \ldots, 20]$
	α	$[1, 2, \ldots, 20]$
Improved watershed	r	$[5, 6, \ldots, 20]$
	σ	$[1, 2, \ldots, 20]$
	Conditioning method	$[0, 1, 2]$
Watershed by reconstruction	r	$[5, 6, \ldots, 20]$
	σ	$[1, 2, \ldots, 20]$
	Conditioning method	$[0, 1, 2]$

MIL, maximum-intensity linking.

In this case, the morphological opening is replaced by a morphological opening by reconstruction (45) in the preprocessing chain as well as for the watershed input images.

Parameter Optimization and Separation of Data

The segmentation performance of the methods described depends on way in which the free parameters are selected. To avoid a biased set-up of the different segmentation algorithms, the parameters for all of the methods are automatically optimized using a genetic algorithm (GA) (47). Threefold cross-validation was used to separate training from testing data. For this cross-validation, each image database was randomly split into three equal-sized image sets in which two-thirds of the images were used to optimize the parameters and the remaining third was used for testing. This was done with all three possible combinations of the training and testing data.

For the present study, the free steady-state genetic algorithm implementation described by Wall, developed at the Massachusetts Institute of Technology, was used (48). From a genetic algorithm point of view, a parameter set is represented by an individual, whereas a specific parameter can be interpreted as an allele. When mutation and cross-over operations are performed (47), alleles are changed and a new individual is formed. A certain number of individuals (10 in the present study) were combined to form a new generation of individuals. A steady-state genetic algorithm was used, as it is able to preserve a certain percentage (20% in the present study) of the best individuals contained in the previous generation. Twenty generations of individuals were computed for the present evaluation. The parameters required for optimization are summarized in Table 2 for all of the segmentation methods evaluated.

Performance Measurement

In the proposed optimization scheme, a performance measurement has to assess the quality achieved by a specific segmentation method in relation to a manually annotated ground truth. An extended overlap performance measurement is therefore used (49) that describes the three major aspects of segmentation—namely, the amount of overlap as well as the amounts of oversegmentation and undersegmentation.

The area overlap measure (AOM; also known as the Jaccard similarity measure), measuring the ratio of the intersection area of S and T and the joint area of S and T:

$$P_1 = \frac{|S \cap T|}{|S \cup T|}.$$

The ratio of the undersegmented area to the ground-truth area T:

$$P_2 = \frac{|T \backslash (S \cap T)|}{|T|}.$$

The ratio of the ground-truth and the segmented area S, defined as oversegmentation:

$$P_3 = \frac{|S \backslash (S \cap T)|}{|S|}.$$

Depending on the way in which they are defined, the three performance measurements yield values between 0 and 1; hence, P_1, P_2, and $P_3 \in [0, 1]$. A combined performance measurement $P \in [0,1]$ can therefore be defined by a linear combination of these terms. By assigning different weighting factors to the individual terms, it is possible to control the influence of each individual term P_1, P_2, and P_3. This makes it possible to reduce oversegmentation, for example, by assigning a larger weighting factor to P_3. As neither oversegmentation nor undersegmentation was preferred for the present application, equal weighting factors were assigned to each of the three criteria. The combined performance measurement P was therefore defined as:

$$P = \frac{P_1 + (1 - P_2) + (1 - P_3)}{3}.$$

A qualitative segmentation evaluation can be performed using P. The numbers of erroneously detected cells and missed cells are consequently irrelevant. In order to judge the segmentation quantitatively as well as qualitatively, the hit rate is included in the measurement. Assuming that the segmentation results contain n cells and the ground truth contains m cells, $n{:}m$ mapping has to be carried out in the following way. First, the best-matching cell — i.e., the cell with the largest P value — is searched for among the segmented cells for each ground-truth cell. The performance of such an optimal pair of cells is denoted as P_i^{best} with $i \in \{1,2,\ldots,m\}$. Each segmented cell may only be assigned to one ground-truth cell. The number of correctly identified cells ($P_i^{\text{best}} > 0$) is then denoted as N_{TP} ("true positives"). N_{FP} denotes the number of oversegmented

Segmentation Performance P for the different datasets

	L929 (FDA)	HeLa (Phalloidin 568)	Hela (DiD)
☐ MIL	0.7605	0.5169	0.6475
■ Improved MIL	0.7658	0.5456	0.6731
⊠ Improved Watershed	0.836	0.6063	0.7105
⊟ Improved Watershed R	0.8457	0.6428	0.6950
⊠ Inter-Observer	0.8906	0.7877	0.918251
☐ Intra-Observer	0.8785	0.8451	0.889456

A

Accuracy for the different datasets

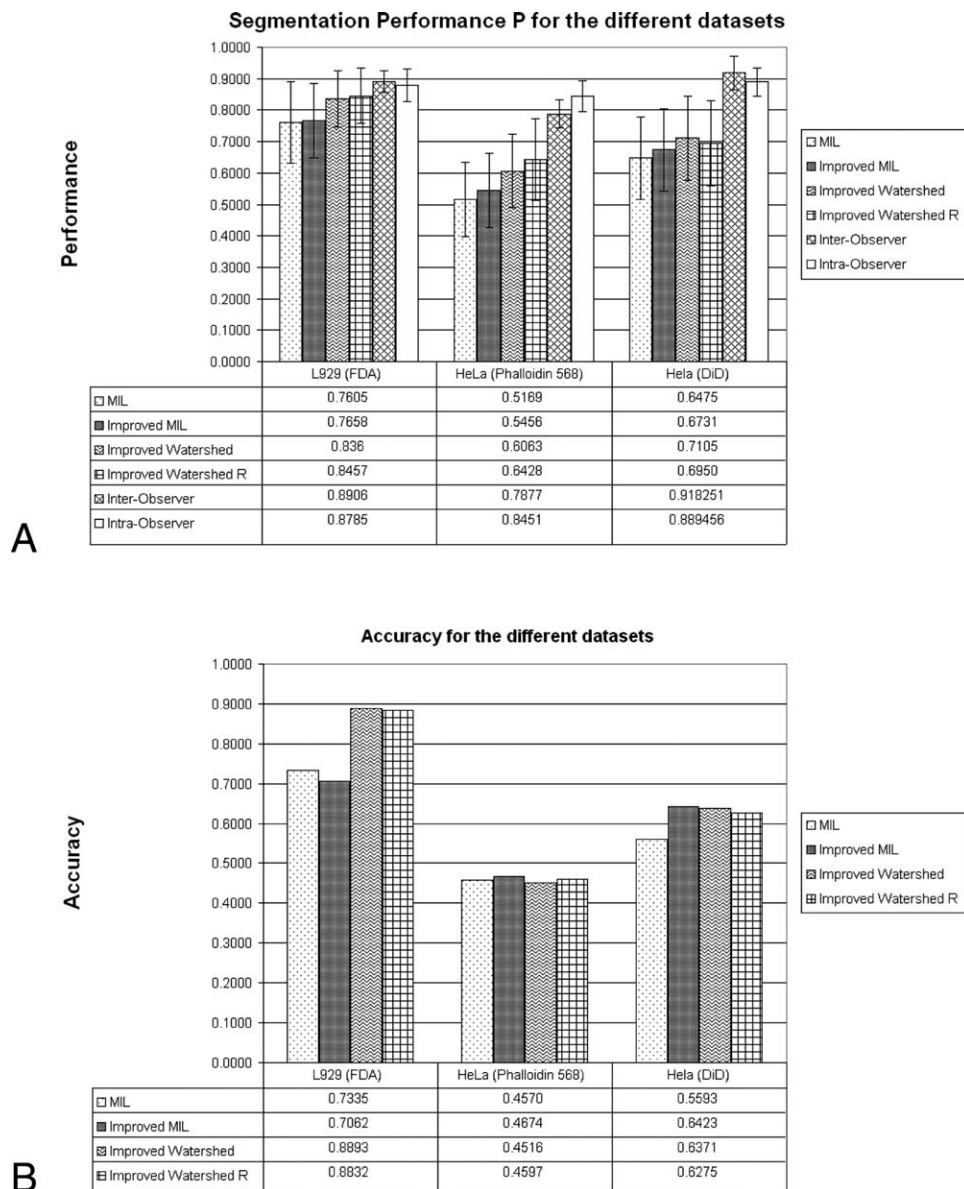|	L929 (FDA)	HeLa (Phalloidin 568)	Hela (DiD)
☐ MIL	0.7335	0.4570	0.5593
■ Improved MIL	0.7062	0.4674	0.6423
⊠ Improved Watershed	0.8893	0.4516	0.6371
⊟ Improved Watershed R	0.8832	0.4597	0.6275

B

Figure 5. Comparison of the segmentation performance using the segmentation methods described and threefold cross-validation for the different datasets. (**A**) The combined segmentation performance P. (**B**) The hit rate. It should be noted that only 10% of the data were used to calculate interobserver and intraobserver variability.

cells ("false positives"). These are cells that have been wrongly detected, as they are not contained in the ground-truth annotations. The number of cells that were not found at all is denoted as N_{FN}. Defining the accuracy as $A = \frac{N_{TP}}{N_{TP} + N_{FP} + N_{FN}}$ leads to the following performance measurement:

$$P_{Punished} = \frac{1}{N_{TP} + N_{FP} + N_{FN}} \sum_{i=1}^{m} P_i^{best}.$$

Maximization of $P_{Punished}$ is equivalent to maximizing the overlapping of the segmentation result while minimizing the numbers of oversegmented and missed cells.

Statistical Analysis

The segmentation methods were analyzed using mixed linear models with the segmentation performance measure P (defined above) as outcome, segmentation method as fixed effect, and ground-truth cells as random effect. These models take into account the fact that each ground-truth cell is used by several segmentation methods. For each dataset a separate model was performed. The segmentation methods were compared with post hoc tests using the Tukey–Kramer method.

All of the tests were two-sided, and a p value <0.05 was considered statistically significant. All of the statistical analyses were carried out using SAS (version 9.2; SAS Institute Inc., Cary, NC).
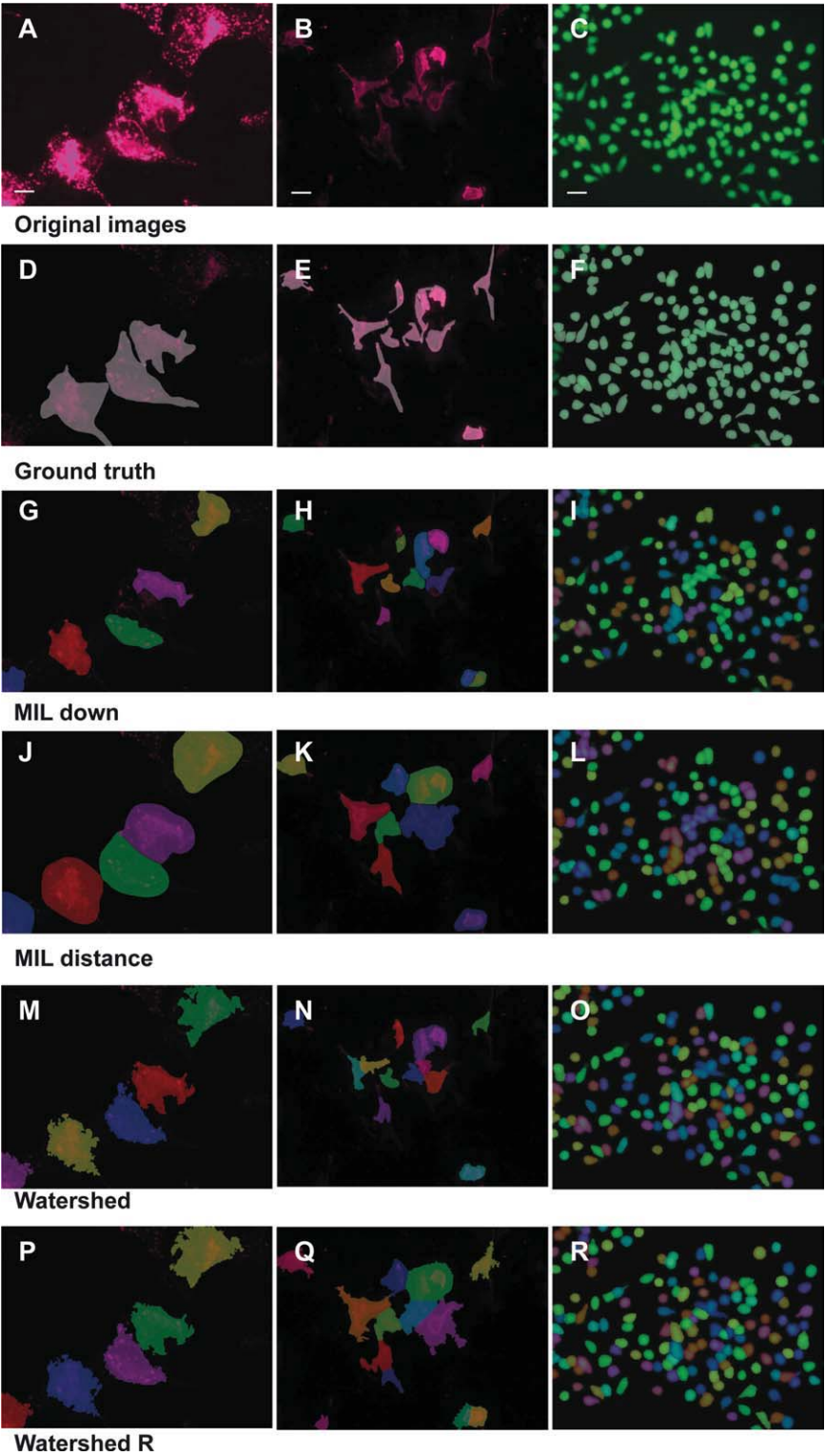
Figure 6. Direct comparison of various segmentation approaches; representative examples of three data sets.

RESULTS

For fair comparison of the segmentation methods presented, their optimal parameters are determined using the genetic algorithm described above. In order to evaluate the qualitative and quantitative segmentation performance, the combined segmentation performance P and the accuracy A

Table 3. Comparison of segmentation methods for the different datasets (mixed linear models)

	MIL/IMPROVED MIL	IMPROVED MIL/WATERSHED	WATERSHED/ WATERSHED WITH RECONSTRUCTION
HeLa (DiD)	<0.0001	<0.0001	0.85
HeLa (phalloidin 568)	<0.0001	<0.0001	0.60
L929 (FDA)	0.64	<0.0001	0.90

p values for Tukey-Kramer post hoc tests are shown. MIL, maximum-intensity linking.

Table 5. Runtime comparison of nonparallelized implementations of the methods described

	MIL	IMPROVED MIL	IMPROVED WATERSHED	WATERSHED BY RECONSTRUCTION
Runtime (s)	1.3	1.5	1.4	1.5

The time required for segmentation of 10 selected images (with 1,376 × 1,032 pixels) was measured and averaged on an Intel Core 2 Duo, 2.66 GHz.

were analyzed. The results show that the segmentation performance of the MIL method can be improved using the proposed preprocessing routines. Further improvement is obtained by using the watershed transform. The watershed-by-reconstruction method can improve performance for the phalloidin 568-stained HeLa cells, while performance deteriorates when segmenting DiD-stained HeLa cells (Fig. 5). A visual comparison of the segmentation results (Fig. 6) confirms that the segmentation performance of the watershed transform is superior to that of the MIL algorithm.

Table 3 shows *p* values for comparisons of the segmentation methods. Testing whether the performance of the MIL method is significantly different from the improved MIL methods yields $p < 0.0001$ on any dataset except for the L929 cells ($p = 0.64$). The tests also show that the watershed method's segmentation performance is different from the performance of the improved MIL method on all datasets ($p < 0.0001$), whereas comparison of watershed and watershed-by-reconstruction does not show any significant differences ($p > 0.05$).

Analyzing the watershed method in more detail, we addressed the question of the optimal input image for the watershed transform. An additional parameter was therefore incorporated that determines whether the preprocessed fluorescence image, the gradient image, or the distance-transformed image is used as input for the watershed algorithm. This parameter was also evaluated for each of the three different combinations of training and testing dataset. The values listed in Table 4 confirm that the best input image depends on

Table 4. Preprocessing steps for the watershed transform achieving the best segmentation performance after parameter optimization for each of the specific methods

	L929 (FDA)	HELA (DID)	HELA (PHALLOIDIN)
Improved watershed	FL	Distance	Distance
Watershed by reconstruction	FL	Distance	FL

''FL'' indicates that the watershed transform is directly performed on the original fluorescence image, whereas ''distance'' shows that the distance-transformed image is used as the optimal input image.

the dataset used, but the gradient image did not outperform competing methods on any dataset.

All of the algorithms presented were developed in relation to performance and efficiency. A runtime comparison (Table 5) showed that all algorithms can segment a 1.3 megapixel image in less than 2 s using a 2.66-GHz processor.

DISCUSSION

In this study, several watershed-based image segmentation methods (MIL, extended MIL, improved watershed, and watershed by reconstruction) were evaluated with regard to their usability in an adaptive segmentation framework (3) for fluorescence-stained cells. A key issue in the selection of these methods was their applicability for the segmentation of different types of fluorescence-stained cells. In addition, the parameter space had to be kept small while maintaining sufficient flexibility to adapt to various cell types and stains, making the methods described suitable for the automated parameter optimization process. As a result of the automatic parameter optimization, a runtime of less than 2 s is required for a typical image with a size of 1.3 megapixels.

These results and the corresponding statistical tests clearly demonstrate that the performance of the MIL method can be improved using the proposed flexible preprocessing routine for most datasets. Using the watershed transform-based segmentation routine can significantly improve performance for all datasets. Incorporating morphological reconstruction operators also improves performance on most datasets, but these differences turned out to be not significant.

Analysis of the optimal preprocessed input image for the watershed method (Table 4) indicates that this parameter depends on the present dataset as well as on the algorithm used. It was notable that the gradient image did not outperform the competing images on any dataset. These observations are consistent with the findings reported by other groups (1,5,17,18,36) using varying input images for the watershed transform.

The results show that all evaluated segmentation methods can be optimized for the application on individual data sets using a genetic algorithm, hence increasing the performance measurements. Nevertheless, it can also be seen, that with rising complexity of the image data (increasing number of touching and overlapping cells, variations and quality of staining, number of cells), a limit is reached in the methodology. Analysis of Figure 6 shows that many errors occur in very

complex scenarios that are challenging even for experienced biologists.

Conclusions

This study outlines a framework for the segmentation of varying cell types based on variations of the watershed transform, combined with an efficient preprocessing routine and automatic parameter optimization using a genetic algorithm. The analyses show that the segmentation schemes evaluated can be adapted effectively to different stains and cell types. Following an automatic adaptation step, an ideal combination of preprocessing methods and the watershed transform can thus be applied robustly to micrographs with the same preparation and cell stains. However, if a high degree of accuracy is required, some interactive correction steps are needed for more complex scenarios, due to an inability to carry out segmentation of overlapping cells. Hence, in order to increase the performance further under such complex side conditions, model-based segmentation routines are needed, which incorporate prior knowledge about the size, form and appearance of the cells to be segmented.

Acknowledgments

Literature Cited

1. Allalou A, Wählby C. BlobFinder, a tool for fluorescence microscopy image cytometry. Comput Methods Programs Biomed 2009;94:58–65.

2. Vidal F, Bello F, Brodlies KW, John NW, Gould D, Phillips R, Avis NJ. Principles and applications of computer graphics in medicine. Comp Graphics Forum 2006;25:113–137.

3. Wittenberg T, Becher F, Hensel M, Steckhan D. Image segmentation of cell nuclei based on classification in the color space. In: Vander Sloten J, Verdonck P, Nyssen M, Haueisen J, editors.4th European Conference of the International Federation for Medical and Biological Engineering (ECIFMBE), Vol22/1; November 23–27, 2008; Antwerp, Belgium. Berlin: Springer; 2008. pp 613–614.

4. Srinivasa G. Active Mask Framework for Segmentation of Fluorescence Microscope Images. Pittsburgh: Carnegie-Mellon University; 2008.

5. Wang Q, Niemi J, Tan CM, You L, West M. Image segmentation and dynamic lineage analysis in single-cell fluorescence microscopy. Cytometry A J Int Soc Anal Cytol 2010;77:101–110.

6. Romanelli M, Buitelaar P, Sintek M. Modeling linguistic facets of multimedia content for semantic annotation. In: Falcidieno B, Spagnuolo M, Avrithis Y, Kompatsiaris I, Buitelaar P, editors. Semantic Multimedia: Second International Conference on Semantic and Digital Media Technologies (SAMT); December 5–7, 2007; Genoa, Italy. Berlin: Springer; 2007. pp 240–251.

7. Garbay C. Computer vision: A plea for a constructivist view. In: Combi C, Shahar Y, Abu-Hanna A, editors. Artificial Intelligence in Medicine: 12th Conference on Artificial Intelligence in Medicine (AIME); July 18–22, 2009; Verona, Italy. Berlin: Springer; 2009. pp 6–15.

8. Restif C.Segmentation and evaluation of fluorescence microscopy images [MD Dissertation]. Oxford: Oxford Brookes University; 2006. Available at: http://cms.brookes.ac.uk/staff/PhilipTorr/Papers/thesis%20Restif.pdf ; [accessed Oct 2011].

9. Harder N, Mora-Bermúdez F, Godinez WJ, Ellenberg J, Eils R, Rohr K. Automated analysis of the mitotic phases of human cells in 3D fluorescence microscopy image sequences. Med Image Comput Comput Assist Interv 2006;9:840–848.

10. Harder N, Mora-Bermúdez F, Godinez WJ, Ellenberg J, Eils R, Rohr K. Determination of mitotic delays in 3D fluorescence microscopy images of human cells using an error-correcting finite state machine. In: Horsch A, Deserno TM, Handels H, Meinzer HP, Tolxdorff T, editors. Bildverarbeitung für die Medizin 2007: Algorithmen—Systeme—Anwendungen. Proceedings des Workshops vom 25–27. März 2007 in München. Berlin: Springer; 2007. pp 242–246.

11. Venkatraman R, Raman R, Raman B, Moss RB, Rubin GD, Mathers LH, Robinson TE. Fully automated system for three-dimensional bronchial morphology analysis using volumetric multidetector computed tomography of the chest. J Digit Imaging 2006;19:132–139.

12. Wittenberg T, Grobe M, Münzenmayer C, Kuziela H, Spinnler K. A semantic approach to segmentation of overlapping objects. Methods Inf Med 2004;43:343–353.

13. Otsu N. A threshold selection method from gray-level histograms. IEEE Trans Syst Man Cybern 1979;9:62–66.

14. Nandy K, Gudla PR, Meaburn KJ, Misteli T, Lockett SJ. Automatic nuclei segmentation and spatial FISH analysis for cancer detection. Conf Proc IEEE Eng Med Biol Soc 2009;2009:6718–6721.

15. Du X, Dua S. Segmentation of fluorescence microscopy cell images using unsupervised mining. Open Med Inform J 2010;4:41–49.

16. Malpica N, de Solórzano CO, Vaquero JJ, Santos A, Vallcorba I, García-Sagredo JM, del Pozo F. Applying watershed algorithms to the segmentation of clustered nuclei. Cytometry 1997;28:289–297.

17. Wählby C, Lindblad J, Vondrus M, Begtsson E, Björkesten L. Algorithms for cytoplasm segmentation of fluorescence labelled cells. Anal Cell Pathol 2002;24:101–111.

18. Lindblad J, Wählby C, Bengtsson E, Zaltsman A Image analysis for automatic segmentation of cytoplasms and classification of Rac1 activation. Cytometry A J Int Soc Anal Cytol 2004;57:22–33.

19. Zhu D, Jarmin S, Ribeiro A, Prin F, Xie SQ, Sullivan K, Briscoe J, Gould AP, Marelli-Berg FM, Gu Y. Applying an adaptive watershed to the tissue cell quantification during T-cell migration and embryonic development. Methods Mol Biol 2010;616:207–228.

20. Soille P. Morphological Image Analysis: Principles and Applications. Berlin: Springer; 1999.

21. Metzler V, Bienert H, Lehmann T, Mottaghy K, Spitzer K. A novel method for quantifying shape deformation applied to biocompatibility testing. ASAIO J 1999;45:264–271.

22. Metzler V, Lehmann T, Bienert H, Mottaghy K, Spitzer K. Scale-independent shape analysis for quantitative cytology using mathematical morphology. Comput Biol Med 2000;30:135–151.

23. Zhang G, Jia X, Pham T, Crane D. Multistage spatial property based segmentation for quantification of fluorescence distribution in cells. In: Pham T, Zhou X, editors. 2009 International Symposium on Computational Models for Life Sciences (CMLS); July 28–29, 2009; Sofia, Bulgaria. Berlin: Springer; 2010. pp 3–12. (American Institute of Physics Conference Proceedings, Vol.1210.)

24. Srinivasa G, Fickus M, Kovacevic J. Active contour-based multiresolution transforms for the segmentation of fluorescence microscope images. In: Van De Ville D, Goyal VK, Papadakis M, editors. Wavelets XII; August 26–29, 2007; San Diego, California. Bellingham, WA: Society of Photo-optical Instrumentation Engineers (SPIE); 2007. pp 254–261. (Proceedings of SPIE—The International Society for Optical Engineering, Vol.6701.)

25. Srinivasa G, Fickus MC, Guo Y, Linstedt AD, Kovacević J. Active mask segmentation of fluorescence microscope images. IEEE Trans Image Process 2009;18:1817–1829.

26. Möller B, Gress O, Stöhr N, Hüttelmaier S, Posch S. Adaptive segmentation of cells and particles in fluorescent microscope images. In: Richard P, Braz J, editors.Proceedings of the International Conference on Computer Vision Theory and Applications (VISAPP), Vol.2; May 17–21, 2010; Angers, France. Berlin: Springer; 2010. pp 97–106.

27. Ersoy I, Bunyak F, Chagin V, Cardoso MC, Palaniappan K. Segmentation and classification of cell cycle phases in fluorescence imaging. Lect Notes Comput Sci 2009;5762:617–624.

28. Yu W, Lee H, Hariharan S, Bu W, Ahmed S. Quantitative neurite outgrowth measurement based on image segmentation with topological dependence. Cytometry A J Int Soc Anal Cytol 2009;75:289–297.

29. Yu W, Lee H, Hariharan S, Bu W, Ahmed S. Evolving generalized Voronoi diagrams for accurate cellular image segmentation. Cytometry A J Int Soc Anal Cytol 2010;77:379–386.

30. Cheng J, Rajapakse JC. Segmentation of clustered nuclei with shape markers and marking function. IEEE Trans Biomed Eng 2009;56:741–748.

31. Dzyubachyk O, Essers J, van Cappellen WA, Baldeyron C, Inagaki A, Niessen WJ, Meijering E. Automated analysis of time-lapse fluorescence microscopy images: From live cell images to intracellular foci. Bioinformatics 2010;26:2424–2430.

32. Dzyubachyk O, van Cappellen WA. Essers J. Niessen WJ, Meijering E. Advanced level-set-based cell tracking in time-lapse fluorescence microscopy. IEEE Trans Med Imaging 2010;29:852–867. Erratum in: IEEE Trans Med Imaging 2010;29:1331.

33. Nattkemper TW.A neural network-based system for high throughput fluorescence micrograph evaluation [MD dissertation]. Bielefeld, Germany: University of Bielefeld, Faculty of Technology; 2001.

34. Nattkemper TW, Twellmann T, Ritter H, Schubert W. Human vs machine: Evaluation of fluorescence micrographs. Comput Biol Med 2003;33:31–43.

35. Nattkemper TW, Wersing H, Schubert W, Ritter H. A neural network architecture for automatic segmentation of fluorescence micrographs. In: Verleysen M, editor. 8th European Symposium on Artificial Neural Networks; April 26–28, 2000; Bruges. Brussels: D Facto Publications; 2000. pp 177–182.

36. Wang M, Zhou X, Li F, Huckins J, King RW, Wong ST. Novel cell segmentation and online SVM for cell cycle phase identification in automated microscopy. Bioinformatics 2008;24:94–101.

37. Coelho LP, Shariff A, Murphy RF. Nuclear segmentation in microscope cell images: A hand-segmented dataset and comparison of algorithms. Proc IEEE Int Symp Biomed Imaging 2009;5193098:518–521.

38. Roerdink J, Meijster A. The watershed transform: Definitions, algorithms and parallelization strategies. Fundam Inf 2001;41:187–228.

39. Elter M, Daum V, Wittenberg T. Maximum-intensity-linking for segmentation of fluorescence-stained cells. In: Metaxas DN, Whitaker RT, Rittscher J, Sebastian TB, editors. Proceedings of 1st Workshop on Microscopic Image Analysis with Applications in Biology (in conjunction with MICCAI); October 5, 2006; Copenhagen. Copenhagen: MIAAB; 2006. pp 46–49. Available at:http://www.miaab.org/miaab-2006-papers/miaab-2006-11-elter.pdf. [accessed Oct 2011].

40. Habeler G, Natter K, Thallinger GG, Crawford ME, Kohlwein SD, Trajanoski Z. YPL.db: The yeast protein localization database. Nucleic Acids Res 2002;30:80–83.

Segmentation in Fluorescence Micrographs

41. Kals M, Natter K, Thallinger GG, Trajanoski Z, Kohlwein SD. YPL.db2: The yeast protein localization database, version 2.0. Yeast 2005;22:213–218.

42. Riffle M, Davis TN. The yeast resource center public image repository: A large database of fluorescence microscopy images. BMC Bioinformatics 2010;11:263.

43. Singh AK, Manjunath BS, Murphy RF. A distributed database for biomolecular images. SIGMOD Rec 2004;33:65–71. Available at: http://sigmod.org/sigmod/record/issues/0406/10.AmbujSingh.pdf ; [accessed Oct 2011].

44. Frenkel-Morgenstern M, Cohen AA, Geva-Zatorsky N, Eden E, Prilusky J, Issaeva I, Sigal A, Cohen-Saidon C, Liron Y, Cohen L, Danon T, Perzov N, Alon U. Dynamic proteomics: A database for dynamics and localizations of endogenous fluorescently-tagged proteins in living human cells. Nucleic Acids Res 2010;38:D508–D512.

45. Serra J. Image Analysis and Mathematical Morphology. London: Academic Press; 1982.

46. Hartigan J, Wong M. A k-means clustering algorithm. Appl Stat 1979;28:100–108.

47. Goldberg DE. Genetic Algorithms in Search, Optimization, and Machine Learning. Reading, MA: Addison-Wesley; 1989.

48. Wall M. A C++ Library of Genetic Algorithm Components. Cambridge, MA: Massachusetts Institute of Technology; 1996. Available at:http://lancet.mit.edu/ga/(version 2.4.7). [accessed Oct 2011].

49. Rojas Dominguez A, Nandi AK. Improved dynamic-programming-based algorithms for segmentation of masses in mammograms. Med Phys 2007;34:4256–4269.