Self-learning Segmentation and Classification of Cell-Nuclei in 3D Volumetric Data Using Voxel-Wise Gray Scale Invariants

Janis Fehr[1], Olaf Ronneberger[1], Haymo Kurz[2], and Hans Burkhardt[1]

[1] Albert-Ludwigs-Universität Freiburg, Institut für Informatik,
Lehrstuhl für Mustererkennung und Bildverarbeitung,
Georges-Koehler-Allee Geb. 052, 79110 Freiburg, Deutschland
`fehr@informatik.uni-freiburg.de`
`http://lmb.informatik.uni-freiburg.de/`
[2] Albert-Ludwigs-Universität Freiburg, Institut für Anatomie und Zell Biologie,
79104 Feiburg i.Br., Deutschland

Abstract. We introduce and discuss a new method for segmentation and classification of cells from 3D tissue probes. The anisotropic 3D volumetric data of fluorescent marked cell nuclei is recorded by a confocal laser scanning microscope (LSM). Voxel-wise gray scale features (see accompaning paper [1][2]), invariant towards 3D rotation of its neighborhood, are extracted from the original data by integrating over the 3D rotation group with non-linear kernels.

In an interactive process, support-vector machine models are trained for each cell type using user relevance feedback. With this reference database at hand, segmentation and classification can be achieved in one step, simply by classifying each voxel and performing a connected component labelling, automatically without further human interaction. This general approach easily allows adoption of other cell types or tissue structures just by adding new training samples and re-training the model. Experiments with datasets from chicken chorioallantoic membrane show encouraging results.

1 Introduction

In biological and medical research as well as in histopathologic diagnosis, the localization and classification of cells is an everyday business. A vast number of research techniques and treatment methods require detailed information on the amount, type, localization and state of cells in a given probe of tissue or dilution. Locating, classifying and analyzing cells is not a simple task, very time consuming, and in most cases a human expert is needed. The demand for automation is continuously growing with the large number of applications in biotechnology and medical research. But so far this problem is not satisfyingly solved in general. Although there are various methods around, which perform quite well for simple tasks like counting or the segmentation of cells in dilution, most problems are still subject to basic research.

Concerning the algorithms introduced in the literature so far, most of them suffer from the fact that they have been developed for one special purpose only and cannot be easily generalized to other cell types or tissues. Besides that, segmentation is often an unsolved problem as well, and many algorithms require manual interaction. We introduce a new general purpose algorithm using voxel-wise gray scale invariants([1][2]) for both, segmentation and classification of cells in 2D and 3D probes, and provide some first, promising experimental results. Motivated by the work of [3] and [4], gray scale invariants were very successfully applied to individual pollen recognition in [5] [6]. A major problem of cell classification in tissue probes is segmentation. In order to achieve good classification results, supervised-learning classifiers rely on proper segmented training samples and classification probes. Proper segmentation is hard to realize without higher semantic knowledge about the object to segment. But the use of a-priori knowledge or manual segmentation is not suited for a fully automatic general purpose approach.

For this reason we developed a self learning segmentation algorithm by use of gray scale invariants, which is capable of performing segmentation and classification in one step. Gray scale invariant features are extracted from the surrounding neighborhood of each pixel/voxel. In an interactive procedure, a support-vector machine model is trained. Once this model has been obtained for the requested types of cells, segmentation and classification can be performed automatically without any further human interaction.

This paper is structured as follows. Section 2 gives a brief introduction to voxel-wise gray scale invariants. In section 3 we introduce the actual segmentation using an interactive training method and support-vector machines. Finally, in section 4 we present some experimental results.

2   Voxel-Wise Gray Scale Invariants

Gray scale features, invariant towards Euclidean motion, using Haar-integration over the whole transformation group of an n-dimensional data set $\mathbf{X}$, are calculated as follows: [3] [5]

$$T[f](\mathbf{X}) := \int\limits_{G} f(g\mathbf{X})dg \tag{1}$$

where G denotes the transformation group, g one element of G, f a nonlinear kernel function and $g\mathbf{X}$ the transformed n-dimensional data set. If the kernel function f only depends on a few points of the image or volume, i.e., if we can rewrite $f(\mathbf{X})$ as $f\left(\mathbf{X}(\boldsymbol{x}_1), \mathbf{X}(\boldsymbol{x}_2), \mathbf{X}(\boldsymbol{x}_3), \ldots\right)$, where $\mathbf{X}(\boldsymbol{x}_i)$ is the gray value[1] at position $\boldsymbol{x}_i$ we only need to transform the kernel points $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \ldots$ accordingly, instead of the whole data set $\mathbf{X}$. This transformation of the kernel points is denoted as $s_g(\boldsymbol{x}_i)$, rewriting (1) as

---

[1] We use the term "gray value" even for color or other multi-channel data. In this case one "gray value" has multiple components.

$$T[f](\mathbf{X}) := \int_G f\left(\mathbf{X}(s_g(\boldsymbol{x}_1)),\ \mathbf{X}(s_g(\boldsymbol{x}_2)),\ \mathbf{X}(s_g(\boldsymbol{x}_3)),\ldots\right) dg\ . \qquad (2)$$

The direct evaluation of the integral (2) is usually too slow for real applications. [5] presented a fast calculation method (using FFTs) for a certain class of kernel-functions (so called separable two-point-kernel functions) of the form

$$f(\mathbf{X}) = f_a\left(\mathbf{X}(\mathbf{0})\right) \cdot f_b\left(\mathbf{X}(\boldsymbol{q})\right) \qquad \begin{array}{l} f_a, f_b : \text{any nonlinear functions that} \\ \qquad \text{transform the gray values} \\ \boldsymbol{q} : \text{span of the kernel function} \end{array} \qquad (3)$$
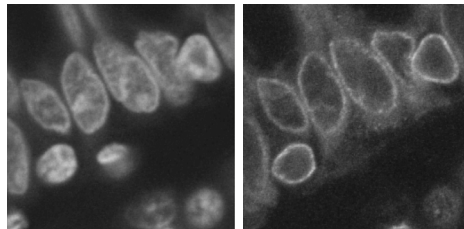
Calculation of Voxel-Wise Gray Scale Invariants: The voxel-wise extraction of invariant features follows the same theory as above, restricting the transformation group to rotation. A major drawback of voxel-wise calculation with two-point-kernel functions is that the resulting features are not only invariant towards rotation, but also towards arbitrary permutation of neighboring gray values. To overcome this problem, we introduced a fast approximation using FFT and 3D *separable three-point-kernel functions* [2] of the type (see accompaning paper [1])

$$f(\mathbf{X}) = f_a\left(\mathbf{X}(\mathbf{0})\right) \cdot f_b\left(\mathbf{X}(\boldsymbol{q}_1)\right) \cdot f_c\left(\mathbf{X}(\boldsymbol{q}_2)\right) \qquad (4)$$

Multichannel Features: As illustrated in Fig. 1, biological probes are often stained with different fluorescent markers, which are recorded as multi channel datasets. Kernels of the form (4) can be evaluated over several channels, using the voxel-wise gray-scale representation of the recorded volumetric datasets $\mathbf{X_v}$, where $\mathbf{X_{v_i}}$ gives the gray-value for the i-th channel.

Kernel Functions: To increase separability, several features with different spans and non-linear mappings are combined to feature vectors for each voxel.

In the case of a compact transformation group, like rotation, any kernel function returning a scalar value may be used, because after parameterization, an integral with fixed borders (e.g., integration from 0 to 360 degrees) can be

Fig. 1. By staining with different fluorescent markers and variation of the ecitation wave length, several data channels can be recorded from a single sample at once
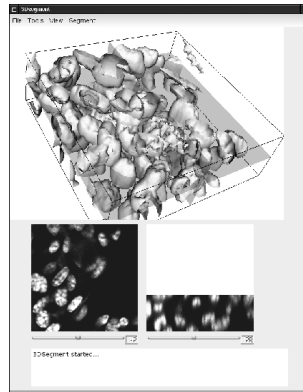
Fig. 2. From left to right: original data, classification of features without exponential kernels, classification with the same training samples but with some features calculated on the "inverse data"

found. For the later described segmentation we use simple non-linear functions like $f(\boldsymbol{v}) = v_i^2, v_i^3, v_i^4, \dots or \sqrt{v_i}$ with spans from 2,4,8 up to 32. While small spans extract local object features (high frequencies), it is useful to have some larger spans covering the entire object (low frequencies). In addition, we perform Gaussian filtering previous to the non-linear mappings for increased local support [1][2]. For datasets with high valued object gray-values and low background values an additional problem arises: due to the nature of Haar-integration, the foreground values dominate the result of the voxel-wise features which leads to a reduced separability of the background close to objects. A solution is provided by calculating some features which are sensitive to the background. This can be achieved by use of an appropriate kernel function like:

$$f(\mathbf{X}) = \sqrt{\mathbf{X(0)}} \cdot e^{-\mathbf{X(q_1)}^2} \quad \text{two-point exponential kernel function} \qquad (5)$$
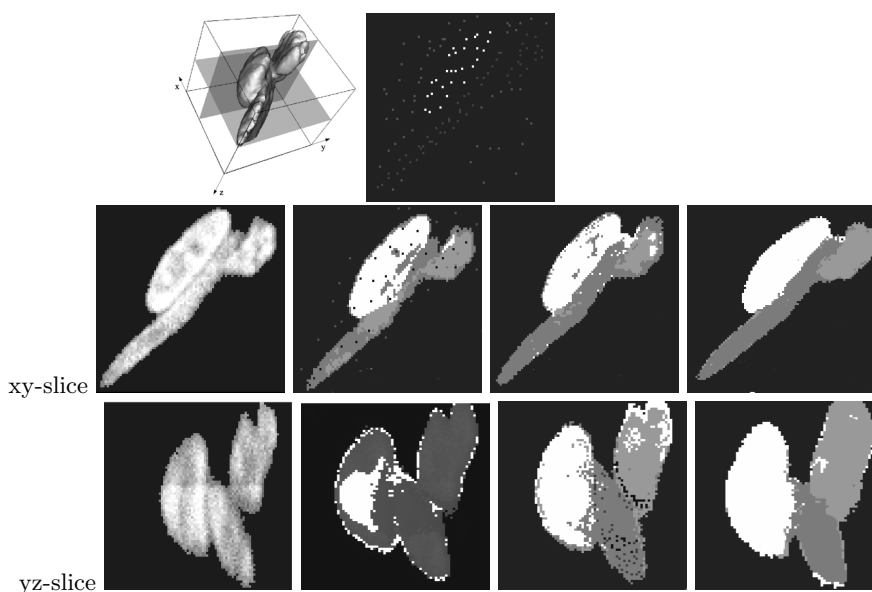
3   Segmentation

After the voxel-wise extraction of feature-vectors using two- and three-point kernels, a support-vector machine (SVM) [7][8] model is trained in an interactive

Fig. 3. Framework for interactive model training: xy-slices (bottom left) and yz-slices (bottom right) are moved through the volumetric dataset (top) and training samples are selected manually via "mouse clicks".

procedure over several iterations: First a small number of training samples (voxels) is manually selected for each class (Fig. 3). Second, a SVM model is trained based on the training feature-vectors. In the last step of one iteration, all voxels are classified against the previously trained model. After each iteration new training samples can be added in order to improve segmentation and classification results until the model reaches a "stable" state, e.g. the support-vectors do not change after adding new samples. In order to avoid overfitting and to find the

Fig. 4. The interactive training process - 1st row: 3D reconstruction of the original data, 311 training samples set for the first iteration of training. 2nd line - from left to right: section of xy-slice of original data as indicated in the 3D reconstruction, result after the first iteration (56 support vectors in model), result after 2nd iteration, result after the 3rd iteration (642 training samples, 129 support vectors in model). 3rd line: section of yz-slice of original data, results after 1st to 3rd iterations in yz-slice.

optimal SVM model, we perform a grid-search over SVM-kernel parameters and cost-function with cross-validation model selection in each training round. One of the major advantages of our approach is, that the obtained model can easily be extended with samples from other datasets and and even new classes, simply by executing additional training rounds. With the model at hand, objects in datasets which have been recorded under similar conditions (staining, excitation, etc.) can be segmented and classified fully automatically: voxel-wise features are extracted and classified. In a final step the labeled voxels are combined to closed objects by connected component labeling.

4   Experiments

In this section we present some experimental results and compare the performance of the previously described algorithms with a standard Watershed region growing approach.
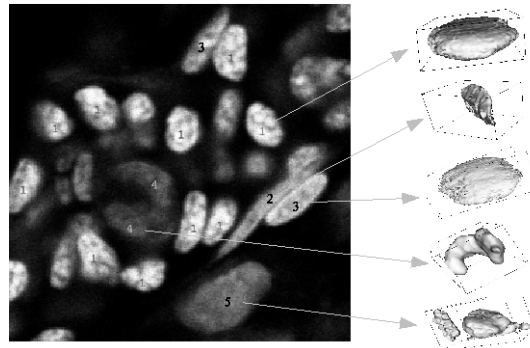
Table 1. Overview of all used three- and two-pint gray-scale invariants of type $f(\mathbf{X}) = f_a(\mathbf{X}(\mathbf{0})) \cdot f_b(\mathbf{X}(\boldsymbol{q}_1)) \cdot f_c(\mathbf{X}(\boldsymbol{q}_2))$. q_α denotes the size of the scope, v_i the i-th channel.

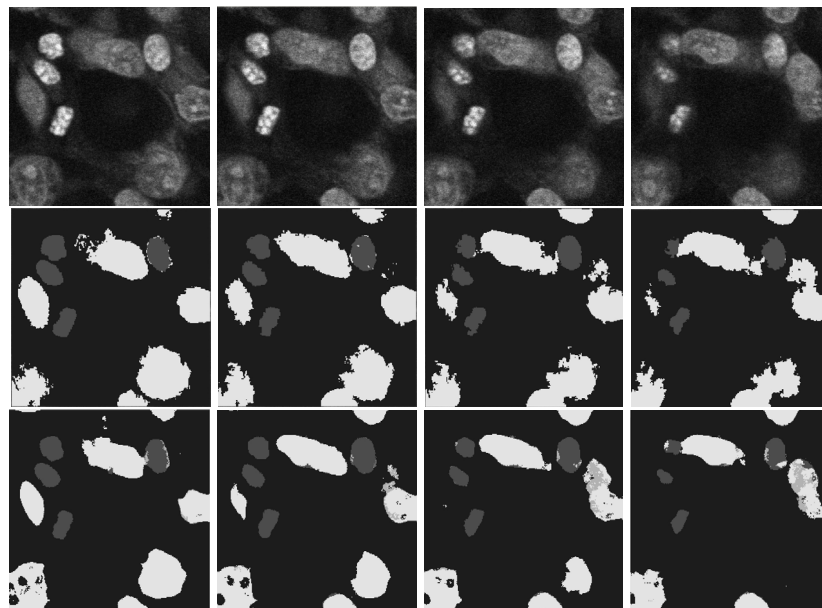|	f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8
f_a	$X_{v_1}(0)$	$(X_{v_1}(0))^5$	$(X_{v_1}(0))^5$	$\sqrt{X_{v_1}(0)}$	$(X_{v_1}(0))^2$	$\sqrt{X_{v_1}(0)}$	$X_{v_1}(0)$	$(X_{v_1}(0))^2$
f_b	$X_{v_1}(0)$	$(X_{v_1}(q_2))^2$	$(X_{v_1}(q_2))^5$	$e^{(-X_{v_1}(q_4))^2}$	$(X_{v_1}(q_4))^5$	$\sqrt{X_{v_1}(q_4)}$	$X_{v_1}(q_4)$	$(X_{v_1}(q_8))^5$
f_c	-	$(X_{v_1}(q_2))^2$	-	-	$(X_{v_1}(q_4))^5$	$\sqrt{X_{v_1}(q_4)}$	$X_{v_1}(q_4)$	$(X_{v_1}(q_8))^5$

	f_{10}	f_{11}	f_{12}	f_{13}	f_{14}	f_{15}	f_{16}
f_a	$(X_{v_1}(0))^5$	$\sqrt{X_{v_1})(0)}$	$(X_{v_1}(0))^5$	$X_{v_1}(0)$	$e^{(-X_{v_1}(0))^2}$	$(X_{v_1}(0))^5$	$(X_{v_1}(0))^5$
f_b	$e^{-X_{v_1}(q_{16})}$	$\sqrt{X_{v_1}(q_{16})}$	$(X_{v_1}(q_{16}))^2$	$X_{v_2}(q_2)$	$e^{(-X_{v_2}(q_2))^2}$	$(X_{v_2}(q_2))^2$	$(X_{v_2}(q_2))^5$
f_c	-	$\sqrt{X_{v_1}(q_{16})}$	$(X_{v_1}(q_{16}))^2$	$X_{v_2}(q_2)$	$e^{(-X_{v_2}(q_2))^2}$	-	$(X_{v_2}(q_2))^5$

	f_{18}	f_{19}	f_{20}	f_{21}	f_{22}	f_{23}	f_{24}
f_a	$(X_{v_1}(0))^2$	$e^{(-X_{v_1}(0))^2}$	$(X_{v_1}(0))^5$	$(X_{v_1}(0))^2$	$\sqrt{X_{v_1}(0)}$	$(X_{v_1}(0))^5$	$(X_{v_1}(0))^5$
f_b	$(X_{v_2}(q_4))^5$	$e^{(-X_{v_2}(q_4))^2}$	$(X_{v_2}(q_4))^2$	$(X_{v_2}(q_8))^5$	$(X_{v_2}(q_8))^5$	$e^{-X_{v_2}(q_{16})}$	$(X_{v_2}(q_{16}))^2$
f_c	-	-	$(X_{v_2}(q_4))^2$	-	$(X_{v_2}(q_8))^5$	-	$(X_{v_2}(q_{16}))^2$

	f_{26}	f_{27}	f_{28}	f_{29}	f_{30}	f_{31}	f_{32}
f_a	$X_{v_2}(0)$	$e^{(-X_{v_2}(0))^2}$	$(X_{v_2}(0))^5$	$(X_{v_2}(0))^5$	$\sqrt{X_{v_2}(0)}$	$(X_{v_2}(0))^2$	$e^{(-X_{v_2}(0))^2}$
f_b	$X_{v_1}(q_2)$	$e^{(-X_{v_1}(q_2))^2}$	$(X_{v_1}(q_2))^2$	$(X_{v_1}(q_2))^5$	$e^{(-X_{v_1}(q_4))^2}$	$(X_{v_1}(q_4))^5$	$e^{(-X_{v_1}(q_4))^2}$
f_c	$X_{v_1}(q_2)$	$e^{(-X_{v_1}(q_2))^2}$	$(X_{v_1}(q_2))^5$	-	$e^{(-X_{v_1}(q_4))^2}$	-	-

Data: The experiments were performed on 3D volumetric data samples of chicken embryo chorioallantoic membrane (CAM) probes recorded by a confocal laser scanning microscope (LSM). The CAM is a widely used model for angiogenesis research. For angiogenesis research at cellular level, an automatic localization and identification of the different cell types is crucial. Understanding angiogenesis has been found key to treatment of many frequent diseases, including cancer and heart ischemia. The samples were prepared as described in [9][10] and treated with YoPro-1 and SMACy3 fluorescent markers.

Methods: A model was built performing the interactive training procedure on several training data sets. Other samples were classified against this model. As reference to our approach we performed seeded watershed segmentation with about one hundred manually set seeds for each cell. A median filter was applied prior to the watershed procedure. For this part we omitted the classification, since the classes were set manually.

Fig. 5. Sample data, cross section of a capillary. Cell types with 3D reconstruction: 1. erythrocyte (**Ery**), 2. endothelial cell (**EC**), 3. pericyte (**PC**), 4. fibroblast (**FB**), 5. macrophage (**M**Φ).

Fig. 6. Results of voxel-wise gray-scale invarinat segmentation in xy-slices. First line: raw data. Second line: watershed reference. Third line: results of our approach.

Fig. 7. Results of voxel-wise gray-scale invarinat segmentation in zy-slices. First line: two channels of raw data. Second line: watershed reference and results of our approach.

Results: For the experiment sown in (Fig. 6), the features in (Table 1) were used. Three-point kernels were restricted to have the points on one straight line and were approximated only by the first coefficient of the series [1].

Conclusion and Outlook. Our algorithm is able to automatically detect previously learned objects. Low fluorescent activity and strong intra cellular structures do not cause false or partial segmentation results. But still the low z-resolution is responsible for miss-classifications at object borders and some noise (Fig. 6). The rather simple approach of connected component labeling is the major drawback at this state - it is neither capable of suppressing small fractions of noise, nor splitting touching objects of the same class.

References

1. Ronneberger, O., Fehr, J., Burkhardt, H.: Voxel-wise gray scale invariants for simultaneous segmentation and classification. In: Pattern Recognition, Proceedings of the 27th DAGM Symposium, Vienna, Austria, Lecture Notes in Computer Science. (2005)
2. Ronneberger, O., Fehr, J., Burkhardt, H.: Voxel-wise gray scale invariants for simultaneous segmentation and classification – theory and application to cell-nuclei in 3d volumetric data. Internal report 2/05, IIF-LMB, University Freiburg (2005)
3. Schulz-Mirbach, H.: Invariant features for gray scale images. In Sagerer, G., Posch, S., Kummert, F., eds.: 17. DAGM - Symposium "Mustererkennung", Bielefeld, Reihe Informatik aktuell, Springer (1995) 1–14
4. Burkhardt, H., Siggelkow, S.: Invariant features in pattern recognition – fundamentals and applications. In Kotropoulos, C., Pitas, I., eds.: Nonlinear Model-Based Image/Video Processing and Analysis, John Wiley & Sons (2001) 269–307
5. Ronneberger, O., Burkhardt, H., Schultz, E.: General-purpose Object Recognition in 3D Volume Data Sets using Gray-Scale Invariants – Classification of Airborne Pollen-Grains Recorded with a Confocal Laser Scanning Microscope. In: Proceedings of the International Conference on Pattern Recognition, Quebec, Canada (2002)
6. Ronneberger, O., Schultz, E., Burkhardt, H.: Automated Pollen Recognition using 3D Volume Images from Fluorescence Microscopy. Aerobiologia **18** (2002) 107–115
7. Vapnik, V.N.: The nature of statistical learning theory. Springer (1995)
8. Ronneberger, O.: Libsvmtl - a support vector machine template library. download at: http://lmb.informatik.uni-freiburg.de/lmbsoft/libsvmtl/ (2004)
9. Kurz, H., et al.: Pericytes in experimental mda-mb231 tumor angiogenesis. Histochem Cell Biol (2002) 117:527–534
10. Kurz, H., et al.: Automatic classification of cell nuclei and cells during embryonic vascular development. Ann Anat 2005; 187 (Suppl): 130. (2005)