Voxel-Wise Gray Scale Invariants for Simultaneous Segmentation and Classification

Olaf Ronneberger, Janis Fehr, and Hans Burkhardt

Albert-Ludwigs-Universität Freiburg, Institut für Informatik,
Lehrstuhl für Mustererkennung und Bildverarbeitung,
Georges-Köhler-Allee Geb. 052, 79110 Freiburg, Deutschland
ronneber@informatik.uni-freiburg.de
http://lmb.informatik.uni-freiburg.de

Abstract. 3D volumetric microscopical techniques (e.g. confocal laser scanning microscopy) have become a standard tool in biomedical applications to record three-dimensional objects with highly anisotropic morphology. To analyze these data in high-throughput experiments, reliable, easy to use and generally applicable pattern recognition tools are required. The major problem of nearly all existing applications is their high specialization to exact one problem, and the their time-consuming adaption to new problems, that has to be done by pattern recognition experts. We therefore search for a tool that can be adapted to new problems just by an interactive training process. Our main idea is therefore to combine object segmentation and recognition into one step by computing voxel-wise gray scale invariants (using nonlinear kernel functions and Haar-integration) on the volumetric multi-channel data set and classify each voxel using support vector machines.

After the selection of an appropriate set of nonlinear kernel functions (which allows to integrate previous knowledge, but still needs some expertise), this approach allows a biologist to adapt the recognition system for his problem just by interactively selecting several voxels as training points for each class of objects. Based on these points the classification result is computed and the biologist may refine it by selecting additional training points until the result meets his needs. In this paper we present the theoretical background and a fast approximative algorithm using FFTs for computing Haar-integrals over the very rich class of nonlinear 3-point-kernel functions. The approximation still fulfils the invariance conditions. The experimental application for the recognition of different cell cores of the chorioallantoic membrane is presented in the accompanying paper [1] and in the technical report [2]

1 Introduction

Three-dimensional microscopical techniques, e.g., confocal laser scanning microscopy, has become a standard tool in biomedical applications within the last few years. Due to the increasing need of high throughput experiments, e.g. in the analysis of 3D gene expression patterns, the gap between the automated

W. Kropatsch, R. Sablatnig, and A. Hanbury (Eds.): DAGM 2005, LNCS 3663, pp. 85–92, 2005.

recording of the data and the tedious and subjective manual evaluation becomes larger and larger.

Due to the rapidly changing requirements for an automatic evaluation, the traditional way of developing highly specialized model-based solutions for exactly one problem with dozens of manually selected morphological processing steps and thresholds does usually not meet the needs of the biologists.

A step towards a generally applicable and easy to use system is presented in this paper. Many of the problems can be reduced to a rotation and translation invariant recognition of certain 3D structures, that are trained by a manually labeled database (e.g., counting or localization of different cell cores). Therefore, we use gray scale invariants [3,4], that have already been successfully applied to the recognition of pollen grains in volumetric data sets [5]. The main limitation of this approach was its need for objects that are already segmented from the background. A good segmentation on the other hand needs already much information about the object to isolate it from the background.

To overcome this classical dilemma, we integrated the segmentation and recognition of the objects into one step by calculating voxel-wise gray scale invariants: For each voxel several rotation invariant features from its surrounding are extracted and the resulting feature vector is classified using support vector machines [6]. The result is a label for each voxel (or several probabilities per voxel). A simple connected component analysis in the next step then searches regions with the same label to segment the objects.

The main advantage of this approach is its direct operation on the raw data and the avoidance of manually selected thresholds. Instead it learns all necessary informations from a labeled training data set.

2 Theory

2.1 Construction of Gray Scale Invariants

The precondition for the use of gray scale invariants in recognizing n-dimensional structures in the real world are:

1. One or more reproducible measurement techniques, that are able to measure certain properties within the structure at definite positions independent from the orientation of the structure (e.g., to measure the fluorescence activity at a certain wavelength at the focal point of a confocal laser scanning microscope) resulting in an n-dimensional multi-channel data set
2. Knowledge of those mathematical transformations, which do not change the meaning of the structure (e.g. rotation and translation)

If these preconditions are fulfilled, we can find a feature extraction, that is able to map all representations of the same object (given by the transformation group) into one point of the feature space by using a nonlinear kernel function and a Haar-integration over the whole transformation group [5].

$$T[f](\mathbf{X}) := \int_G f(g\mathbf{X})dg \qquad \begin{aligned} G &: \text{transformation group} \\ g &: \text{one element of the transformation group} \\ f &: \text{nonlinear kernel function} \\ \mathbf{X} &: n\text{-dim, multi-channel data set} \\ g\mathbf{X} &: \text{the transformed } n\text{-dim data set} \end{aligned} \tag{1}$$

For each nonlinear kernel function f this integral returns a scalar value that describes a certain feature of the n-dimensional data set invariant under the given transformations, as long as the integral exists and is finite.

Reduction to Kernel-Functions with sparse support. 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 $f\big(\mathbf{X}(\boldsymbol{x}_1), \mathbf{X}(\boldsymbol{x}_2), \mathbf{X}(\boldsymbol{x}_3), \ldots\big)$, 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 $s_g(\boldsymbol{x}_i)$ such that

$$(g\mathbf{X})(\boldsymbol{x}_i) := \mathbf{X}(s_g(\boldsymbol{x}_i)) \quad \forall g, \boldsymbol{x}_i \ . \tag{2}$$

With this we can rewrite (1) as

$$T[f](\mathbf{X}) := \int_G f\Big(\mathbf{X}(s_g(\boldsymbol{x}_1)),\ \mathbf{X}(s_g(\boldsymbol{x}_2)),\ \mathbf{X}(s_g(\boldsymbol{x}_3)), \ldots\Big)\, dg \ . \tag{3}$$

This considerably speeds up the computation and results for a given kernel in linear complexity $O(N)$ of the algorithm, where N is the number of voxels in the data set.

Multi Scale Approach. In the general formulation of the gray scale invariants (1) appropriate kernel functions can be used in order to sense any features of the structures at any scales. Computable kernel functions (3) depend only on a few gray values at certain points. To use them for sensing also large-scale informations, a multi scale approach is applied [5]. In the continuous domain this is equivalent to applying a certain low-pass filter (e.g. convolution with a Gaussian) to the data set before evaluating the kernel functions (see Fig. 1)

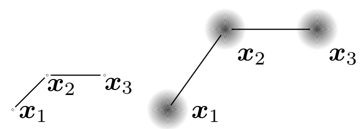

Fig. 1. Computable kernels rely on a small number of sampling points. To sense informations at multiple scales, the sampling points are "enlarged" with Gaussians of mutliple size

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

Voxel-Wise Gray Scale Invariants. For voxel-wise invariants the transformation group is just a rotation, where the origin of the coordinate system is shifted to a certain voxel in advance. The resulting features from the different kernel functions are collected in a feature vector, which then describes the surrounding of this voxel in a unique and rotation invariant way. This is done for all voxels in a volume.

2.2 Computation of Gray Scale Invariants

To compute the gray scale invariants from (3) we first have to select a parameterization $\boldsymbol{\lambda}$ of s_g so that (3) can be rewritten as

$$T[f](\mathbf{X}) := \int f\Big(\mathbf{X}(s_{\boldsymbol{\lambda}}(\boldsymbol{x}_1)),\ \mathbf{X}(s_{\boldsymbol{\lambda}}(\boldsymbol{x}_2)),\ \mathbf{X}(s_{\boldsymbol{\lambda}}(\boldsymbol{x}_3)), \dots\Big)\, d\boldsymbol{\lambda}\,. \tag{4}$$

For a given data set $\mathbf{X}$ and given kernel points $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, $\boldsymbol{x}_3$, ... we can substitute $\mathbf{X}(s_{\boldsymbol{\lambda}}(\boldsymbol{x}_i))$ with $v_i(\boldsymbol{\lambda})$, where $\mathbf{X}(s_{\boldsymbol{\lambda}}(\boldsymbol{x}_i))$ are the gray values, that are touched by the i'th kernel point $\boldsymbol{x}_i$ during all transformations described by $\boldsymbol{\lambda}$, resulting in

$$T = \int f\Big(v_1(\boldsymbol{\lambda}),\ v_2(\boldsymbol{\lambda}),\ v_3(\boldsymbol{\lambda}), \dots\Big)\, d\boldsymbol{\lambda}\,. \tag{5}$$

A simple example for the resulting one-dimensional curves $v_1(\varphi)$, $v_2(\varphi)$ and $v_3(\varphi)$ when using a 3-point kernel on a 2D image and the transformation group of rotations is given in Fig. 2.

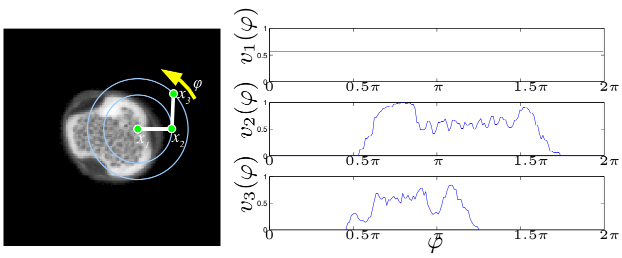

Fig. 2. When using a 3-point kernel on a continuous 2D image and the transformation group of rotations, the gray values, that are touched by the kernel points $\boldsymbol{x}_1$, $\boldsymbol{x}_2$ and $\boldsymbol{x}_3$, are one-dimensional functions $v_1(\varphi)$, $v_2(\varphi)$ and $v_3(\varphi)$

Two-Point Kernel Functions. The direct evaluation of the integral (3) is usually too slow for real applications. A fast calculation method (using FFTs) for a certain class of kernel-functions (so called separable two-point-kernel functions) of the form

$$f(\mathbf{X}) = f_a\Big(\mathbf{X}(\mathbf{0})\Big) \cdot f_b\Big(\mathbf{X}(\boldsymbol{q})\Big) \qquad \begin{aligned} f_a, f_b &: \text{any nonlinear functions that} \\ &\quad \text{transform the gray values} \\ \boldsymbol{q} &: \text{span of the kernel function} \end{aligned} \tag{6}$$

and for Euclidean transformations was presented in [5]: For this purpose we define $\mathbf{A}(\boldsymbol{x}) := f_a(\mathbf{X}(\boldsymbol{x}))$ and $\mathbf{B}(\boldsymbol{x}) := f_b(\mathbf{X}(\boldsymbol{x}))$ The resulting Haar integral (3) is

$$T[f](\mathbf{X}) = \int d\boldsymbol{x}\ \mathbf{A}(\boldsymbol{x}) \cdot (\mathbf{B} * \mathbf{S})(\boldsymbol{x}), \qquad \text{with } \mathbf{S}(\boldsymbol{x}) := \delta\big(|\boldsymbol{x}| - q\big) \tag{7}$$

which is the convolution (denoted as '$*$') of $\mathbf{B}$ with $\mathbf{S}$ (which is a surface of a sphere in 3D or a circle in 2D) and the point-wise multiplication with $\mathbf{A}$. For the evaluation of voxel-wise gray scale invariants the final integration over $\boldsymbol{x}$ is omitted.

Three-Point-Kernel Functions Two-point kernel functions perform very well in Haar integrals over Euclidean motions. For the voxel-wise invariants, they are somewhat limited in their discrimination power, because the resulting invariants are not only invariant to rotation of the surrounding but also to any random permutation of the gray values at the same radius. In contrast to this, three-point kernel functions, where the first point is located at the rotation center[2]

$$f(\mathbf{X}) = f_a\Big(\mathbf{X}(\mathbf{0})\Big) \cdot f_b\Big(\mathbf{X}(\boldsymbol{q}_2)\Big) \cdot f_c\Big(\mathbf{X}(\boldsymbol{q}_3)\Big) \tag{8}$$

are sensitive to such permutations but they cannot be computed directly with the above mentioned fast algorithm, because both factors in the product $f_b(\mathbf{X}(\boldsymbol{q}_1)) \cdot f_c(\mathbf{X}(\boldsymbol{q}_2))$ change when rotating the kernel and therefore cannot be calculated by a simple convolution.

In the following we present an expansion into a series of simple coonvolutions with the nice property, that every truncated evaluation of this series still fulfills the invariance criterion.

For volumetric data the rotations must be parameterized by three angles $\boldsymbol{\lambda} = (\varphi_1, \varphi_2, \varphi_3)^T$ (see Fig. 3).

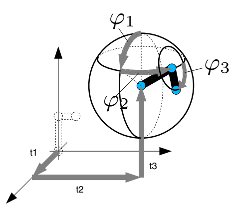

Fig. 3. Parameterization of the 3D rotation with $\boldsymbol{\lambda} = (\varphi_1, \varphi_2, \varphi_3)^T$

The first kernel point is always shifted to the rotation center, which results in $v_1(\boldsymbol{\lambda}) = v_1(\mathbf{0})$. Without changing the result we can rotate the kernel function, such that the second kernel point is located on the z-axis, which makes

[2] For Euclidean transformations, a translation of the kernel function does not change the integral. We use the same terminology for voxel-wise invariants. If you do not plan to integrate over translations in a later step, you could leave out the first kernel point.

$v_2(\boldsymbol{\lambda})$ insensitive to φ_3-rotations, resulting in $v_2(\boldsymbol{\lambda}) = v_2(\varphi_1, \varphi_2, 0)$. With this parameterization the Haar integral becomes

$$T = f_a\Big(v_1(\mathbf{0})\Big) \int_0^\pi d\varphi_1 \sin(\varphi_1) \int_{-\pi}^{\pi} d\varphi_2 \; f_b\Big(v_2(\varphi_1, \varphi_2, 0)\Big) \int_{-\pi}^{\pi} d\varphi_3 \; f_c\Big(v_3(\varphi_1, \varphi_2, \varphi_3)\Big) \,. \tag{9}$$

The integration in the last term can be rewritten as a convolution with a circle on the sphere surface,

$$\begin{aligned} v_{cc}(\varphi_1, \varphi_2) &:= \int_{-\pi}^{\pi} d\varphi_3 \; f_c\Big(v_3(\varphi_1, \varphi_2, \varphi_3)\Big) \\ &= \int_0^\pi d\psi_1 \sin(\psi_1) \int_{-\pi}^{\pi} d\psi_2 \int_{-\pi}^{\pi} d\psi_3 \; f_c\Big(v_3(\psi_1, \psi_2, \psi_3)\Big) \\ &\qquad \cdot \delta\Big(\mathrm{dist}\big((\psi_1, \psi_2, \psi_3), (\varphi_1, \varphi_2, 0)\big) - r\Big) \end{aligned} \tag{10}$$

where "dist" is the distance between two points on a sphere surface and r the distance of the third kernel-point to the "north pole" of the sphere. This reduces the evaluation of the Haar integral to a pixel-wise multiplication and subsequent integration of two gray value data sets on the sphere surfaces. By defining $v_a = f_a\Big(v_1(\mathbf{0})\Big)$ and $v_b(\varphi_1, \varphi_2) = f_b\Big(v_2(\varphi_1, \varphi_2, 0)\Big)$, the Haar integral becomes

$$T = v_a \int_0^\pi d\varphi_1 \sin(\varphi_1) \int_{-\pi}^{\pi} d\varphi_2 \;\; v_b(\varphi_1, \varphi_2) \cdot v_{cc}(\varphi_1, \varphi_2) \tag{11}$$

Analogous to Fourier series in 2D, this can be approximated with spherical harmonics as basis functions. The coefficients are

$$W_{blm} = \int_0^\pi d\varphi_1 \sin(\varphi_1) \int_{-\pi}^{\pi} d\varphi_2 \;\; v_b(\varphi_1, \varphi_2) \, {Y_m^l}^*(\varphi_1, \varphi_2) \tag{12}$$

(W_{cclm} if defined analogously) which allows to write the Haar integral as

$$\begin{aligned} T = v_a \int_0^\pi d\varphi_1 \sin(\varphi_1) \int_{-\pi}^{\pi} d\varphi_2 & \left(\sum_{l_1=0}^{\infty} \sum_{m_1=-l_1}^{l_1} W_{bl_1m_1} Y_{m_1}^{l_1}(\varphi_1, \varphi_2) \right) \\ & \cdot \left(\sum_{l_2=0}^{\infty} \sum_{m_2=-l_2}^{l_2} W_{ccl_2m_2} Y_{m_2}^{l_2}(\varphi_1, \varphi_2) \right) \end{aligned} \tag{13}$$

Using the orthogonality relationships between the basis functions Y_m^l and the precondition, that our data is real-valued, allows to reduce this integral to a simple summation

$$T \sim v_a \sum_{l=0}^{N} \sum_{m=0}^{l} \Re\left(W_{blm} W_{cclm}^*\right) \tag{14}$$

where the series may be truncated after the N'th coefficient without violating the rotation invariance. The full operations are shown in the scheme in Fig. 4.

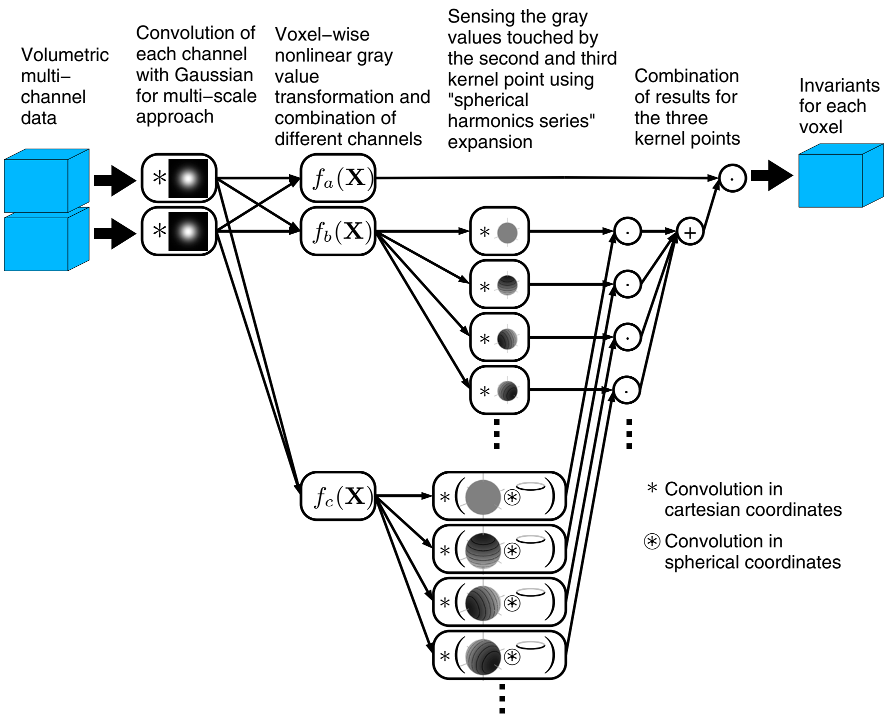

Fig. 4. Computation of three-point-kernel invariants $f(\mathbf{X}) = f_a(\mathbf{X}(\mathbf{0})) \cdot f_b(\mathbf{X}(\boldsymbol{q}_2)) \cdot f_c(\mathbf{X}(\boldsymbol{q}_3))$ on multi-channel volumetric data. For each kernel function this scheme simultaneously calculates the invariants for all voxels.

Voxel-Wise Classification After the extraction of multiple voxel-wise invariants, the feature vector for each voxel is classified with a SVM, that was trained on manually labeled data. A simple connected component analysis on these voxel-wise classification results is used to segment the different objects in the volume.

3 Experiments

Recognition of different cell cores on confocal recordings of the chicken embryo chorioallantoic membrane are promising and show good discrimination performance even for difficult constellations. For details see the accompanying paper [1] or technical report [2].

4 Conclusion and Outlook

Voxel-wise gray scale invariants allow to recognize objects in volumetric multi-channel data without prior segmentation. The presented fast computation algorithms allow to use them in real-world applications. Therefore they build an important step towards a reliable, generally applicable and easy to use pattern recognition system that can be adapted to new problems by a biologist just by an interactive point-and-click procedure.

The next obvious extension will be the use of the voxel-wise classification results (or probabilities) as additional synthetic data channels for additional feature extraction steps. This gives the biologist an easy possibility to integrate his previous knowledge just by decomposing the recognition task into single steps. E.g., in a first step he trains the system to recognize small low-level structures (like cell cores or cell walls) and then combines in the next steps these intermediate results for the recognition of higher level structures.

Present limitations of this framework are the need for manually selected kernel functions. Even though there are already some sets of kernel functions for several applications, our current research focusses on a completely automatic selection of the best kernel functions for a given training data set.

Another problem that cannot be solved with the current approach is segmentation of two neighboring objects of the same class, when there is no significant border between them, that can be trained and classified as an extra class or as background. One solution may be to train seeding points in the center of each object as an additional class. A seeded watershed on the classification results might then be used to crop the two objects at the correct position.

References

1. Fehr, J., Ronneberger, O., Kurz, H., Burkhardt, H.: Self-learning segmentation and classification of cell-nuclei in 3d volumetric data using voxel-wise gray scale invariants. In Kropatsch, W., Sablating, R., eds.: Pattern Recognition - Proc. of the 27th DAGM Symposium, Vienna, Austria, Springer, Berlin (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. Vapnik, V.N.: The nature of statistical learning theory. Springer (1995)