

3D Invariants with High Robustness to Local Deformations for Automated Pollen Recognition

Olaf Ronneberger, Qing Wang, and Hans Burkhardt

Albert-Ludwigs-Universität Freiburg, Institut für Informatik, Lehrstuhl für
Mustererkennung und Bildverarbeitung, Georges-Köhler-Allee Geb. 052,
79110 Freiburg, Deutschland
{ronneber,qwang,burkhardt}@informatik.uni-freiburg.de

Abstract. We present a new technique for the extraction of features from 3D volumetric data sets based on group integration. The features are invariant to translation, rotation and global radial deformations. They are robust to local arbitrary deformations and nonlinear gray value changes, but are still sensitive to fine structures. On a data set of 389 confocally scanned pollen from 26 species we get a precision/recall of 99.2% with a simple 1NN classifier. On volumetric transmitted light data sets of about 180,000 airborne particles, containing about 22,700 pollen grains from 33 species, recorded with a low-cost optic in a fully automated online pollen monitor the mean precision for allergenic pollen is 98.5% (recall: 86.5%) and for the other pollen 97.5% (recall: 83.4%).

1 Introduction

Nearly all worldwide pollen forecasts are still based on manual counting of pollen in air samples under the microscope. Within the BMBF-funded project “OM-NIBUSS” a first demonstrator of a fully automated online pollen monitor was developed, that integrates the collection, preparation and microscopic analysis of air samples. Due to commercial interests, no details of the developed pattern recognition algorithms were published within the last three years. This is the first time that we show how this machine works behind the scenes.

Challenges in pollen recognition. Due to the great intra class variability and only very subtle inter-class differences, automated pollen recognition is a very challenging but still largely unsolved problem. As most pollen grains are nearly spherical and the subtle differences are mainly found near the surface, a pollen expert needs the full 3D information (usually by “focussing through” the transparent pollen grain). An additional difficulty is that pollen grains are often agglomerated and that the air samples contain lots of other airborne particles. For a reliable measurement of high allergenic pollen (e.g. *Artemisia*. A few such pollen grains per m^3 of air can already cause allergic reactions) the avoidance of false positives is one of the most important requirements for a fully automated system.

State of the art. Almost all published articles concerning pollen recognition deal with very low numbers of pollen grains from only a few species and use

manually prepared pure pollen samples, e.g. [1]. Only [4] used a data set from real air samples containing a reasonable number of pollen grains (3686) from 27 species. But even on a reduced data set containing only 8 species and dust particles, the recall was only 64,9% with a precision of 30%.

Main Contribution. In this paper we describe the extension of the Haar-integration framework [9,6,7,8] (further denoted as “HI framework”) to global and local deformations. This is achieved by creating synthetic channels containing the segmentation borders and employing special parameterized kernel functions. Due to the sparsity of non-zero-values in the synthetic channels the resulting integral features are highly localized in the real space, while the framework automatically guarantees the desired invariance properties.

For efficient computation of these integrals we make use of the sparsity of the data in the synthetic channels and use a Fourier or spherical harmonics (“SH”) series expansion (for the desired rotation invariance) to compute multiple features at the same time.

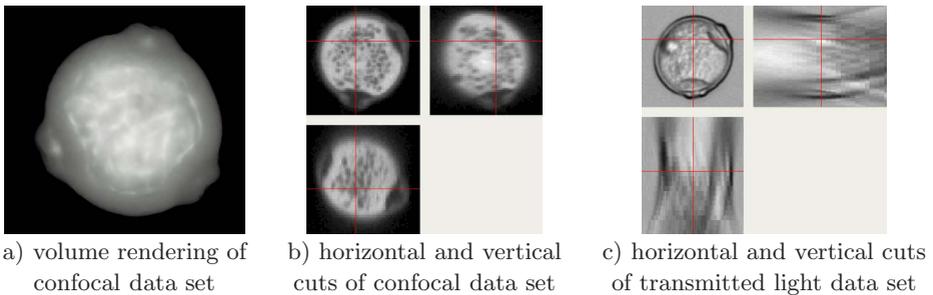


Fig. 1. 3D recordings of *Betula* pollen grains. In transmitted light microscopy the recording properties in z -direction (the direction of the optical axis) are significantly different from those in the xy -direction, because the effects of diffraction, refraction and absorption depend on the direction of the transmitted light. Furthermore there is a significant loss of information in z -direction due to the low-pass property of the optical transfer function.

2 Material and Methods

Data Sets. To demonstrate the generality of the proposed invariants and compare them to earlier results, we use two different pollen data sets in this article. Both contain 3D volumetric recordings of pollen grains.

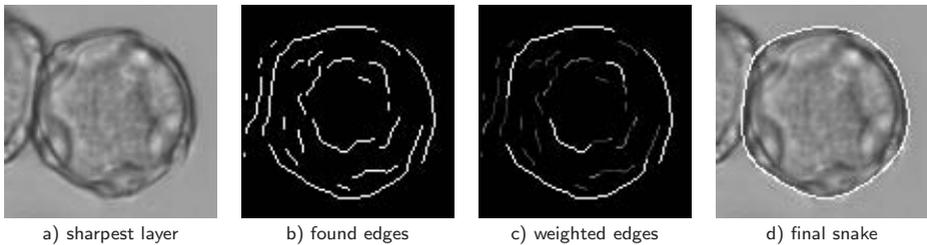
The “*confocal data set*” contains 389 pollen grains from 26 German pollen taxa, recorded with a confocal laser scanning microscope (fig 1a,b). For further details on this data set refer to [6].

The “*pollen monitor data set*” contains about 180,000 airborne particles including about 22,700 pollen grains from air samples that were collected, prepared

and recorded with transmitted light microscopy from the online pollen monitor from March to September 2006 in Freiburg and Zürich (fig. 1c). All 180,000 particles were manually labeled by pollen experts.

Segmentation. To find the 3D surface of the pollen grains in the *confocal data set*, we use the graph cut algorithm described in [2]. The original data were first scaled down. The edge costs to source and sink were modeled by a Gaussian distribution relative to the mean and minimum gray value. We added voxel-to-voxel edges to the 124 neighborhood, where the weight was a Gaussian of the gray differences. The resulting binary mask was then smoothly scaled up to the original size.

The first step in processing the *pollen monitor data set* is the detection of circular objects with voxel-wise vector based gray-scale invariants, similar to those in [8]. For each detected circular object the precise border in the sharpest layer is searched: As parts of the object border are often missing or not clear, we use snakes to find a smooth and complete border. To avoid the common problem of snakes being attracted to undesired edges (if plain gradient magnitude is used as force field), we take the steps depicted in fig 2.



1. Applying modified Canny edge detection.

As pollen grains have a nearly round shape, the edges that are approximately perpendicular to the radial direction are more relevant. We replace the gradient with its radial component in the original Canny edge detection algorithm.

2. Model-based weighting of the edges.

The curvatures and relative locations of the edges are analyzed and each edge is given a different weight. Some edges are even eliminated. As a result, a much clearer weighted edge image is obtained.

3. Employing snakes to find the final border.

The initial contour is chosen to be the circle found in the detection step. The external force field is the so-called "gradient vector flow" [10] computed from the weighted edge image

Fig. 2. Segmentation of transmitted light microscopic images

2.1 Construction of Invariants

For the construction of invariants we use the combination of a normalization and Haar-integration [9,6,7,8](see eq. (1)) over a transformation group containing rotations and deformations (Haar-integration has nothing to do with Haar wavelets). In contrast to the very general approach in [6], we now use the

object center and the outer border found in the segmentation step to extract more distinctive features describing certain regions of the object.

$$T[f](X) := \int_G f(gX) dg \quad \begin{array}{l} G : \text{transformation group} \\ g : \text{one element of the transformation group} \\ dg : \text{Haar measure} \\ f : \text{nonlinear kernel function} \\ X : n\text{-dim, multi-channel data set} \end{array} \quad (1)$$

Invariance to translations. Invariance to translations is achieved by moving the center of mass of the segmentation mask to the origin. The final features are quite insensitive to errors in this normalization step, because they are computed “far” away from this center and only the direction to it is used.

Invariance to rotation. Invariance to rotation around the object center is achieved by integration over the rotation group. In the *confocal data set* we can model a 3D rotation of a real-world object by a 3D rotation of the recorded volumetric data set (see fig. 1b). In contrast to this, the transmitted light microscopic image stacks from the *pollen monitor data set* show very different characteristics in xy- and z-direction, (see fig. 1c). A rotation around the x- or y-axis of the real-world object results in so different gray value distributions, that it is more reasonable to model only the rotation around the z-axis, resulting in a planar rotation invariance.

Invariance to global Deformations and Robustness to local Deformations. The deformation model consists of two parts. The global deformations are modeled by a simple shift in radial direction \mathbf{e}_r , which depends only on the angular coordinates (see figure 3a). For full 3D-rotations described in spherical coordinates $\mathbf{x} = (x_r, x_\varphi, x_\vartheta)$ this model is

$$\mathbf{x}' = \mathbf{x} + \boldsymbol{\gamma}(\mathbf{x}) \quad \text{with} \quad \boldsymbol{\gamma}(\mathbf{x}) = \gamma(x_\varphi, x_\vartheta) \cdot \mathbf{e}_r(x_\varphi, x_\vartheta) . \quad (2)$$

For rotations around the z-axis described in cylindrical coordinates $\mathbf{x} = (x_r, x_\varphi, x_z)$ we get

$$\mathbf{x}' = \mathbf{x} + \boldsymbol{\gamma}(\mathbf{x}) \quad \text{with} \quad \boldsymbol{\gamma}(\mathbf{x}) = \gamma(x_\varphi) \cdot \mathbf{e}_r(x_\varphi) . \quad (3)$$

Please note, that this deformation is well defined only for $r > -\gamma(\varphi)$, which is no problem in the present application, because the features are computed “far” away from the center.

The smaller local deformations are described by an arbitrary displacement field $\mathbf{D}(\mathbf{x})$ such that

$$\mathbf{x}' = \mathbf{x} + \mathbf{D}(\mathbf{x}) \quad (4)$$

(see fig. 3b). For the later partial Haar-integration [3] over all possible realizations of this displacement field, it is sufficient to know only the probability for the occurrence of a certain relative displacement \mathbf{r} within this field as

$$p(\mathbf{D}(\mathbf{x} + \mathbf{d}) - \mathbf{D}(\mathbf{x}) = \mathbf{r}) = p_d(\mathbf{r}; \|\mathbf{d}\|) \quad \forall \mathbf{x}, \mathbf{d} \in \mathbb{R}^3 , \quad (5)$$

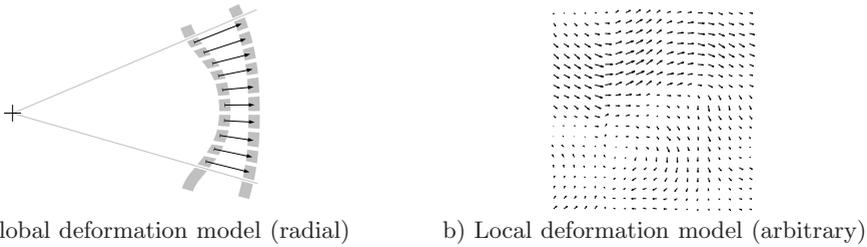


Fig. 3. Possible realizations of the deformation models

where we select $p_d(\mathbf{r}; \|\mathbf{d}\|)$ to be a rotationally symmetric Gaussian distribution with a standard deviation $\sigma = \|\mathbf{d}\| \cdot \sigma_d$.

While we achieve full invariance to radial deformations by full Haar-integration we can only reach robustness to local deformations by partial Haar-integration. But this non-invariance in the second case is exactly the desired behavior. In combination with appropriate kernel functions this results in a continuous mapping of objects (with weak or strong local deformations) into the feature space.

The kernel functions. Instead of selecting a certain fixed number of kernel functions, we introduce parameterized kernel functions here. Embedded into the HI framework, each new combination of kernel parameters results in a new invariant feature. For multiple kernel parameters, we now have a multidimensional invariant feature array describing the object.

Robustness to gray value transformations. To become robust to gray value transformations the information is split into gradient direction (which is very robust even under nonlinear gray value transformations) and gradient magnitude. This was already successfully applied to the HI framework in [8] and to confocal pollen data sets in [5].

Synthetic channels with segmentation results. To feed the segmentation information into the HI framework we simply render the surface (confocal data set) or the contour of the sharpest layer (transmitted light data set) as delta-peaks into a new channel S and extend the kernel-function with two additional points that sense the gray value in this channel. The only condition for this technique is that the computation of the synthetic channel and the action of transformation group can be exchanged without the result being changed (i.e., we must get the same result if we first extract the surface and then rotate and deform the volume and vice versa).

Resulting kernel function. To achieve the requested properties we construct 4-point kernels, where 2 points of the kernel \mathbf{a}_1 and \mathbf{a}_2 sense the segmentation

channel and the other 2 points $\mathbf{b}_1 = \mathbf{a}_1 + \mathbf{q}_1$ and $\mathbf{b}_2 = \mathbf{a}_2 + \mathbf{q}_2$ sense the gradient ∇X of the gray values relative to the information in the segmentation channel,

$$k_1[\mathbf{p}](S, X) = S(\mathbf{a}_1) \cdot \|\nabla X\|(\mathbf{b}_1) \cdot \delta \left(c_1 - \frac{\mathbf{a}_1}{\|\mathbf{a}_1\|} \cdot \frac{\nabla X}{\|\nabla X\|}(\mathbf{b}_1) \right) \cdot S(\mathbf{a}_2) \cdot \|\nabla X\|(\mathbf{b}_2) \cdot \delta \left(c_2 - \frac{\mathbf{a}_2}{\|\mathbf{a}_2\|} \cdot \frac{\nabla X}{\|\nabla X\|}(\mathbf{b}_2) \right) \quad (6)$$

while the delta-functions restrict the kernel to “see” only gradients with the given “direction” c_1 and c_2 . Not all combinations of \mathbf{a}_1 , \mathbf{a}_2 , \mathbf{q}_1 , \mathbf{q}_2 , c_1 and c_2 make sense, because the Haar integration returns identical features for all kernels that are equivalent under the given transformation group. Furthermore for certain combinations, only trivial features will be returned. To ensure, that only non-trivial and non-identical features are created, we introduce the low-dimensional parameterization \mathbf{p} . Examples of this kernel function are depicted in figure 4.

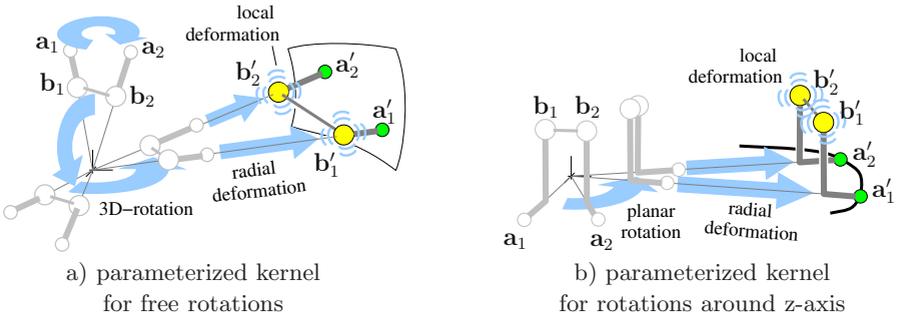


Fig. 4. Action of the transformation group on the selected 4-point kernel functions

As mentioned above, the resulting “structural” features are fully invariant to the global radial deformations. To also extract the shape information, we select the parameterized kernel

$$k_2[\mathbf{a}_1, \mathbf{a}_2](S) = \|\boldsymbol{\gamma}(\mathbf{a}_1)\|S(\mathbf{a}_1) \cdot \|\boldsymbol{\gamma}(\mathbf{a}_2)\|S(\mathbf{a}_2) \quad , \quad (7)$$

that operates on the synthetic channel. When we use the simple scheme of creating the synthetic channels described above, the resulting features are equivalent to the magnitude of the Fourier coefficients of the contour in the 2D case and spherical harmonic (“SH”) coefficients for the surface in the 3D case.

2.2 Fast Simultaneous Computation of the invariants

With the group of radial deformations G_γ , the group of arbitrary deformations G_D and the group of rotations G_R the final Haar integral becomes:

$$T = \int_{G_R} \int_{G_\gamma} \int_{G_D} f(g_R g_\gamma g_D S, g_R g_\gamma g_D X) p(\mathbf{D}) dg_\gamma dg_R \quad , \quad (8)$$

where $p(\mathbf{D})$ is the probability for the occurrence of the local displacement field \mathbf{D} . The transformation of the data set is described by $(gX)(\mathbf{x}) =: X(\mathbf{x}')$, where

$$\mathbf{x}' = \underbrace{\mathbf{R}\mathbf{x}}_{\text{rotation}} + \underbrace{\boldsymbol{\gamma}(\mathbf{R}\mathbf{x})}_{\text{global deformation}} + \underbrace{\mathbf{D}(\mathbf{R}\mathbf{x} + \boldsymbol{\gamma}(\mathbf{R}\mathbf{x}))}_{\text{local deformation}} . \quad (9)$$

To ensure a strong coupling of \mathbf{a}'_1 and \mathbf{b}'_1 (and $\mathbf{a}'_2, \mathbf{b}'_2$ accordingly) we only use kernels, where these two points will be treated equally by the global transformation, i.e., the kernel must fulfill the condition $\boldsymbol{\gamma}(\mathbf{R}\mathbf{a}_i) = \boldsymbol{\gamma}(\mathbf{R}\mathbf{b}_i), \forall \mathbf{R}$ (illustrated by thick connections in fig 4). The transformed kernel points are

$$\begin{aligned} \mathbf{a}'_i(\mathbf{R}, \boldsymbol{\gamma}, \mathbf{D}) &= \mathbf{R}\mathbf{a}_i + \boldsymbol{\gamma}(\mathbf{R}\mathbf{a}_i) + \mathbf{D}(\mathbf{R}\mathbf{a}_i + \boldsymbol{\gamma}(\mathbf{R}\mathbf{a}_i)) \\ \mathbf{b}'_i(\mathbf{R}, \boldsymbol{\gamma}, \mathbf{D}) &= \mathbf{R}\mathbf{b}_i + \boldsymbol{\gamma}(\mathbf{R}\mathbf{a}_i) + \mathbf{D}(\mathbf{R}\mathbf{b}_i + \boldsymbol{\gamma}(\mathbf{R}\mathbf{a}_i)) . \end{aligned} \quad (10)$$

Now inserting the kernel into the Haar integral gives:

$$\begin{aligned} T &= \int_{G_{\mathbf{R}}} \int_{G_{\boldsymbol{\gamma}}} \int_{G_{\mathbf{D}}} S(\mathbf{a}'_1) \cdot \|\nabla_{\mathbf{x}'} X\|(\mathbf{b}'_1) \cdot \delta \left(c_1 - \frac{\mathbf{a}_1}{\|\mathbf{a}_1\|} \cdot \frac{\nabla_{\mathbf{x}'} X}{\|\nabla_{\mathbf{x}'} X\|}(\mathbf{b}'_1) \right) \\ &\quad \cdot S(\mathbf{a}'_2) \cdot \|\nabla_{\mathbf{x}'} X\|(\mathbf{b}'_2) \cdot \delta \left(c_2 - \frac{\mathbf{a}_2}{\|\mathbf{a}_2\|} \cdot \frac{\nabla_{\mathbf{x}'} X}{\|\nabla_{\mathbf{x}'} X\|}(\mathbf{b}'_2) \right) \\ &\quad \cdot p(\mathbf{D}) \cdot d\mathbf{D}(\mathbf{R}\mathbf{a}_1 + \boldsymbol{\gamma}(\mathbf{R}\mathbf{a}_1)) \cdot d\mathbf{D}(\mathbf{R}\mathbf{b}_1 + \boldsymbol{\gamma}(\mathbf{R}\mathbf{a}_1)) \\ &\quad \cdot d\mathbf{D}(\mathbf{R}\mathbf{a}_2 + \boldsymbol{\gamma}(\mathbf{R}\mathbf{a}_2)) \cdot d\mathbf{D}(\mathbf{R}\mathbf{b}_2 + \boldsymbol{\gamma}(\mathbf{R}\mathbf{a}_2)) \\ &\quad \cdot d\boldsymbol{\gamma}(\mathbf{R}\mathbf{a}_1) \cdot d\boldsymbol{\gamma}(\mathbf{R}\mathbf{a}_2) \\ &\quad \cdot d\mathbf{R} , \end{aligned} \quad (11)$$

where $\nabla_{\mathbf{x}'}$ denotes the del operator in the transformed coordinate system. The uncommon notation like $d\boldsymbol{\gamma}(\mathbf{R}\mathbf{a}_1)$ is necessary, because each displacement field is described here with an infinite number of parameters (one displacement for each location in the 3D space). During the integration the outer integral continuously “selects” the integration parameter for the inner integral.

If the synthetic channel is created from a single surface or contour and if we can assume a star-shaped object (which is granted for all considered pollen types) we will find for every given \mathbf{R} only one nonzero response of S during the integration over all deformations $\boldsymbol{\gamma}(\mathbf{R}\mathbf{a}_i)$. By defining this coordinate as $\mathbf{s}(\mathbf{R}\mathbf{a}_i)$ we see that the integral only returns nonzero values for

$$\begin{aligned} \mathbf{s}(\mathbf{R}\mathbf{a}_i) &= \mathbf{a}'_i \\ \Rightarrow \boldsymbol{\gamma}(\mathbf{R}\mathbf{a}_i) &= \mathbf{s}(\mathbf{R}\mathbf{a}_i) - \mathbf{R}\mathbf{a}_i - \mathbf{D}(\mathbf{R}\mathbf{a}_i + \boldsymbol{\gamma}(\mathbf{R}\mathbf{a}_i)) , \end{aligned} \quad (12)$$

which allows to eliminate the direct dependency of \mathbf{b}'_i on $\boldsymbol{\gamma}(\mathbf{R}\mathbf{a}_i)$:

$$\mathbf{b}'_i = \mathbf{s}(\mathbf{R}\mathbf{a}_i) + \mathbf{R}\mathbf{b}_i - \mathbf{R}\mathbf{a}_i + \mathbf{D}(\mathbf{R}\mathbf{b}_i + \boldsymbol{\gamma}(\mathbf{R}\mathbf{a}_i)) - \mathbf{D}(\mathbf{R}\mathbf{a}_i + \boldsymbol{\gamma}(\mathbf{R}\mathbf{a}_i)) . \quad (13)$$

With the additional precondition that the probability for the occurrence of a certain relative displacement \mathbf{r} only depends on the distance of the two considered points (5) we can fully eliminate the dependency of the global transformation $\boldsymbol{\gamma}(\mathbf{R}\mathbf{a}_i)$ and replace the four integrals over the local displacement field

$p(\mathbf{D}) \cdot d\mathbf{D}(\mathbf{R}\mathbf{a}_1 + \boldsymbol{\gamma}(\mathbf{R}\mathbf{a}_1)) \dots$ by the integration over all relative displacements \mathbf{r}_i weighted with their probability $p_d(\mathbf{r}_1; \|\mathbf{d}_1\|) \cdot p_d(\mathbf{r}_2; \|\mathbf{d}_2\|) d\mathbf{r}_1 d\mathbf{r}_2$. The resulting \mathbf{b}'_i is

$$\mathbf{b}'_i(\mathbf{R}, \mathbf{r}_i) = \mathbf{s}(\mathbf{R}\mathbf{a}_i) + \mathbf{R}\mathbf{b}_i - \mathbf{R}\mathbf{a}_i + \mathbf{r}_i \quad (14)$$

and

$$\|\mathbf{d}_i\| = \|\mathbf{R}\mathbf{b}_i + \boldsymbol{\gamma}(\mathbf{R}\mathbf{a}_i) - \mathbf{R}\mathbf{a}_i - \boldsymbol{\gamma}(\mathbf{R}\mathbf{a}_i)\| = \|\mathbf{b}_i - \mathbf{a}_i\|$$

By the substitution of $\mathbf{q}_i := \mathbf{b}_i - \mathbf{a}_i$ the full Haar integral can be written as

$$T = \int_{\mathbf{R}} \int_{\mathbb{R}^3} \|\nabla_{\mathbf{x}'} X\|(\mathbf{b}'_1) \delta\left(c_1 - \frac{\mathbf{a}_1}{\|\mathbf{a}_1\|} \cdot \frac{\nabla_{\mathbf{x}'} X}{\|\nabla X\|}(\mathbf{b}'_1)\right) p_d(\mathbf{r}_1; \|\mathbf{q}_1\|) d\mathbf{r}_1 \\ \int_{\mathbb{R}^3} \|\nabla_{\mathbf{x}'} X\|(\mathbf{b}'_2) \delta\left(c_2 - \frac{\mathbf{a}_2}{\|\mathbf{a}_2\|} \cdot \frac{\nabla_{\mathbf{x}'} X}{\|\nabla X\|}(\mathbf{b}'_2)\right) p_d(\mathbf{r}_2; \|\mathbf{q}_2\|) d\mathbf{r}_2 d\mathbf{R} \quad (15)$$

After integration over the local deformations, this results in two scalar functions defined on a sphere (or a cylinder), that are “scanned” by a simple two-point-kernel, which allows to use the framework introduced in [7] for fast but still fully rotation invariant approximation of the solution. For 3D rotations this framework uses a spherical-harmonics series expansion, and for planar rotations around the z-axis it is simplified to a Fourier series expansion.

Parameterization. For the experiments described in this paper we only used kernels with $\|\mathbf{q}_1\| = \|\mathbf{q}_2\|$ and $c_1 = c_2$. For the application on the *confocal data set* (allowing full 3D rotations) this results in 3 parameters for the kernel: The distance q to the segmentation surface, the relative direction of the gradient c and the desired angular resolution n . For the application on the *pollen monitor data set* (rotational invariance only around the z-axis), q is split into a radial distance q_r to the segmentation border and the z-distance to the central plane q_z .

For the computation, each voxel of the dense 3D data is first projected into the sparse representation in the 4D kernel parameter space, defined by each “arm” of the kernel function (q, c, φ, ϑ for confocal data and q_r, q_z, c, φ for the pollen monitor data). The advantage of this sparseness is that fine detail information from the original images “survive” the smoothing effects of the partial Haar-integration over the local deformation model and the extraction of rotation invariant features.

For the selected kernels with $\|\mathbf{q}_1\| = \|\mathbf{q}_2\|$ and $c_1 = c_2$ a further reduction of the complexity can be achieved, because the final features are only a nonlinear combination of the magnitudes of the computed SH- / Fourier coefficients. As this final recombination does not introduce additional informations we can omit it, and instead use the magnitudes of the SH- / Fourier coefficients directly.

The best sampling of the parameter space of the kernel functions (corresponding to the inner class deformations of the objects), was found by cross validation on the training data set, resulting in $N_{q_r} \times N_{q_z} \times N_c \times n = 31 \times 11 \times 16 \times 16 = 87296$ “structural” features (using kernel function k_1) and 8 “shape” features (using

kernel function k_2). For combination into one feature vector the structural features were normalized to unit sum and the shape features were multiplied by 0.01 . For the *confocal data set* this resulted in $N_q \times N_c \times n = 64 \times 7 \times 2 = 896$ “structural” features (using kernel function k_1). The “shape” features were not yet used here. For both data sets, σ_d , describing the allowed local deformations, was set to 0.1 .

3 Experiments

Experimental setup. For the *confocal data set* a simple 1 Nearest Neighbor classifier (using the L1-norm) was used.

The *pollen monitor data set* was split to approximately equal sized sets by using the air samples with an even index as training set and that with odd index as test set. From the training set only the “clean” (not agglomerated, not contaminated) pollen and the “non-pollen” particles from a few samples were used to train the support vector machine (SVM) using the RBF-kernel (radial basis function) and the one-vs-rest multi-class approach. The best parameters were selected using cross-validation on the training data set. After that the resulting SVM was used to classify all particles (about 100,000) in the training set and the false classified “non-pollen” objects were added to the final SVM-training set.

The detection step (before segmentation) only finds circular objects. For the very few non-circular (fortunately also non-allergenic) species like pinus, often only fragments are segmented. These fragments are simply labeled as “non-pollen”.

Results. On the *confocal data set* we got a recognition rate of **99.2%** with the simple 1NN classifier using a leave-one-out test, which is a significant improvement to best published 1NN result on the same data set: 94,5% [5]

Table 1. Confusion matrix for pollen monitor samples. The pollen grains that the biologists were not able to recognize (“indeterm.”) were left out from the statistics. Due to space limitations only the recall of the results with “no rejection” are given.

	with rejection								no rejection
	rej.	no p.	Cory.	Alnu.	Betu.	Poac.	Arte.	other	recall (%)
indeterm.	(826)	(1518)	(5)	(8)	(23)	(40)	(3)	(461)	
no pollen	1882	77430	1	7	3	1	0	119	
Corylus	14	6	75	0	0	0	0	0	78.9
Alnus	96	41	0	751	2	0	0	0	84.4
Betula	86	10	1	4	933	1	0	0	90.1
Poaceae	84	7	0	0	1	576	0	10	85.0
Artemisia	6	0	0	0	0	0	24	0	80.0
other pollen	814	195	1	4	5	3	1	5126	83.4
precision (%)			96.2	98.0	98.8	99.1	96.0	97.5	
mean precision (allergenic):					98.5%				93.8%
mean recall (allergenic):					86.5%				93%

On the *pollen monitor data set*, for allergenic pollen species, we got a mean precision of **98.5%** at a recall of 86.5% or a precision of 93.8% at a recall of 93% depending on the selected rejection scheme. The details are shown in table 1. Due to the limited space only the results of the 5 allergenic pollen taxa are explicitly listed, the remaining 28 pollen taxa were combined into the row “other pollen”. Objects were rejected, when the SVM returned no or more than one positive decision values. The results in the small table on the right “no rejection” were obtained by always assigning the class with the highest decision value.

There were several air samples with 100% precision and 100% recall and other with very low recall, mainly caused by extreme climate conditions (e.g. snow flakes, that melted on the air sample and created a big cluster of particles in multiple layers) or malfunction of the pollen monitor (e.g. vibrations during the recording of the image stack, misadjustment of the optics , etc.)

4 Conclusions and Outlook

The integration of deformation models into the feature extraction seems to be a central step for a reliable recognition of biological structures. At least for the presented application on automated pollen recognition this technique produced results that are better than all comparable results published elsewhere (e.g., [1,4]). Furthermore it has proven to also work outside the clean laboratory world in a real routine application. Anyhow, we can expect that pollen recognition still remains a challenging research area. In pollen monitor data sets from more than only one machine, one year and two cities, the variations will be even larger.

References

1. Allen, G.P., Hodgson, R.M., Marsland, S.R., Arnold, G., Flemmer, R.C., Flenley, J., Fountain, D.W.: Automatic recognition of light-microscope pollen images. In: Proc of Image Vision and Computing New Zealand (2006)
2. Boykov, Y., Kolmogorov, V.: An experimental comparison of min-cut/max- flow algorithms for energy minimization in vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 26(9), 1124–1137 (2004)
3. Haasdonk, B., Halawani, A., Burkhardt, H.: Adjustable invariant features by partial Haar-integration. In: Proc. of the 17th ICPR, Cambridge (2004)
4. Ranzato, M., Taylor, P.E., House, J.M., Flagan, R.C., LeCun, Y., Perona, P.: Automatic recognition of biological particles in microscopic images. *Pattern Recognition Letters* 28(1), 31–39 (2007)
5. Reiser, M., Burkhardt, H.: Invariant features for 3d-data based on group integration using directional information and spherical harmonic expansion. In: Proc. of the ICPR, Hong Kong (2006)
6. 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: Proc. of the ICPR, Quebec (2002)

7. Ronneberger, O., Fehr, J., Burkhardt, H.: Voxel-wise gray scale invariants for simultaneous segmentation and classification. In: Kropatsch, W.G., Sablatnig, R., Hanbury, A. (eds.) *Pattern Recognition*. LNCS, vol. 3663, Springer, Heidelberg (2005)
8. Schulz, J., Schmidt, T., Ronneberger, O., Burkhardt, H., Pasternak, T., Dovzhenko, A., Palme, K.: Fast scalar and vectorial grayscale based invariant features for 3d cell nuclei localization and classification. In: Franke, K., Müller, K.-R., Nickolay, B., Schäfer, R. (eds.) *Pattern Recognition*. LNCS, vol. 4174, Springer, Heidelberg (2006)
9. Schulz-Mirbach, H.: Invariant features for gray scale images. In: Sagerer, G., Posch, S., Kummert, F. (eds.) *17. DAGM-Symposium "Mustererkennung"*, Informatik aktuell, pp. 1–14. Springer, Heidelberg (1995)
10. Xu, C., Prince, J.L.: Snakes, shapes, and gradient vector flow. *IEEE Transactions on Image Processing* 7(3), 359–369 (1998)