ALBERT-LUDWIGS-UNIVERSITÄT FREIBURG
INSTITUT FÜR INFORMATIK
Lehrstuhl für Mustererkennung und Bildverarbeitung

Harmonic Filters in 3D - Theory and Applications

Internal Report 1/09

Marco Reisert

February 2009

Harmonic Filters in 3D - Theory and Applications

Marco Reisert
Computer Science Department
Albert-Ludwigs-University Freiburg
79110 Freiburg, Germany
reisert@informatik.uni-freiburg.de

February, 2009

Abstract

This report proposes a concept for $SE(3)$-equivariant non-linear filters for multiple purposes, especially in the context of feature and object detection. The idea of the approach is to compute local descriptors as projections onto a local harmonic basis. These descriptors are mapped in a non-linear way onto new local harmonic representations, which then contribute to the filter output in a linear way. This approach may be interpreted as a kind of voting procedure in the spirit of the generalized Hough transform, where the local harmonic representations are interpreted as a voting function. On the other hand, the filter has similarities with classical low-level feature detectors (like corner/blob/line detectors), just extended to the generic feature/object detection problem. The proposed approach fills the gap between low-level feature detectors and high-level object detection systems based on the generalized Hough transform. As an introductory example we use the proposed approach for edge preserving denoising. Secondly, we will apply the proposed filter to a feature detection task on confocal microscopical images of airborne pollen and compare the results to a 3D-extension of a popular GHT-based approach and to a classification per voxel solution.

1 Introduction

The theory of non-linear filters is well developed for image translations. It is known as Volterra theory. Volterra theory states that any non-linear translation-invariant system can be modelled as an infinite sum of multidimensional convolution integrals. More precisely, a filter H is said to be equivariant with respect to some group $\mathcal{G}$, if $gH\{f\} = H\{gf\}$ holds for all images f and all $g \in \mathcal{G}$, where

gf denotes the action of the group to the image f. For the group of translations (or the group of time-shifts) such filters are called Volterra series. In this paper we want to develop non-linear filters that are invariant with respect to Euclidean motion $SE(3)$, therefore, we need a generalization of Volterra's principle to $SE(3)$. In [1] a 2D non-linear filter was proposed that is $SE(2)$-equivariant. The filter was derived from the general concept of group integration which replaced Volterra's principle. In this paper we want to generalize this filter to $SE(3)$. The generalization is not straightforward because the two-dimensional rotation group $SO(2)$ essentially differs from its three-dimensional counterpart $SO(3)$.

As already mentioned the derivation of the filter in [1] was based on the principle of group integration. In this paper we want to follow a more pragmatic way and directly propose the 3D filter guided by its 2D analogon. Let us recapitulate the workflow of the holomorphic filter and give a sketch of its 3D counterpart. In a first step the holomorphic filter computes several convolutions with functions of the form $z^j e^{-|z|^2}$ where $z = x + \mathbf{i}y$ is the pixel coordinate in complex notation. Note, that the monomial $z^j = r^j e^{\mathbf{i}j\phi}$ is holomorphic. The results of these convolutions show a special rotation behavior, e.g. for $j = 1$ it behaves like a gradient field or for $j = 2$ it behaves like a 2nd rank tensor field. Several products of these convolution results are computed. These products show again a special rotation behavior. For example, if we multiply a gradient field ($j_1 = 1$) and a 2-tensor-field ($j_2 = 2$) we obtain a third-order field with $j = j_1 + j_2 = 3$. According to the transformation behavior of the products they are again convolved with functions of the form $\overline{z}^j e^{-|z|^2}$ such that the result of the convolution transforms like a scalar ($j = 0$). This is the principle of the holomorphic filter which we want to generalize to 3D.

The first question is, what are the function corresponding to $z^j e^{-|z|^2}$ in 3D? We know that the real and imaginary part of a holomorphic polynomial are harmonic polynomials. Harmonic polynomials solve the Laplace equation. As $z^j e^{-|z|^2}$ is a Gaussian-windowed holomorphic monomial we will use instead a Gaussian-windowed harmonic polynomial for the 3D filter. The second question is, how can we form products of convolutions with harmonic polynomials that entail their transformation behavior? We will find out that the Clebsch-Gordan coefficients that are known from quantum mechanics provide such products. Given two tensor fields of a certain degree we are able to form a new tensor field of another degree by a certain multiplication and weighted summation of the input fields. The weights in the summations are the Clebsch Gordan coefficients. In [1] and [2] it was shown that the convolutions with the Gaussian-windowed holomorphic basis can be computed efficiently with complex derivatives. In fact, there is a very similar approach in 3D by so called spherical derivatives [3].

The paper is organized as follows: in the following section we give a small overview about related work. In Section 2 we introduce the basics in spherical tensor analysis. We introduce the spherical product which couples spherical tensor fields and introduce basics about spherical harmonics. We also introduce so

called spherical derivatives that are the counterpart to the usual complex derivatives in 2D. They will help us to compute the occurring convolutions in an efficient manner. In Section 3 we introduce the Harmonic filter. We show how the filter can be generalized to tensor valued inputs and outputs and how its parameters can be trained to a specific problem. Section 4 shows how the filter can be implemented efficiently and how it can be applied for feature detection in confocal microscopical images. In Section 5 we conclude and give an outlook for future work.

1.1 Related Work

We figured out four frameworks that are closely related to ours: Volterra filters, Steerable filters, Tensor Voting and the generalized Hough transform.

Volterra filters are the canonical generalization of the linear convolution to a nonlinear mapping. They are widely used in the signal processing community and also find applications in image processing tasks [4, 5, 6]. The filter proposed in this work might be interpreted as a kind of 'joint' Volterra filter for translation and rotation.

Steerable filters, introduced in [7], are a common tool in early vision and image analysis. A generalization for non-group like deformations was proposed in [8] using an approximative scheme. Applications of steerable filters are widespread, an example in 3D is [9]. Our filter computes a certain subset of gaussian-windowed spherical moments in a first step which is actually a steerable filter.

The generalized Hough transform (GHT) [10] is a major tool for the detection of arbitrary shapes. Many modern approaches [11, 12] for object detection and recognition are based on the idea that local parts of the object cast votes for the putative center of the object. If the proposed algorithm is used in the context of object detection, it may be interpreted as some kind of voting procedure for the object center. This voting interpretation also relates our approach to the Tensor Voting [13] framework (TV). However, in TV the voting function does not depend on the local context. Contrarily the proposed filter is able to cast context dependend votes. In particular, in [14] the tensor voting process in 2D is formulated with the use of steerable filters. In a similar way our filter can also interpreted as an implementation of a voting process using steerable filters, but in 3D.

2 Spherical Tensor Analysis

In the following we shortly repeat the basic notions in 3D harmonic analysis as they were introduced and proved in [3]. For introductory reading we recommend mostly literature [15, 16] concerning the quantum theory of the angular momentum, while our representation tries to avoid terms from quantum theory to also give the non-physicists a chance to follow. See e.g. [17, 18] for an introduction from an image processing/engineering viewpoint.

2.1 Preliminaries

Let $\mathbf{D}_g^j$ be the unitary irreducible representation of a $g \in SO(3)$ of order j with $j \in \mathbb{N}$. They are also known as the *Wigner D-matrices* (see e.g. [16]). The representation $\mathbf{D}_g^j$ acts on a vector space V_j which is represented by $\mathbb{C}^{2j+1}$. The standard basis of $\mathbb{C}^{2j+1}$ is written as $\mathbf{e}_m^j$. We write the elements of V_j in bold face, e.g. $\mathbf{u} \in V_j$ and write the $2j+1$ components in unbold face $u_m \in \mathbb{C}$ where $m = -j, \dots j$. For the transposition of a vector/matrix we write $\mathbf{u}^T$; the joint complex conjugation and transposition is denoted by $\mathbf{u}^\top = \overline{\mathbf{u}}^T$. In this terms the unitarity of $\mathbf{D}_g^j$ is expressed by the formula $(\mathbf{D}_g^j)^\top \mathbf{D}_g^j = \mathbf{I}$.

Note, that we treat the space V_j as a real vector space of dimensions $2j+1$, although the components of $\mathbf{u}$ might be complex. This means that the space V_j is only closed under weighted superpositions with real numbers. As a consequence we observe that the components are interrelated by $\overline{u_m} = (-1)^m u_{-m}$. From a computational point of view this is an important issue. Although the vectors are elements of $\mathbb{C}^{2j+1}$ we just have to store just $2j+1$ real numbers. So, the standard coordinate vector $\mathbf{r} = (x, y, z)^T \in \mathbb{R}^3$ has a natural relation to elements $\mathbf{u} \in V_1$ in the form of

$$\mathbf{u} = \begin{pmatrix} \overline{w} \\ z \\ -w \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}}(x - \mathbf{i}y) \\ z \\ -\frac{1}{\sqrt{2}}(x + \mathbf{i}y) \end{pmatrix} = \mathbf{S}\mathbf{r} \in V_1$$

Note, that $\mathbf{S}$ is an unitary coordinate transformation. Actually, the representation $\mathbf{D}_g^1$ is directly related to the real valued rotation matrix $\mathbf{U}_g \in \mathbb{R}^{3\times 3}$ by $\mathbf{D}_g^1 = \mathbf{S}\mathbf{U}_g\mathbf{S}^\top$

Definition 1. *A function $\mathbf{f} : \mathbb{R}^3 \mapsto V_j$ is called a spherical tensor field of rank j if it transforms with respect to rotations as follows*

$$(g\mathbf{f})(\mathbf{r}) := \mathbf{D}_g^j \mathbf{f}(\mathbf{U}_g^T \mathbf{r})$$

for all $g \in SO(3)$. The space of all spherical tensor fields of rank j is denoted by $\mathcal{T}_j$.

2.2 Spherical Tensor Coupling

We define a family of symmetric bilinear forms that connect tensors of different ranks.

Definition 2. *For every $j \geq 0$ we define a family of symmetric bilinear forms of type*

$$\bullet_j : V_{j_1} \times V_{j_2} \mapsto V_j$$

where $j_1, j_2 \in \mathbb{N}$ has to be chosen according to the triangle inequality $|j_1 - j_2| \leq j \leq j_1 + j_2$ and $j_1 + j_2 + j$ has to be even. It is defined by

$$(\mathbf{e}_m^j)^\top (\mathbf{v} \bullet_j \mathbf{w}) := \sum_{m=m_1+m_2} \frac{\langle jm \mid j_1 m_1, j_2 m_2 \rangle}{\langle j0 \mid j_1 0, j_2 0 \rangle} v_{m_1} w_{m_2}$$

where $\langle jm \mid j_1 m_1, j_2 m_2 \rangle$ are the Clebsch Gordan coefficients (see e.g. [16]).

Up to the factor $\langle j0 \mid j_1 0, j_2 0 \rangle$ this definition is just the usual spherical tensor coupling equation which is very well known in quantum mechanics of the angular momentum. The additional factor is for convenience. It normalizes the product such that it shows a more gentle behavior with respect to the spherical harmonics as we will see later.

The characterizing property of these products is that they respect the rotations of the arguments, namely

Proposition 1. *Let $\mathbf{v} \in V_{j_1}$ and $\mathbf{w} \in V_{j_2}$, then for any $g \in SO(3)$*

$$(\mathbf{D}_g^{j_1}\mathbf{v}) \bullet_j (\mathbf{D}_g^{j_2}\mathbf{w}) = \mathbf{D}_g^{j}(\mathbf{v} \bullet_j \mathbf{w})$$

holds.

For the special case $j = 0$ the arguments have to be of the same rank due to the triangle inequality. Actually, in this case the new product coincides with the standard inner product $\mathbf{v} \bullet_0 \mathbf{w} = \mathbf{w}^\top \mathbf{v}$. Further note, that if one of the arguments of $\bullet$ is a scalar, then $\bullet$ reduces to the standard scalar multiplication, i.e. $v \bullet_j \mathbf{w} = v\mathbf{w}$, where $v \in V_0$ and $\mathbf{w} \in V_j$. Another remark is that $\bullet$ is not associative.

The introduced product can also be used to combine tensor fields of different rank by point-wise multiplication.

Proposition 2. *Let $\mathbf{v} \in \mathcal{T}_{j_1}$ and $\mathbf{w} \in \mathcal{T}_{j_2}$ and j chosen such that $|j_1 - j_2| \leq j \leq j_1 + j_2$, then*

$$\mathbf{f}(\mathbf{r}) = \mathbf{v}(\mathbf{r}) \bullet_j \mathbf{w}(\mathbf{r})$$

is in $\mathcal{T}_j$, i.e. a tensor field of rank j.

If we consider a translation of the image function, that is

$$(\tau\mathbf{f})(\mathbf{r}) := \mathbf{f}(\mathbf{r} - \mathbf{t}_\tau).$$

we can verify that for any two tensor fields $\mathbf{v}$ and $\mathbf{u}$ and any $g \in SE(3)$ we obtain that $(g\mathbf{v}) \bullet_j (g\mathbf{w}) = g(\mathbf{v} \bullet_j \mathbf{w})$. Actually, there is another way to combine two tensor fields: by convolution. For the convolution the evolving product respects the translation in a different sense.

Proposition 3. *Let $\mathbf{v} \in \mathcal{T}_{j_1}$ and $\mathbf{w} \in \mathcal{T}_{j_2}$ and j chosen such that $|j_1 - j_2| \leq j \leq j_1 + j_2$, then*

$$(\mathbf{v}\widetilde{\bullet}_j\mathbf{w})(\mathbf{r}) := \int_{\mathbb{R}^3} \mathbf{v}(\mathbf{r}' - \mathbf{r}) \bullet_j \mathbf{w}(\mathbf{r}')\, d\mathbf{r}'$$

is in $\mathcal{T}_j$, i.e. a tensor field of rank j.

Now, consider the behavior of $\widetilde{\bullet}_j$ with respect to a translation τ. The following relation is simple to be verified

$$\mathbf{v}\widetilde{\bullet}_j(\tau\mathbf{w}) \quad = \quad (\tau\mathbf{v})\widetilde{\bullet}_j\mathbf{w} = \tau(\mathbf{v}\widetilde{\bullet}_j\mathbf{w}) \tag{1}$$

due to properties of the convolution.

2.3 Spherical and Solid Harmonics

We denote the well-known spherical harmonics by $\mathbf{Y}^j : S^2 \to V_j$. We write $\mathbf{Y}^j(\mathbf{r})$, where $\mathbf{r}$ may be an element of $\mathbb{R}^3$, but $\mathbf{Y}^j(\mathbf{r})$ is independent of the magnitude of $\mathbf{r}$. We know that the $\mathbf{Y}^j$ provide an orthogonal basis of scalar functions on the 2-sphere S^2. Thus, any real scalar field $f \in \mathcal{T}_0$ can be expanded in terms of spherical harmonics in a unique manner. In the following, we use Racah's normalization (also known as semi-Schmidt normalization), i.e. $\langle Y^j_m, Y^{j'}_{m'} \rangle_{S^2} = \frac{1}{2j+1}\delta_{jj'}\delta_{mm'}$. One important and useful property is that $\mathbf{Y}^j = \mathbf{Y}^{j_1} \bullet_j \mathbf{Y}^{j_2}$. We can use this formula to iteratively compute higher order $\mathbf{Y}^j$ from given lower order ones. Note that $\mathbf{Y}^0 = 1$ and $\mathbf{Y}^1 = \mathbf{Sr}$, where $\mathbf{r} \in S^2$. The spherical harmonics have a variety of nice properties. One of the most important ones is that each $\mathbf{Y}^j$, interpreted as a tensor field of rank j is a fix-point with respect to rotations, i.e. $(g\mathbf{Y}^j)(\mathbf{r}) = \mathbf{Y}^j(\mathbf{r})$ or in other words $\mathbf{Y}^j(\mathbf{U}_g\mathbf{r}) = \mathbf{D}^j_g\mathbf{Y}^j(\mathbf{r})$. The spherical harmonics naturally arise from the solutions from the Laplace equation as the so called solid harmonics $\mathbf{R}^j(\mathbf{r}) := r^j\mathbf{Y}^j(\mathbf{r})$. The generalization $\mathbf{R}^j_i(\mathbf{r}) := r^{j+i}\mathbf{Y}^{j-i}(\mathbf{r})$ is a complete basis of the analytical functions on $\mathbb{R}^3$, i.e. we can write any analytical function as

$$f(\mathbf{r}) = \sum_{j \geq i} (\mathbf{a}^j_i)^\top \mathbf{R}^j_i(\mathbf{r}) \tag{2}$$

In this way we can project any analytical function on its harmonic part by restricting on the summands with $i = 0$. Note that the harmonic part depends on the choice of the origin. Thus we cannot imagine that there is a unique harmonic part of an analytical functions, moreover there are many ways to compute this projection. Further note that the $\mathbf{R}^j_i$ are homogeneous polynomials of order $j + i$, meaning $\mathbf{R}^j_i(\lambda\mathbf{r}) = \lambda^{j+i}\mathbf{R}^j_i(\mathbf{r})$ for any $\lambda \in \mathbb{R}$.

2.4 Spherical Derivatives

This section proposes the basic tools for dealing with derivatives in the context of spherical tensor analysis. In [3] the spherical derivatives are introduced. They connect spherical tensor fields of different ranks by differentiation. Specifically, we will examine the spherical Gaussian derivatives. They are just the Gaussian-windowed harmonic polynomials that we have mentioned in the introduction.

Proposition 4 (Spherical Derivatives)**.** *Let $\mathbf{f} \in \mathcal{T}_j$ be a tensor field. The spherical up-derivative $\boldsymbol{\nabla}^1 : \mathcal{T}_j \to \mathcal{T}_{j+1}$ and the down-derivative $\boldsymbol{\nabla}_1 : \mathcal{T}_j \to \mathcal{T}_{j-1}$ are defined as*

$$\boldsymbol{\nabla}^1\mathbf{f} := \nabla \bullet_{j+1} \mathbf{f} \tag{3}$$

$$\boldsymbol{\nabla}_1\mathbf{f} := \nabla \bullet_{j-1} \mathbf{f}, \tag{4}$$

where

$$\nabla = (\frac{1}{\sqrt{2}}(\partial_x - \mathbf{i}\partial_y), \partial_z, -\frac{1}{\sqrt{2}}(\partial_x + \mathbf{i}\partial_y))$$

is the spherical gradient and $\partial_x, \partial_y, \partial_z$ the standard partial derivatives.

Note, that for a scalar function the spherical up-derivative is just the spherical gradient, i.e. $\nabla f = \boldsymbol{\nabla}^1 f$.

In the Fourier domain the spherical derivatives act by point-wise $\bullet$-multiplication with a solid harmonic $\mathbf{i}k\mathbf{Y}^1(\mathbf{k}) = \mathbf{i}\mathbf{R}^1(\mathbf{k}) = \mathbf{i}\mathbf{S}\mathbf{k}$ where $k = \|\mathbf{k}\|$ the frequency magnitude:

Proposition 5 (Multiple Spherical Derivatives)**.** *For $j \geq i$ we define $\boldsymbol{\nabla}_i^j : \mathcal{T}_0 \to \mathcal{T}_{j-i}$ by*

$$\boldsymbol{\nabla}_i^j := \boldsymbol{\nabla}_i \boldsymbol{\nabla}^j := \underbrace{\boldsymbol{\nabla}_1 \dots \boldsymbol{\nabla}_1}_{i-\textit{times}} \underbrace{\boldsymbol{\nabla}^1 \dots \boldsymbol{\nabla}^1}_{j-\textit{times}}.$$

In the Fourier domain these multiple derivatives are given by

$$(\widetilde{\boldsymbol{\nabla}}_i^j \widetilde{f})(\mathbf{k}) \quad = \quad (\mathbf{i})^{j+i}\, \mathbf{R}_i^j(\mathbf{k}) \widetilde{f}(\mathbf{k}), \tag{5}$$

Using this one can show that $\boldsymbol{\nabla}_i^j = \boldsymbol{\nabla}^{j-i}\Delta^i$, where Δ is the Laplace operator.

The above proposition is mainly due to the fact $\mathbf{Y}^j = \mathbf{Y}^{j_1} \bullet_j \mathbf{Y}^{j_2}$. The introduced spherical derivatives can be used to compute the expansion coefficients of the expansion introduced in equation 2. In fact the $\mathbf{a}_i^j$ are given by

$$\mathbf{a}_i^j = \frac{(2(j-i)+1)}{i!2^i(2j+1)!!}(\boldsymbol{\nabla}_i^j f)(\mathbf{0}).$$

(see [3] for a proof). In this way we obtain the spherical equivalent of the usual Taylor expansion in cartesian coordinates.

2.5 Spherical Gaussian Derivatives

As a prerequisite to the Harmonic filter we have to show that the Gaussian-windowed solid harmonics have a very special behavior with respect to the Fourier transform. Actually, they are the $\boldsymbol{\nabla}^j$-derivatives of the Gaussian.

Proposition 6. *The Gaussian windowed harmonic of width σ is defined as*

$$\mathbf{G}_\sigma^j(\mathbf{r}) := \frac{1}{(\sqrt{2\pi}\sigma)^3} \left(\frac{-1}{\sigma^2}\right)^j \mathbf{R}^j(\mathbf{r}) e^{-\frac{r^2}{2\sigma^2}},$$

then

$$\widetilde{\mathbf{G}}_\sigma^j(\mathbf{k}) = \langle e^{\mathbf{i}\mathbf{k}^\top \mathbf{r}}, \mathbf{G}^j(\mathbf{r}) \rangle_{L_2} = (\mathbf{i})^j \mathbf{R}^j(\mathbf{k}) e^{-\frac{(\sigma k)^2}{2}}.$$

is the Fourier transformation of $\mathbf{G}^j(\mathbf{r})$.

In fact, for $\sigma = 1$ the $\mathbf{G}^j$s are eigenfunctions of the Fourier transformation with eigenvalues $\sqrt{2\pi}^3 \mathbf{i}^j$. They form a orthogonal basis of the space of Gaussian windowed harmonic functions (with respect to the standard inner product). Or, in other words they are a orthogonal basis of the space of harmonic functions with respect to the Gaussian measure $d_\sigma \mathbf{x} \propto e^{-\frac{r^2}{\sigma^2}} d\mathbf{x}$. Using the above proposition it is also easy to show that the $\mathbf{G}^j$ are just the jth order spherical derivatives of a Gaussian.

Corollary 1 (Spherical Gaussian Derivative)**.** *The spherical derivative $\boldsymbol{\nabla}^j$ of a Gaussian computes to*

$$\boldsymbol{\nabla}^j e^{-\frac{r^2}{2\sigma^2}} = (\sqrt{2\pi}\sigma)^3 \mathbf{G}^j_\sigma(\mathbf{r}) = \left(-\frac{1}{\sigma^2}\right)^j \mathbf{R}^j(\mathbf{r})\, e^{-\frac{r^2}{2\sigma^2}}$$

An implication is that convolutions with the $\mathbf{G}^j_\sigma$ are derivatives of Gaussian-smoothed functions, namely

$$\mathbf{G}^j_\sigma * f = \boldsymbol{\nabla}^j (G_\sigma * f),$$

where $f \in \mathcal{T}_0$. Note that we use the convention $\mathbf{G}^0_\sigma = G_\sigma = \frac{1}{(\sqrt{2\pi}\sigma)^3} e^{-\frac{r^2}{2\sigma^2}}$.

3 Harmonic Filters

Our goal is to build non-linear image filters that are equivariant to Euclidean motion. An $SE(3)$-equivariant image filter is given by the following

Definition 3 ($SE(3)$-Equivariant Image Filter)**.** *An image filter $\mathbf{F}$ is a mapping from $\mathcal{T}_{\ell_1}$ onto $\mathcal{T}_{\ell_2}$. We call such a mapping $SE(3)$-equivariant if $\mathbf{F}\{g\mathbf{f}\} = g\mathbf{F}\{\mathbf{f}\}$ for all $g \in SE(3)$ and $\mathbf{f} \in \mathcal{T}_{j_1}$.*

First, consider a scalar filter $H : \mathcal{T}_0 \to \mathcal{T}_0$ that takes a scalar field f as input and also gives a scalar-valued output.

Our approach may be interpreted as a kind of context-dependend voting scheme. The intuitive idea is as follows: Compute for each position in the 3D space the projection onto the Gaussian windowed harmonic basis $\mathbf{G}^j_\sigma$ for $j = 0, \ldots, n$. You can do this by a simple convolution of the image $\mathbf{f}$ with the basis, i.e.

$$\mathbf{p}^j := \mathbf{G}^j_\sigma * f.$$

Imagine this set of projections $\mathbf{p}^j$ as some local descriptor images, where the set of $[\mathbf{p}^0(\mathbf{r}), \mathbf{p}^1(\mathbf{r}), \ldots, \mathbf{p}^n(\mathbf{r})]$ of coefficients describe the harmonic part of the neighborhood of the voxel $\mathbf{r}$. Then, for each voxel map these projections on some new harmonic descriptors $\mathbf{V}^j(\mathbf{r}) = \mathbf{V}^j[\mathbf{p}^0(\mathbf{r}), \mathbf{p}^1(\mathbf{r}), \ldots, \mathbf{p}^n(\mathbf{r})]$ which can be interpreted as a local expansion of a kind of voting function that contributes into

the neighborhood of $\mathbf{r}$. The contribution stemming from the voter at voxel $\mathbf{r}'$ at position $\mathbf{r}$ is

$$V_{\mathbf{r}'}(\mathbf{r}) = G_\eta(\mathbf{r}-\mathbf{r}') \sum_{j=0}^{\infty} \mathbf{V}^j(\mathbf{r}') \bullet_0 \mathbf{R}^j(\mathbf{r}-\mathbf{r}'), \tag{6}$$

i.e. the voting function is just a Gaussian-windowed harmonic function. The final step is to render the contribution from all pixels $\mathbf{r}$ in an additive way together by integration to arrive at

$$H\{f\}(\mathbf{r}) \;:=\; \int_{\mathbb{R}^3} V_{\mathbf{r}'}(\mathbf{r}) d\mathbf{r}' = \sum_{j=0}^{n} \mathbf{G}_\eta^j \,\widetilde{\bullet}_0\, \mathbf{V}^j.$$

To ensure rotation-equivariance the $\mathbf{V}^j[\cdot]$ has to obey the following equivariance constraint:

Definition 4. *We call a mapping $\mathbf{V}^j$ of type*

$$\mathbf{V}^j : V_0 \times \ldots \times V_n \to V_j,$$

equivariant if it behaves as

$$\mathbf{V}^j[\mathbf{D}_g^0\mathbf{p}^0, \ldots, \mathbf{D}_g^n\mathbf{p}^n] = \mathbf{D}_g^j \mathbf{V}^j[\mathbf{p}^0, \ldots, \mathbf{p}^n].$$

We will use the spherical product $\bullet$ as the basic building block for the equivariant non-linearities $\mathbf{V}^j$. There are many possibility to combine several spherical tensors by the products $\bullet$ in an equivariant way. Later we will discuss this in detail.

Finally, we want to verify the $SE(3)$-equivariance of the filter. We can consider translation and rotation separately. First, examine the translation: if the image f is shifted, also the descriptor images $\mathbf{p}^j$ are shifted by the same amount, which comes due to the properties of the convolution. As $\mathbf{V}^j$ is working in a pointwise manner it is also obvious that $\mathbf{V}^j[\tau\mathbf{p}^0, \ldots, \tau\mathbf{p}^n] = \tau\mathbf{V}^j[\mathbf{p}^0, \ldots, \mathbf{p}^n]$. And finally, due to equation (1) we also have that $\mathbf{G}_\eta^j \widetilde{\bullet}_0(\tau\mathbf{V}^j) = \tau(\mathbf{G}_\eta^j \widetilde{\bullet}_0 \mathbf{V}^j)$ which shows the translation-equivariance. Now, consider the rotation: we know that the $\mathbf{G}_\sigma^j$ are fix points with respect to rotations, i.e. $g\mathbf{G}_\sigma^j = \mathbf{G}_\sigma^j$ for any rotation g. And so, due to Proposition 3 the descriptor fields $\mathbf{p}^j$ are spherical tensor fields, because $\mathbf{G}_\sigma^j * (gf) = g(\mathbf{G}_\sigma^j * f)$. Hence, also $\mathbf{V}^j[\mathbf{p}^0, \ldots, \mathbf{p}^n]$ is a spherical tensor field of order j. And again, due to the fix point property of $\mathbf{G}_\eta^j$ and Proposition 3 we know that $H\{f\} \in \mathcal{T}_0$ which proves the rotation-equivariance.

3.1 Differential Formulation

A computational expensive part of the filter are the convolutions. On the one hand, the projection onto the harmonic basis of the input and, secondly, the rendering of the output, also by convolution. Equation (6) shows that there is another way

to compute such projections: by the use of the spherical derivative. So, we can reformulate the filter as follows:

$$H\{f\} := G_\eta * \sum_{j=0}^{n} \boldsymbol{\nabla}_j \mathbf{V}^j[\boldsymbol{\nabla}^0 f_s, \ldots, \boldsymbol{\nabla}^n f_s] \tag{7}$$

with $f_s = G_\sigma * f$. In Algorithm 1 we depict the computation of the filter. Note, that we just have to compute n spherical derivatives $\boldsymbol{\nabla}^1$ if we implement them by repeated applications. And actually the same holds for the down-derivative $\boldsymbol{\nabla}_1$ if we follow Algorithm 1. In Figure 1 we illustrate the workflow of the filter.

Algorithm 1 Filter Algorithm $y = H\{f\}$

Input: scalar volume image f
Output: scalar volume imag y
1: Initialize $\mathbf{y}^n := 0 \in \mathcal{T}_n$
2: Convolve $\mathbf{p}^0 := G_\sigma * f$
3: **for** $j = 1 : n$ **do**
4: $\quad \mathbf{p}^j = \boldsymbol{\nabla}^1 \mathbf{p}^{j-1}$
5: **end for**
6: **for** $j = n : -1 : 1$ **do**
7: $\quad \mathbf{y}^{j-1} := \boldsymbol{\nabla}_1(\mathbf{y}^j + \mathbf{V}^j[\mathbf{p}^0, \ldots, \mathbf{p}^n])$
8: **end for**
9: Let $y := \mathbf{y}^0 + \mathbf{V}^0[\mathbf{p}^0, \ldots, \mathbf{p}^n])$
10: Convolve $y := G_\eta * y$

3.2 Generalization to Tensor Fields

The generalization of the filter to tensor-valued input and output is simple. In fact, there are many possibilities to do this. We want to point out one alternative. Assume, we want to construct a filter of type $\mathbf{H} : \mathcal{T}_{\ell_1} \to \mathcal{T}_{\ell_2}$. Starting with the differential formulation we can compute the descriptor images as before $\mathbf{p}^j = \boldsymbol{\nabla}^j(G_\sigma * \mathbf{f})$, where now $\mathbf{p}^j \in \mathcal{T}_{j+\ell_1}$. And now, we have to design local non-linearities $\mathbf{V}^j$ of type $\mathcal{T}_{0+\ell_1} \times \ldots \times \mathcal{T}_{n+\ell_1} \to \mathcal{T}_{j+\ell_2}$ and can use just the same formula in eqn. (7) as before for an assembly of the output.

3.3 The Harmonic Subspace

We proposed to consider just the harmonic part to compute the local descriptor images $\mathbf{p}^j$. One should ask what advantages and disadvantages follow from this restriction. Of course, the complexity of the algorithm is reduced if we just consider the harmonic part which is an advantage, but what kind of properties of the local neighborhood are neglected? Or, what subspace of functions is described

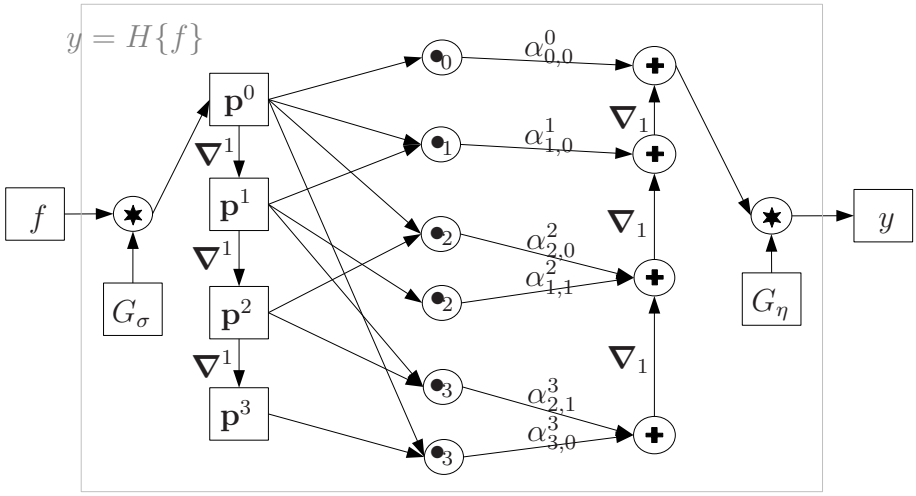

Figure 1: The workflow of a second order-filter ($N = 2$). The harmonic function is expanded up to a degree of $n = 3$. The star '*' in the circle is indicating a convolution of the two incoming images. The plus '$+$' indicate a addition of the incoming images, and the $\bullet_j$ a spherical multiplication of incoming channels. The labels at the arrows indicate a multiplication or spherical differentiation, respectively.

completely by its harmonic part. Of course, the harmonic functions are such a class. But there are also analytical functions which can form such a subspace, namely, those functions whose analytical part has a deterministic dependency on the harmonic part.

Let us be more precise. We denote the analytical local descriptor images by $\mathbf{p}^j_i = \boldsymbol{\nabla}^j_i f_s$ and by $\mathbf{p}^j_0 = \mathbf{p}^j$ the harmonic ones. As we are considering rotation-invariant problems the subspace of functions should be closed under rotations, i.e. the functional dependency has to be equivariant with respect to rotations. Let us consider a linear functional dependency between the harmonic and disharmonic part. One can imagine that the only rotation-equivariant linear dependency is described by

$$\mathbf{p}^j_i = \alpha^j_i \mathbf{p}^{j-i} \quad \text{with} \quad i > 0,$$

where α^j_i are some fixed real coefficients. We know that the spherical derivative $\boldsymbol{\nabla}^j_i$ can be written as $\boldsymbol{\nabla}^{j-i}\Delta^i$ and hence $\mathbf{p}^j_i = \boldsymbol{\nabla}^{n-i}\Delta^i f_s = \Delta^i \boldsymbol{\nabla}^{n-i} f_s = \Delta^i \mathbf{p}^{j-i}_0$. So, a simple guess for a function class which fulfills the above linear constraint are those functions which are eigenfunctions of the Laplace operator with respect to the same eigenvalue. The plane waves $e^{\mathbf{i}\mathbf{k}^\top \mathbf{r}}$ with fixed frequency magnitude $k = \|\mathbf{k}\|$ are these functions, that is $\alpha^j_i = k^{2i}$. These subspaces are known to be the irreducible subspaces of the motion group $SE(3)$, that is, there are no smaller subspaces that are closed under the action of the Euclidean motion. More practically, functions that have locally a dominating wave length in its Fourier domain are well described by the harmonic part. The functions look like a sea of

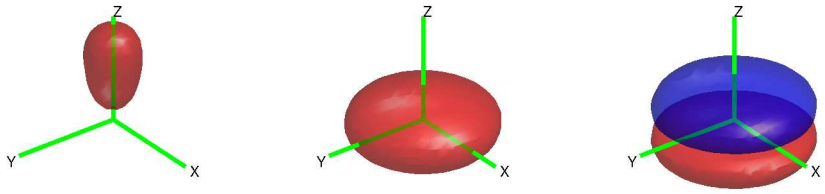

Figure 2: Isosurface Plots of the harmonic kernel in z-direction $\mathbf{r}_z$. Left: $f(\mathbf{r}) = K(\mathbf{r}, \mathbf{r}_z)$. Middle: $f(\mathbf{r}) = \Re(K(\mathbf{r}, \mathbf{i}\mathbf{r}_z))$. Right: $f(\mathbf{r}) = \Im(K(\mathbf{r}, \mathbf{i}\mathbf{r}_z))$.

bubbles of approximately the same size. The size is determined by the magnitude k of the wave number. For physicists: they are the steady state solutions of the Schrödinger equation of the free particle.

On the other hand we suggested to use harmonic functions to model locally the output, i.e. the voting function. Here also the question arises how limiting this fact is. To understand this more deeply we want consider a different basis. Recall the expansion as given in eq. (6). What kind of function we can model with this type of basis? From kernel theory we know that the reproducing kernel is a kind of basic building block, in our case, for the harmonic function space. Following Burbea [19] the reproducing kernel can be easily computed in terms of an orthogonal basis. Given the square norm $\langle G^j_{\sigma,m}, G^j_{\sigma,m} \rangle = \frac{(2j-1)!!}{(\sqrt{2\pi}\sigma)^3}$ the reproducing kernel of the harmonic subspace is

$$K_\sigma(\mathbf{r}, \mathbf{r}') = \sum_{j=0}^{\infty} \frac{(\sqrt{2\pi}\sigma)^3}{(2j-1)!!} \mathbf{G}^j_\sigma(\mathbf{r})^\top \mathbf{G}^j_\sigma(\mathbf{r}'). \tag{8}$$

It fulfills the reproducing property $f(\mathbf{r}) = \int K(\mathbf{r}, \mathbf{r}') f(\mathbf{r}') d\mathbf{r}'$ for any function f within the harmonics and hence provides an orthogonal projection onto this space. In Figure 2 we show a surface plot of the harmonic kernel for fixed evaluation in z-direction. One can see that the real harmonic kernel as given in eq. (8) is a bulb like structure oriented in z-direction. The bulb gets more elongated and moves away from the origin if the magnitude of $\mathbf{r}_z$ is increased. With linear combinations of this bulb-like functions one can model any harmonic function. As mentioned above (Proposition 6) the Gaussian-windowed harmonics are eigenfunctions of the Fourier transform for $\sigma = 1$, hence it is easy to compute the projection of the plane wave onto the harmonics, namely,

$$K_1(\mathbf{r}, \mathbf{i}\mathbf{k}) = \sqrt{2\pi}^3 \int_{\mathbb{R}^3} K_1(\mathbf{r}, \mathbf{r}') e^{\mathbf{i}\mathbf{k}^\top \mathbf{r}'} d\mathbf{r}',$$

i.e. the harmonic projection of a Gaussian-windowed plane wave is just the imaginary evaluation $K(\mathbf{r}, \mathbf{i}\mathbf{k})$ of the reproducing kernel. In Figure 2 we also show the

real and imaginary part of this complex kernel. It gives a kind of wave of just one period evolving in z-direction. The magnitude of $\mathbf{r}_z$ determines the wavelength.

4 An Introductory Example

We already explained the workflow of the algorithm roughly. The question of how to compute the spherical derivatives was already answered in [3] by a finite difference scheme. In the following introductory experiment we use a central difference scheme. The implementation of the spherical product is straightforward and follows strictly definition 2. Note, that we just have to store $2j+1$ real numbers for spherical tensors of degree j and use the relation $v_{-m} = (-1)^m \overline{v_m}$ to compute the product. The Clebsch-Gordan coefficients can be precomputed and do not cost any computation time. The pre- and post-smoothing with the Gaussian is performed with the aid of an FFT. We implemented our experiments in *C++* and used the *mex*-interface of *MATLAB* as a frontend.

To get a first understanding how the filter actually works we consider two simple examples. First we start with a linear one. Suppose the voting function $\mathbf{V}$ is linear in its arguments, i.e. $\mathbf{V}^j[\mathbf{p}^0, \ldots, \mathbf{p}^n] = \alpha^j \mathbf{p}^j$ with some $\alpha_j \in \mathbb{R}$. This is the only rotation-equivariant linearity. Then we can compute

$$\begin{aligned} H\{f\} &= G_\eta * \sum_{j=0}^{n} \alpha^j \boldsymbol{\nabla}_j \boldsymbol{\nabla}^j (G_\sigma * f) \\ &= \sum_{j=0}^{n} \alpha^j \Delta^j (G_{\sigma'} * f) \end{aligned}$$

with $\sigma' = \sqrt{\sigma^2 + \eta^2}$. This means that the harmonic filter can model any linear filter with a spectrum of the form

$$\sum_{j=0}^{n} \alpha^j k^{2j} e^{-\frac{(\sigma' k)^2}{2}},$$

where $k = \|\mathbf{k}\|$ is the frequency magnitude. In fact, any rotationally symmetric spectrum can be modelled with such a kind of filter. Or more generally, any rotation-equivariant linear filter is a Harmonic filter. On the one hand we can be really happy with this fact, we know that in the linear case it does not harm the expressivity of the filter if we just consider the harmonic part. But the practical benefit of this fact is limited. Computations of such filters via FFT convolutions are in most cases more practically. But at least the parametric form of the Harmonic filter offers a simple way to adapt the filter coefficients α_j via regression to a specific problem.

Secondly, we consider an image reconstruction problem, edge preserving denoising. The idea is as follows: First, we take the original image and apply a

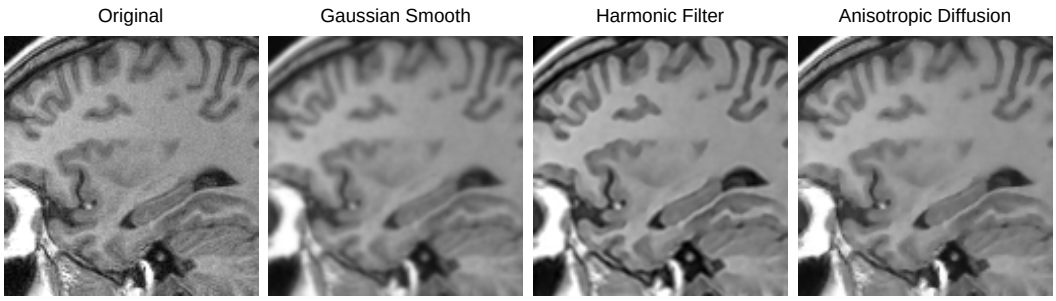

Figure 3: Results for edge preserving smoothing.

classical isotropic smoothing with a Gaussian. Then, based on this image we compute a kind of edge indicator image, providing information about intensity and orientation of the edges in the volume, e.g. just the gradient image. This image is then used for edge enhancement as follows: if an edge is indicated the imaginary part of the harmonic kernel (shown in Figure 2) is superimposed to the local neighborhood, where the harmonic kernel is steered along the indicated edge direction. This approach enhances the edge like classical edge enhancement techniques known from 1D, but here, extended to 3D, the orientation of the kernel is locally adapted (steered) depending on an edge indicator image.

Now, let us translate this approach into terms of the harmonic filter and its voting function. The imaginary part of the harmonic kernel has only odd summands, so the $\mathbf{V}^j$ give only contribution for odd j. The amplitude of the harmonic kernel should be proportional to the edge strength which is encoded by $\|\mathbf{p}^1\|$. For $j = 0$ we just transfer the original blurred image encoded in $\mathbf{p}^0$. Thus, we have for the voting function,

$$\begin{aligned}
\mathbf{V}^0 &= \mathbf{p}^0 \\
\mathbf{V}^j &= 0 \text{ for even } j > 0 \\
\mathbf{V}^j &= \alpha\|\mathbf{p}^1\|\frac{(-1)^{\lfloor j/2 \rfloor}}{(2j-1)!!}\mathbf{Y}^j(\mathbf{n}) \text{ for odd } j > 0
\end{aligned}$$

where $\mathbf{n} : \mathbb{R}^3 \mapsto S_2$ is the gradient direction image and α the edge indicator image. To compute α we follow Weickert [20]. He used $\alpha = \gamma \exp(-\frac{C}{(\|\mathbf{p}^1\|/\lambda)^4})$ to guide his nonlinear diffusion process (for $\gamma = 1$). The parameter $\gamma \in \mathbb{R}^+$ controls the strength of the sharpening effect of the filter.

Similar as in [20] we have chosen $C = 3.3$ and for $\lambda = 0.1$. In Figure 3 we show some results of the filter applied on MRI-data of the human brain and comparison to anisotropic nonlinear diffusion (following [20] Ch.5. pages 96-100), both with the same α. We also show the original image and a Gaussian smoothed version (the harmonic filter with $\gamma = 0$). It is difficult to make a quantitive analysis of the differences. For the human eye the result of the harmonic filter seems to be a little bit sharper due to the over- and undershoots which are inherent with such an approach. But actually the anisotropic diffusion filter shows sharper edges

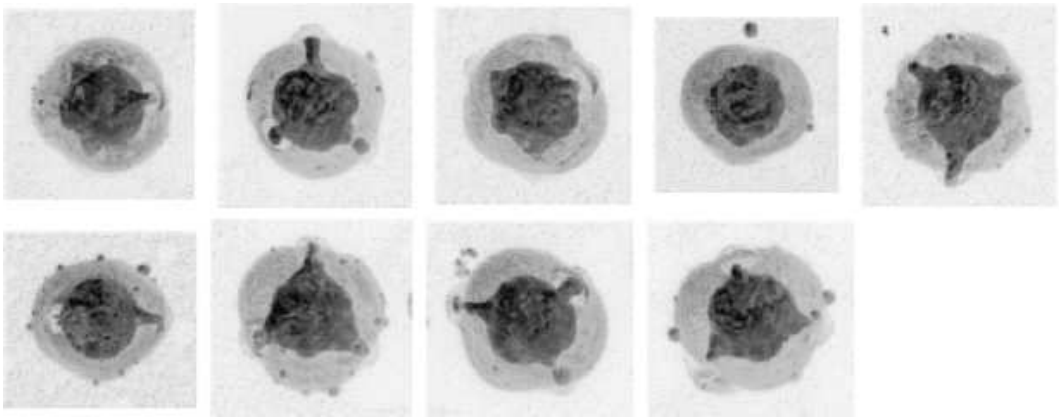

Figure 4: Examples of confocally scanned pollen data. The goal is to detect the three little porates on the 'surface' of the pollen.

with higher details.

5 Pollen Porate Detection in Confocal Data

Analysis techniques for data acquired by microscopy typically demand for a rotation and translation invariant treatment. In this experiment we use the harmonic filter for the analysis of pollen grains acquired with confocal laser scanning microscopy (see [21]). Palynology, the study and analysis of pollen, is an interesting topic with very diverse applications like in Paleoclimatology or Forensics. An important feature of certain types of pollen grain are the so called porates that are small pores on the surface of the grain. Their relative configuration is crucial for the determination of the species. We want to show that our filter is able to detect this structures in a reliable way. In Figure 4 we show volume renderings of some examples. The dataset consists of 45 samples. The images have varying sizes. We scaled the images to a size of about 80^3 voxels. We labeled the porates by hand. The experimental setup is quite simple. We apply on each pollen image the trained harmonic filter and then select local maxima up to a certain threshold as detection hypotheses.

5.1 Reference Approaches

We use the ideas of Ballard et al [10], Lowe et al [11] and Leibe et al [12, 22] and extended them to 3D. The approach is based on the generalized Hough transform (GHT). Based on a selection of interest points local features are extracted and assigned to a codebook entry. Each codebook entry is endowed with a set of votes for the center of object which are casted for each interest point. This approach resembles closely the idea of the implicit shape model by Leibe et al [12], where we used a 3D extension of Lowe's SIFT features [11] as local features. We used

several interest point detectors: Difference of Gaussians [11] (DOG), determinant of Hessian (DHES) and the Harris detector [23]. The interest points were always used in a non-scale invariant manner. This had basically two reasons, we focus on biological problems that demand scale robustness rather than invariance and, secondly, the number of scale-invariant keypoints was for our problems just too low, resulting in very poor results. The rotation invariance for the feature extraction process and the voting procedure is obtained by normalization with respect to the eigensystem of the Hessian or the structure tensor (for the Harris detector). The features are joint gradient direction, position histograms relative to the main eigenvector $\mathbf{n}$ of the Hessian, i.e. the histogram dimensions are associated with $\mathbf{n}^\top \nabla f(\mathbf{r})$ and $\mathbf{r}^\top \mathbf{n}$ and $\sqrt{\|\mathbf{r}\|^2 - (\mathbf{r}^\top \mathbf{n})^2}$. As the main direction of the Hessian/structure tensor is only determined up to a axial flip we normalized $\mathbf{n}$ for each $\mathbf{r}$ such that $\mathbf{r}^\top \mathbf{n}$ is positive. This retains more information than just taking the absolute of the three quantities. As the eigensystem is determined up to eight axial flips we have to cast each vote eight times to reach full invariance against the $O(3)$. After the generation of the voting map the map is blurred and local maximas are used as detection hypotheses.

As a second approach we apply a simple classification scheme per voxel (VC). For each voxel we compute a set of expressive rotation invariant features and train a classifier to discriminate the objects of interest from the background. This idea was for example used by Staal et al [24] for blood vessel detection in retinal images in 2D or by Fehr et al [25, 26] for cell detection in 3D. We use features that are based on a local spherical harmonics expansion of the image. The norms of the spherical harmonic expansion serve as an invariant feature. For the radial part we follow Koenderink et al [27] and use a Gaussian-windowed Laguerre expansion, i.e. the local basis can be written up to a normalization constant as $\mathbf{b}_i^\ell(\mathbf{r}) \propto \mathcal{L}_i^{\ell+1/2}(r^2)\mathbf{R}^\ell(\mathbf{r})e^{-r^2}$, where $\mathcal{L}_i^{\ell+1/2}$ denotes the associated Laguerre polynomials. The $\mathbf{b}_i^\ell(\mathbf{r})$ solve the differential equations connected with the 3D harmonic oscillator in quantum mechanics and are optimal for smooth and local white processes (see [2]). They are the eigenfunctions of the corresponding Hamilton operator, hence they provide an orthogonal and complete set of functions. Convolved with the image and taken the norm the $\|\mathbf{b}_i^\ell * f\|$ deliver rotation invariant features for each voxel. We compute them for different scales, i.e. different size of Gaussians and up to expansion degrees of order $\ell + i \leq 6$, resulting in 80 features (per voxel!). For classification we use a KNN classifier as in [24]. We considered also to use a SVM together with a Histogram-Intersection kernel; HI-kernels allow a fast implementation.

5.2 The Voting Function

The probably most simple generic nonlinear voting function $\mathbf{V}^j$ is a sum of second order products of the descriptor images $\mathbf{p}^j$, namely

$$\mathbf{V}^j[\mathbf{p}^0,\ldots,\mathbf{p}^n] = \sum_{\substack{|j_1-j_2|\le j\le j_1+j_2 \\ j_1+j_2+j \text{ even} \\ j_1,j_2\le n}} \alpha^j_{j_1,j_2}\ \mathbf{p}^{j_1} \bullet_j \mathbf{p}^{j_2} \tag{9}$$

where $\alpha^j_{j_1,j_2} \in \mathbb{R}$ are expansion coefficients. We call the order of the products that are involved in $\mathbf{V}^j$ the order of the filter and denote it by N. Depending on the application they may or may not depend on the absolute intensity values of the input image. To become invariant against additive intensity changes one leaves out the zero order descriptor $\mathbf{p}^0$. For robustness against illumination/contrast changes we introduce a soft normalization of the first order ('gradient') descriptor $\mathbf{p}^1$. This means, that in the for-loop in Alg. 1 from line 3-5 we introduce a special case for $j = 1$, namely

$$\mathbf{p}^1(\mathbf{r}) = \frac{1}{\gamma + s_{\text{dev}}(\mathbf{r})}\boldsymbol{\nabla}^1 f(\mathbf{r}),$$

where $\gamma \in \mathbb{R}$ is a fixed regularization parameter and $s_{\text{dev}}(\mathbf{r})$ denotes the standard deviation computed in a local window around $\mathbf{r}$. The normalization makes the filter robust against multiplicative changes of the gray values and, secondly, emphasizes the 'structural' and 'textural' properties rather than the pure intensities. Besides γ, the filter has three other parameters: the expansion degree n, the width of the input Gaussian σ and the output Gaussian η. In the spirit of the GHT, the parameter σ determines the size of the local features that vote for the center of the object of interest. To assure that every pixel of the object can contribute, the extent of the voting function should be at least half the diameter of the object.

5.3 Training

For the training of the harmonic filter (and for all reference approaches) we selected one(!) good pollen example, i.e. three porate samples. To train the harmonic filter we built an indicator image with pixels set to 1 at the centers of the three porates. The indicator image is just the target image y which should satisfy $H\{f\} = y$. As mentioned before the linearity of the filter in its parameters makes it easy to adapt them. We use an unregularized least square approach. Due to the high dynamic differences between the filter responses corresponding to the individual parameters it is necessary to normalize the equation to avoid numerical problems. We used the standard deviation of the individual filter responses taken over all samples in the training image. The σ parameter determining the size of the local features was chosen to be 2.5 pixels. The output width η determining the range of voting function was chosen to be 8 pixels, this is about half the diameter of the porates.

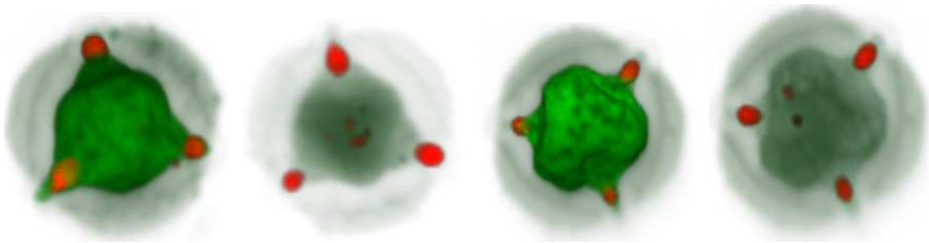

Figure 5: Mugwort pollen (green) with overlayed filter response (red) for two examples. The filter detects the three porates, but there are also some spurious responses within the pollen, because the pollen has also strong inner structures.

The training of the GHT based approach is straightforward. For the codebook generation we used k-means clustering. We found about 65 clusters most beneficial for the application. For the voxel-based classification the training was more difficult. We have chosen a 6-neighborhood around the true labels as positive training samples and a randomly chosen set of non-positive pixels as negative samples. We choose $K = 20$ for the KNN-classifier. With the number of negative samples one can control the a-priori probability of the classifier. We found a rate of $1/5$ most suitable. The SVM approach (we use a C-SVM) has more problems with such kind of unbalanced training sets. We also experienced a high dependency on the randomly generated negative samples. To become less dependend of the sampling procedure we increased the number of negative samples and adapted the class-specific C parameter to control the a-priori probability.

5.4 Evaluation

In Figure 5 we show two examples. The filter detects the porates but shows also some small responses within the pollen, however the results are still acceptable. For quantative results we computed Precision/Recall graphs. A detection was found to be successful if it is at least 8(4) pixels away from the true label. In Figure 6 on the left we show a PR-graph for a varying expansion degree n with a low detection precision of 8 pixels. As one expects the filter improves its performance with growing n. For $n = 8$ no performance gain is observed. The runtime of the filter heavily depends on the number of spherical products to be computed. For example for $n = 6$ we have to compute 46 products. The computation of these products takes on a *P4 2.8Ghz* about 6 seconds. In Figure 6 in the middle we compare the result of the Harmonic filter with $n = 7$ with the reference approaches. The results of the GHT based on DOG interest points are comparable with the Harmonic filter. The voxel classification approach (VC) performs not so well. In particular, for the SVM based classification is performing quite poorly. Finally, we evaluated the PR-graph with a higher detection precision of 4 pixels. As already experienced in [1] the GHT based approach has problems in this case, which has probably to do with the inaccurate and unstable determination of the

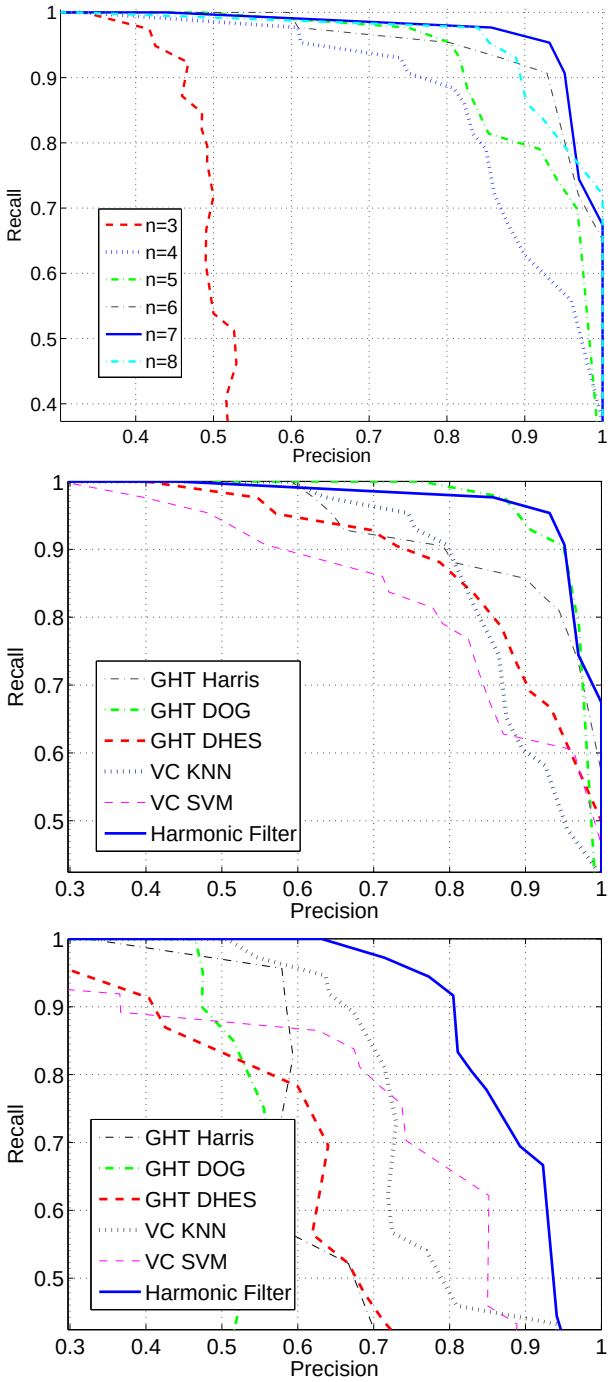

Figure 6: Precision/Recall graphs of the porate detection problem. Top: Comparison of the Harmonic filter for different expansion degrees (precision 8 pixels). Middle: Comparison with reference approaches (precision 8 pixels). Bottom: Comparison with reference approaches (4 pixels).

interest points. Now both VC approaches are outperforming the GHT approaches while the Harmonic Filter is definitely superior over all the others.

6 Conclusion

In this paper we presented a general-purpose non-linear filter that is equivariant with respect to the 3D Euclidean motion. The filter may be seen as a joint Volterra filter for rotation and translation. The filter senses locally a harmonic projection of the image function and maps this projection onto a kind of voting function which is also harmonic. The mapping is modelled by rotation equivariant polynomials in the describing coefficients. The harmonic projections are computed in an efficient manner by the use of spherical derivatives of Gaussian-smoothed images. We applied the filter on a 3D detection problem. For low detection precision the performance is comparable to state of the art approaches, while for high detection precision the approach is definitely outperforming existing approaches.

A Spherical Harmonics

We always use Racah-normalized spherical harmonics. In terms of Legendre polynomials they are written as

$$Y_m^\ell(\phi,\theta) = \sqrt{\frac{(l-m)!}{(l+m)!}} P_m^\ell(\cos(\theta)) e^{\mathbf{i}\phi}$$

Mostly we write $\mathbf{r}/r \in S^2$ instead of (ϕ,θ). The Racah-normalized solid harmonics can be written as

$$R_m^\ell(\mathbf{r}) = \sqrt{(\ell+m)!(\ell-m)!} \sum_{i,j,k} \frac{\delta_{i+j+k,\ell}\delta_{i-j,m}}{i!j!k!2^i 2^j} (x-\mathbf{i}y)^j (-x-\mathbf{i}y)^i z^k,$$

where $\mathbf{r} = (x,y,z)$. They are related to spherical harmonics by $R_m^\ell(\mathbf{r})/r^\ell = Y_m^\ell(\mathbf{r}/r)$

B Clebsch Gordan Coefficients

Orthogonality

$$\sum_{j,m} \langle jm|j_1m_1, j_2m_2\rangle \langle jm|j_1m_1', j_2m_2'\rangle = \delta_{m_1,m_1'}\delta_{m_2,m_2'} \quad (10)$$

$$\sum_{m=m_1+m_2} \langle jm|j_1m_1, j_2m_2\rangle \langle j'm'|j_1m_1, j_2m_2\rangle = \delta_{j,j'}\delta_{m,m'} \quad (11)$$

$$\sum_{m_1,m} \langle jm|j_1m_1, j_2m_2\rangle \langle jm|j_1m_1, j_2'm_2'\rangle = \frac{2j+1}{2j_2'+1}\delta_{j_2,j_2'}\delta_{m_2,m_2'} \quad (12)$$

Special Values

$$\langle \ell m | (\ell-\lambda)(m-\mu), \lambda\mu \rangle = \begin{pmatrix} \ell+m \\ \lambda+\mu \end{pmatrix}^{1/2} \begin{pmatrix} \ell-m \\ \lambda-\mu \end{pmatrix}^{1/2} \begin{pmatrix} 2\ell \\ 2\lambda \end{pmatrix}^{-1/2} \tag{13}$$

$$\begin{aligned} \langle \ell m | (\ell+\lambda)(m-\mu), \lambda\mu \rangle = (-1)^{\lambda+\mu} & \begin{pmatrix} \ell+\lambda-m+\mu \\ \lambda+\mu \end{pmatrix}^{1/2} \\ & \begin{pmatrix} \ell+\lambda+m-\mu \\ \lambda-\mu \end{pmatrix}^{1/2} \begin{pmatrix} 2\ell+2\lambda+1 \\ 2\lambda \end{pmatrix}^{-1/2} \end{aligned} \tag{14}$$

Symmetry

$$\langle jm | j_1 m_1, j_2 m_2 \rangle = \langle j_1 m_1, j_2 m_2 | jm \rangle \tag{15}$$

$$\langle jm | j_1 m_1, j_2 m_2 \rangle = (-1)^{j+j_1+j_2} \langle jm | j_2 m_2, j_1 m_1 \rangle \tag{16}$$

$$\langle jm | j_1 m_1, j_2 m_2 \rangle = (-1)^{j+j_1+j_2} \langle j(-m) | j_1(-m_1), j_2(-m_2) \rangle \tag{17}$$

C Wigner D-Matrix

The components of $\mathbf{D}_g^\ell$ are written D_{mn}^ℓ. They are called the Wigner D-matrix. In Euler angles ϕ, θ, ψ in ZYZ-convention we have

$$D_{mn}^\ell(\phi, \theta, \psi) = e^{\mathbf{i}m\phi} d_{mn}^\ell(\theta) e^{\mathbf{i}n\psi},$$

where $d_{mn}^\ell(\theta)$ are the Wigner d-matrix which is real-valued. Relation to the Clebsch Gordan coefficients:

$$D_{mn}^\ell = \sum_{\substack{m_1+m_2=m \\ n_1+n_2=n}} D_{m_1 n_1}^{\ell_1} D_{m_2 n_2}^{\ell_2} \langle lm | l_1 m_1, l_2 m_2 \rangle \langle ln | l_1 n_1, l_2 n_2 \rangle \tag{18}$$

$$D_{m_1 n_1}^{\ell_1} D_{m_2 n_2}^{\ell_2} = \sum_{l,m,n} D_{mn}^\ell \langle lm | l_1 m_1, l_2 m_2 \rangle \langle ln | l_1 n_1, l_2 n_2 \rangle \tag{19}$$

C.1 Double Factorial

$$(2\ell+1)!! = \Gamma(\ell+3/2) \frac{2^l}{\sqrt{\pi/4}} = (2\ell+1)(2\ell-1)(2\ell-3)\dots$$

References

[1] M. Reisert and H. Burkhardt, "Equivariant holomorphic filters for contour denoising and rapid object detection," *IEEE Trans. Image Processing*, vol. 17, no. 2, 2008.

[2] ——, "Complex derivative filters," *IEEE Trans. Image Processing*, vol. 17, no. 12, 2008.

[3] ——, "Spherical tensor calculus for local adaptive filtering," *Tensors in Image Processing and Computer Vision (to appear)*.

[4] S. Thurnhofer and S. Mitra, "A general framework for quadratic volterra filters for edge enhancment," *IEEE Trans. Image Processing*, pp. 950–963, 1996.

[5] V. J. Mathews and G. Sicuranza, *Polynomial Signal Processing*. J.Wiley, New York, 2000.

[6] J. August, "Volterra filtering of noisy images of curves," in *Proceedings of the ECCV*, A. Heyden, G. Sparr, M. Nielsen, and P. Johansen, Eds. LNCS, Springer, 2002.

[7] W. T. Freeman and E. H. Adelson, "The design and use of steerable filters," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 13, no. 9, pp. 891–906, 1991.

[8] P. Perona, "Deformable kernels for early vision," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 17, no. 5, pp. 488 – 499, 1995.

[9] M. Jacob and M. Unser, "Design of steerable filters for feature detection using canny-like criteria," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 26, no. 82, pp. 1007 – 1019, 2004.

[10] D. Ballard, "Generalizing the hough transform to detect arbitrary shapes," *Pattern Recognition*, vol. 13, no. 2, pp. 111–122, 1981.

[11] D. Lowe, "Distinct image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, pp. 91–110, 2004.

[12] B. Leibe, A. Leonardis, and B. Schiele, "Combined object categorization and segmentation with an implicit shape model," in *Proceedings of the ECCV'04 Workshop on Statistical Learning in Computer Vision, Prague*, T. Pajdla and J. Matas, Eds. LNCS, Springer, 2004.

[13] P. Mordohai, *Tensor Voting: A Perceptual Organization Approach to Computer Vision and Machine Learning*. Morgan and Claypool, ISBN-10: 1598291009, 2006.

[14] E. Franken, M. van Almsick, P. Rongen, L. Florack, and B. ter Haar Romeny, "An efficient method for tensor voting using steerable filters," in *Proceedings of the ECCV 2006*. Lecture Notes in Computer Science, Springer, 2006, pp. 228–240.

[15] P. Wormer, "Angular momentum theory," *Lecture Notes - University of Nijmegen Toernooiveld, 6525 ED Nijmegen, The Netherlands*. [Online]. Available: www.theochem.kun.nl/ pwormer/teachmat.html

[16] M. Rose, *Elementary Theory of Angular Momentum*. Dover Publications, 1995.

[17] W. Miller, R. Blahut, and C. Wilcox, "Topics in harmonic analysis with applications to radar and sonar," *IMA Volumes in Mathematics and its Applications, Springer-Verlag, New York*, 1991.

[18] R. Lenz, *Group theoretical methods in Image Processing*. Springer Verlag, Lecture Notes, 1990.

[19] J. Burbea, "Total positivity of certain reproducing kernels," *Pacific Journal of Mathematics*, vol. 67, no. 1, 1976.

[20] J. Weickert, "Anisotropic diffusion in image processing," Ph.D. dissertation, Universitt Kaiserslautern, Jan. 1996.

[21] O. Ronneberger, H. Burkhardt, and E. Schultz, "General-purpose Object Recognition in 3D Volume Data Sets using Gray-Scale Invariants," in *Proceedings of the International Conference on Pattern Recognition*. Quebec, Canada: IEEE Computer Society, Sept. 2002.

[22] K. Mikolajczyk, B. Leibe, and B. Schiele, "Multiple object class detection with a generative model," in *Proceedings of the CVPR*, vol. 1. IEEE Computer Society, 2006, pp. 26 – 36.

[23] C. Harris and M. Stephens, "A combined corner and edge detector," in *Proceedings of the Alvey Vision Conference*, 1988, pp. 147–152.

[24] J. Staal, B. Ginneken, M. Niemeijer, A. Viegever, and M. Abramoff, "Ridge based vessel segmentation in color images of the retina," *IEEE Trans. Med. Imaging*, vol. 23, no. 4, pp. 501–509, 2004.

[25] J. Fehr, O. Ronneberger, H. Kurz, and H. Burkhardt, "Self-learning segmentation and classification of cell-nuclei in 3d volumetric data using voxel-wise gray scale invariants," in *Pattern Recognition - Proc. of the 27th DAGM Symposium, Vienna, Austria*, W. Kropatsch and R. Sablating, Eds. LNCS, Springer, 2005.

[26] O. Ronneberger, J. Fehr, and H. Burkhardt, "Voxel-wise gray scale invariants for simultaneous segmentation and classification," in *In Proceedings of the 27th DAGM Symposium, Vienna, Austria*, W. Kropatsch, R. Sablatnig, and A. Hanbury, Eds. LNCS, Springer, 2005.

[27] J. Koenderink and A. van Doorn, “Generic neighborhood operators,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 14, pp. 597–605, 1992.