# Harmonic Filters for 3D Multichannel Data: Rotation Invariant Detection of Mitoses in Colorectal Cancer

Matthias Schlachter*, Marco Reisert, Corinna Herz, Fabienne Schlürmann, Silke Lassmann, Martin Werner, Hans Burkhardt, and Olaf Ronneberger

***Abstract*—In this paper, we present a novel approach for a trainable rotation invariant detection of complex structures in 3D microscopic multichannel data using a nonlinear filter approach. The basic idea of our approach is to compute local features in a window around each 3D position and map these features by means of a nonlinear mapping onto new local harmonic descriptors of the local window. These local harmonic descriptors are then combined in a linear way to form the output of the filter. The optimal combination of the computed local harmonic descriptors is determined in previous training step, and allows the filter to be adapted to an arbitrary structure depending on the problem at hand. Our approach is not limited to scalar-valued images and can also be used for vector-valued (multichannel) images such as gradient vector flow fields. We present realizations of a scalar-valued and a vector-valued multichannel filter. Our proposed algorithm was quantitatively evaluated on colorectal cancer cell lines (cells grown under controlled conditions), on which we successfully detected complex 3D mitotic structures. For a qualitative evaluation we tested our algorithms on human 3D tissue samples of colorectal cancer. We compare our results with a steerable filter approach as well as a morphology-based approach.**

***Index Terms*—Equivariant nonlinear image filter, rotation invariant detection, three-dimensional (3D) microscopic multichannel data, tissue, trainable.**

## I. Introduction

IN BIOMEDICAL research quantitative analysis of cells in their natural 3D surrounding becomes more and more important. Specific fluorescence staining of certain structures and 3D microscopic imaging techniques are well established, resulting in a large number of 3D multichannel data sets. In contrast, the automated analysis of such data sets is still in its infancy. In this paper, we address a central task for such analyzes, the detection of complex 3D multicolored structures.

Window-based detectors combined with machine learning approaches like the Viola–Jones detector [8] are very popular in image analysis but have the disadvantage of not being invariant or equivariant to rotations of the object of interest. However rotation invariance is required in many applications. Popular approaches addressing this issue are often based on the principle of template matching and incorporate rotations by means of steering the template. Most of them are low-level feature detectors, e.g., steerable filters, for lines and corners [1] or surfaces [2] which are common in early vision and image analysis. Some approaches have the ability to learn the objects of interest using numerical or least square approaches [3]–[5] in a rotation invariant manner. Many approaches are model based such as the Hough transform [6]. For multichannel images most algorithms are only addressed to 2D problems [7].

Our novel approach is based on the extension of the 3D harmonic filter framework [4], [5] to multichannel data by utilizing spherical tensor analysis [9]. Therefore our approach can be seen as an approach between a classical low-level feature detector (e.g., steerable filter) and a high-level object detector based on the generalized Hough transform [10].

The advantages of this framework are its high flexibility due to the nonlinear filter approach, the straightforward adaptation to new structures due to its learning ability, and its speed, as it does not rely on large numbers of expensive convolutions like steerable filter approaches.

We validate our approach on a well characterized cellular 3D structure, the metaphase of dividing, mitotic cells. In this specific (sub-)phase of mitosis, the sister chromatids of the cell have been arranged at a central position of the cell, and are attached via spindle proteins to two opposite poles, the so called centrosomal complex [11]. The 3D structure of mitotic cells in metaphase encompasses well defined cellular objects (i.e., the central structure of DNA and the bi-directionally located poles, or multiple poles in aberrant mitosis). Therefore, this represents an ideal system to validate our novel filter approach by detecting complex structures of mitotic cells in 3D multichannel data sets.

We compared our results with various reference approaches, such as a steerable filter approach, mathematical morphology based approach and the original single-channel harmonic filter approach.

Manuscript received February 21, 2010; revised April 21, 2010; accepted April 26, 2010. Date of publication June 07, 2010; date of current version August 04, 2010. This study was supported by the Excellence Initiative of the German Federal and State Governments (EXC 294). Parts of the study were supported by a grant to SL and MW of the German Research Foundation (DFG, grant #LA 1290/3-1).*Asterisk indicates corresponding author.*

*M. Schlachter is with the Computer Science Department, University of Freiburg, 79110 Freiburg, Germany.

M. Reisert is with the Department of Diagnostic Radiology, Medical Physics, University Medical Center of Freiburg, 79106 Freiburg, Germany.

C. Herz, F. Schlürmann, S. Lassmann, and M. Werner are with the Institute of Pathology, University Medical Center of Freiburg, 79106 Freiburg, Germany.

H. Burkhardt and O. Ronneberger are with the Computer Science Department and the Centre for Biological Signaling Studies (bioss), University of Freiburg, 79110 Freiburg, Germany.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TMI.2010.2049856

## II. Preliminaries

### A. Spherical Tensor Analysis

This section gives a brief summary of the theoretical background needed to understand our approach. Details and proofs can be found in [4] and [9].

Let $\mathbf{D}_g^j$ be the unitary irreducible representation of a $g \in SO(3)$ (rotation group in three dimensions) of order $j$ with $j \in \mathbb{N}$. They are also known as the *Wigner D-matrices* (see e.g., [12]). The representation $\mathbf{D}_g^j$ acts on the 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 $m = -j, \dots j$. We treat $V_j$ as a real vector space of dimension $2j+1$, despite the fact that the components of $\mathbf{v}$ might be complex. This means that $V_j$ is only closed under weighted superpositions of real numbers. Therefore, components of $u \in V_j$ fulfill the property $\overline{u_m} = (-1)^m u_{-m}$. The corresponding orthogonal space is denoted as $\mathbf{i}V_j$. We write the elements $\mathbf{w} \in \mathbf{i}V_j$ as $\mathbf{w} = \mathbf{i}\mathbf{v}$ with $\mathbf{v} \in V_j$. Analogue to elements of $V_j$, the elements $\mathbf{w} \in \mathbf{i}V_j$ fulfill always $\overline{w_m} = (-1)^{m+1} w_{-m}$. Hence, $\mathbb{C}^{2j+1}$ can be written as direct sum of these two spaces $\mathbb{C}^{2j+1} = V_j \oplus \mathbf{i}V_j$. For the transposition of a vector/matrix we write $\mathbf{u}^T$; the joint complex conjugation and transposition is denoted by $\mathbf{u}^\top = \overline{\mathbf{u}}^T$. The relation of the standard coordinate vector $\mathbf{r} = (x, y, z)^T \in \mathbb{R}^3$ to the complex valued spherical coordinate $\mathbf{u} \in V_1$ is given by an unitary coordinate transformation

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

with $\mathbf{S} = (1/\sqrt{2}) \begin{pmatrix} 1 & -\mathbf{i} & 0 \\ 0 & 0 & \sqrt{2} \\ -1 & -\mathbf{i} & 0 \end{pmatrix}$.

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

$$(g\mathbf{f})(\mathbf{r}) := \mathbf{D}_g^j \mathbf{f}(\mathbf{U}_g^T \mathbf{r}) \,, \text{ for all } g \in SO(3) \,.$$

$\mathbf{U}_g \in \mathbb{R}^{3\times 3}$ denotes the real-valued rotation matrix representing $g$. The space of all spherical tensor fields of rank $j$ is denoted by $\mathcal{T}_j$.

*Definition 2 (Spherical Tensor Coupling):* For $j \geq 0$ we define a family of bilinear forms of type

$$\circ_j : V_{j_1} \times V_{j_2} \to \mathbb{C}^{2j+1}$$

where $j_1, j_2 \in \mathbb{N}$ have to be chosen according the triangle inequality

$$|j_1 - j_2| \leq j \leq j_1 + j_2 \,. \tag{1}$$

For $\mathbf{v} \in V_{j_1}$ and $\mathbf{w} \in V_{j_2}$, $\mathbf{v} \circ_j \mathbf{w}$ is defined by

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

where $\langle jm \mid j_1 m_1, j_2 m_2 \rangle$ are the Clebsch–Gordan coefficients. This is just the ordinary angular momentum coupling known in quantum mechanics (see [12] and [13]). Note, the Clebsch–Gordan coefficients $\langle jm \mid j_1 m_1, j_2 m_2 \rangle$ are only non zero for $m = m_1 + m_2$. In the following, we will present some other useful properties of the products defined by definition 2.

*Proposition 1:* Let $j_1, j_2, j \in \mathbb{N}$ be chosen according to the triangle inequality (1). If $j + j_1 + j_2$ is even, then $\circ_j$ is symmetric and antisymmetric otherwise. The space $V_j$ is closed under symmetric products but not for antisymmetric products. Then we have for $\mathbf{v} \in V_{j_1}$ and $\mathbf{w} \in V_{j_2}$

$$\begin{aligned} j + j_1 + j_2 \text{ is even} &\Rightarrow \mathbf{v} \circ_j \mathbf{w} \in V_j \\ j + j_1 + j_2 \text{ is odd} &\Rightarrow \mathbf{v} \circ_j \mathbf{w} \in \mathbf{i}V_j \,. \end{aligned}$$

For a proof of proposition 1 see [9] or [14]. The most important and useful property of the above defined products is that they respect rotations of their arguments. Let $j_1, j_2, j \in \mathbb{N}$ be chosen according to the triangle inequality (1). For any $\mathbf{v} \in V_{j_1}$ and $\mathbf{w} \in V_{j_2}$ and $g \in SO(3)$ holds

$$\left(\mathbf{D}_g^{j_1} \mathbf{v}\right) \circ_j \left(\mathbf{D}_g^{j_2} \mathbf{w}\right) = \mathbf{D}_g^j \left(\mathbf{v} \circ_j \mathbf{w}\right) \,.$$

The previously introduced products are giving us the ability to not only couple tensors of different ranks, but also to couple tensor fields of different ranks. Let $\mathbf{v} \in \mathcal{T}_{j_1}$ and $\mathbf{w} \in \mathcal{T}_{j_2}$ and $j_1, j_2, j \in \mathbb{N}$ be chosen according to the triangle inequality (1), then

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

defines a tensor field of rank $j$, i.e., $\mathbf{f} \in \mathcal{T}_j$. Tensor fields can also be combined by convolution. The advantage of combining tensor fields by convolution is that the so defined "convolution" products also respect translation in a certain sense. Therefore, let $\mathbf{v} \in \mathcal{T}_{j_1}$ and $\mathbf{w} \in \mathcal{T}_{j_2}$ and $j_1, j_2, j \in \mathbb{N}$ be chosen according to the triangle inequality (1), then

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

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

### B. Spherical Derivatives

In [9] the spherical derivatives are introduced. They connect spherical tensor fields of different ranks by differentiation. They represent a basic tool in the context of spherical tensor analysis, and will lead to an efficient formulation of our proposed filter.

*Proposition 2 (Spherical Derivatives):* Let $\mathbf{f} \in \mathcal{T}_j$ be a spherical 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

$$\boldsymbol{\nabla}^1 \mathbf{f} := \nabla \bullet_{j+1} \mathbf{f} \tag{2}$$

$$\boldsymbol{\nabla}_1 \mathbf{f} := \nabla \bullet_{j-1} \mathbf{f} \tag{3}$$

where

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

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

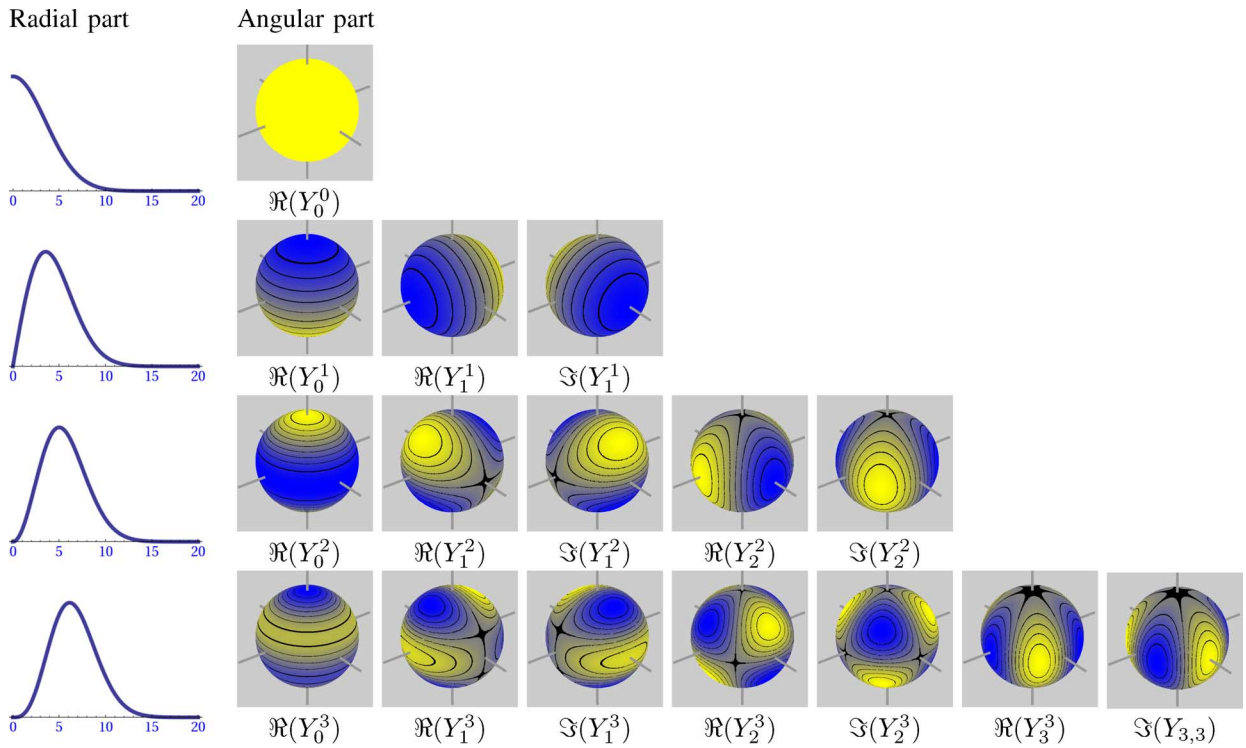



Fig. 1. The radial and the angular part of the first three Gaussian windowed solid harmonics. The radial part is depicted in the left column. The angular part (which are spherical harmonics $Y_m^j$) is depicted in the right column. The color of the angular part indicates the value of the spherical harmonic (yellow corresponds to "$+1$" and blue to "$-1$"). $\Re$ denotes the real and $\Im$ the imaginary part of complex numbers.

$\bullet_j : V_{j_1} \times V_{j_2} \to V_j$ are normalized symmetric products given by

$$\mathbf{v} \bullet_j \mathbf{w} := \frac{1}{\langle j\,0 \mid j_1 0, j_2 0\rangle}\, \mathbf{v} \circ_j \mathbf{w}\,.$$

They are only defined for $j+j_1+j_2 =$ even (see [9] for details). Note, for a scalar function the spherical up-derivative is just the spherical gradient, i.e., $\nabla f = \boldsymbol{\nabla}^1 f$.

### C. Spherical Gaussian Derivatives

As a prerequisite to the Harmonic filter it is necessary to mention that the spherical derivative $\boldsymbol{\nabla}^j$ of a Gaussian is just a Gaussian-windowed solid harmonic $\mathbf{G}_\sigma^j$

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

where $\mathbf{R}^j(\mathbf{r}) := \|\mathbf{r}\|^j \mathbf{Y}^j(\mathbf{r}/\|\mathbf{r}\|)$ are the so called solid harmonics, and $\mathbf{Y}^j$ are the well-known spherical harmonics. In our paper, we use tensor notation for spherical harmonics, therefore we write $\mathbf{Y}^j = (Y_{-j}^j, \ldots, Y_j^j)$ where $Y_m^j$ is a spherical harmonic of band $j$ and $m = -j, \ldots, j$. Using this notation we are able to write $\mathbf{Y}^j$ as a mapping from the 2-sphere $S^2$ onto $\mathbb{C}^{2j+1}$

$$\mathbf{Y}^j : S^2 \to \mathbb{C}^{2j+1}\,.$$

The solid harmonics $\mathbf{R}^j$ are a basis for harmonic functions. Fig. 1 shows the radial and the angular part of the first three Gaussian windowed solid harmonics. An implication is that convolutions with $\mathbf{G}_\sigma^j$ are derivatives of Gaussian-smoothed functions, namely

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

where $f \in \mathcal{T}_0$. We use the convention

$$\mathbf{G}_\sigma^0 = G_\sigma = \frac{1}{(\sqrt{2\pi}\sigma)^3} e^{-r^2/2\sigma^2}\,.$$

## III. Multichannel Harmonic Filters

Our goal is to build nonlinear image filters that are equivariant to Euclidean motion. The group of Euclidian motion of degree 3 is denoted as $SE(3)$. An $SE(3)$-equivariant multichannel image filter is given by the following definition.

*Definition 3 ($SE(3)$-Equivariant Multichannel Filter):* A scalar multichannel image filter is a mapping

$$H : \mathcal{T}_0 \times \cdots \times \mathcal{T}_0 \to \mathcal{T}_0\,.$$

We call such a mapping $SE(3)$-equivariant if

$$H\{gf_1, \ldots, gf_k\} = gH\{f_1, \ldots, f_k\}$$

for all $g \in SE(3)$ and $f_i \in \mathcal{T}_0$ for $i = 1, \ldots, k$.

For building such a filter we follow the idea of [4]. By treating each channel as a separate scalar-valued function $f_i : \mathbb{R}^3 \to \mathcal{T}_0$ for $i = 1, \ldots, k$ ($k \in \mathbb{N}$ denotes the number of channels), we can integrate this into the harmonic filter framework under

the assumption that all channels are transformed by the same $g \in SE(3)$.

Therefore, we first compute for each position $\mathbf{r} \in \mathbb{R}^3$ the projection onto the Gaussian windowed harmonic basis $\mathbf{G}_\sigma^j$ for $j = 0, \ldots, n$ by a convolution of the input channels $f_i$ with the basis elements

$$\mathbf{p}_i^j := \mathbf{G}_\sigma^j * f_i \,.$$

In a second step, we map these projections $\mathbf{p}_i^j$ onto some new local harmonic descriptors by means of a nonlinear mapping

$$\begin{aligned}\mathbf{N}^j(\mathbf{r}) := \mathbf{N}^j[&\mathbf{p}_1^0(\mathbf{r}), \mathbf{p}_1^1(\mathbf{r}), \ldots, \mathbf{p}_1^n(\mathbf{r}), \ldots, \\ &\mathbf{p}_k^0(\mathbf{r}), \mathbf{p}_k^1(\mathbf{r}), \ldots, \mathbf{p}_k^n(\mathbf{r})]\end{aligned} \tag{4}$$

which describe the harmonic part of $f_i$ in a local neighborhood. This can be interpreted as a local expansion of a kind of voting function. The contribution of the voter at position $\mathbf{r}'$ to the voxel $\mathbf{r}$ is then given by

$$V_{\mathbf{r}'}(\mathbf{r}) = G_\eta(\mathbf{r} - \mathbf{r}') \sum_{j=0}^{n} \mathbf{N}^j(\mathbf{r}') \bullet_0 \mathbf{R}^j(\mathbf{r} - \mathbf{r}') \,. \tag{5}$$

In a final step, the contribution of all voters that contribute to a voxel $\mathbf{r}$ is collected by convolution

$$\begin{aligned}H\{f_1, \ldots, f_k\}(\mathbf{r}) &:= \int_{\mathbb{R}^3} V_{\mathbf{r}'}(\mathbf{r}) d\mathbf{r}' \\ &= \sum_{j=0}^{n} \int_{\mathbb{R}^3} \mathbf{G}_\eta^j(\mathbf{r} - \mathbf{r}') \bullet_0 \mathbf{N}^j(\mathbf{r}') \\ &= \sum_{j=0}^{n} \mathbf{G}_\eta^j \,\tilde{\bullet}_0\, \mathbf{N}^j.\end{aligned} \tag{6}$$

The output of the filter reflects the probability of the occurrence of the trained structure in a local window. In order to ensure $SE(3)$-equivariance, $\mathbf{N}^j$ has to obey the equivariance constraint (see [4])

$$\begin{aligned}\mathbf{N}^j[\mathbf{D}_g^0\mathbf{p}_1^0, \ldots, \mathbf{D}_g^n\mathbf{p}_1^n, \ldots, \mathbf{D}_g^0\mathbf{p}_k^0, \ldots, \mathbf{D}_g^n\mathbf{p}_k^n] \\ = \mathbf{D}_g^j \mathbf{N}^j[\mathbf{p}_1^0, \ldots, \mathbf{p}_1^n, \ldots, \mathbf{p}_k^0, \ldots, \mathbf{p}_k^n] \,.\end{aligned} \tag{7}$$

For an efficient computation we make use of the spherical derivatives in order to compute the projections onto the harmonic basis. Furthermore, using the commuting property for convolutions (see [9]) we can rewrite (6) as

$$\begin{aligned}\sum_{j=0}^{n} \mathbf{G}_\eta^j \,\tilde{\bullet}_0\, \mathbf{N}^j &= \sum_{j=0}^{n} (\boldsymbol{\nabla}^j G_\eta) \,\tilde{\bullet}_0\, \mathbf{N}^j \\ &= \sum_{j=0}^{n} G_\eta \,\tilde{\bullet}_0\, (\boldsymbol{\nabla}_j \mathbf{N}^j) \\ &= G_\eta * \sum_{j=0}^{n} \boldsymbol{\nabla}_j \mathbf{N}^j.\end{aligned}$$

The filter algorithm in terms of spherical derivatives is then given by

$$\begin{aligned}H\{f_1, \ldots, f_k\} := G_\eta * \sum_{j=0}^{n} \boldsymbol{\nabla}_j \mathbf{N}^j[&\boldsymbol{\nabla}^0 f_s^1, \ldots, \boldsymbol{\nabla}^n f_s^1, \\ &\vdots \\ &\boldsymbol{\nabla}^0 f_s^k, \ldots, \boldsymbol{\nabla}^n f_s^k]\end{aligned} \tag{8}$$

with $f_s^i = G_\sigma * f_i$ are the Gaussian smoothed channels and $\boldsymbol{\nabla}^0 f_s^i, \ldots, \boldsymbol{\nabla}^n f_s^i$ are their projections onto the Gaussian windowed harmonic basis.

The parameter $\sigma$ of the input Gaussian determines the size of the local features which vote for the center of the object of interest. In order to assure that every voxel of the object can contribute to the resulting output, the size of the voting function $\eta$ should be at least half the diameter of the object. The parameter $n$ determines the expansion degree of our input.

We will utilize the spherical products $\circ$ to build the $SE(3)$-equivariant nonlinear mappings $\mathbf{N}^j[\cdot]$ for multichannel images.

### A. Multichannel Voting Functions

In this section, we will discuss how to construct the nonlinear mapping $\mathbf{N}^j[\cdot]$ for multichannel images. We will do this by exemplary for $k = 2$ channels and give two possible formulations.

For the first one, let $\mathbf{f}_1, \mathbf{f}_2 \in \mathcal{T}_0$ and let $\mathbf{p}_1^0, \ldots, \mathbf{p}_1^n$ and $\mathbf{p}_2^0, \ldots, \mathbf{p}_2^n$ be their expansion in the Gaussian windowed harmonic basis. We define $\mathbf{N}_{\text{Multi}}^j$ as

$$\mathbf{N}_{\text{Multi}}^j[\cdot] := \sum_{\substack{|j_1 - j_2| \le j \le j_1 + j_2 \\ j_1 + j_2 + j \text{ even} \\ j_1, j_2 \le n}} \alpha_{j_1, j_2}^j \, \mathbf{p}_1^{j_1} \bullet_j \mathbf{p}_2^{j_2} \tag{9}$$

where $\alpha_{j_1,j_2}^j \in \mathbb{R}$ are expansion coefficients. Since $\mathbf{p}_1^m, \mathbf{p}_2^m \in \mathcal{T}_m$ for $m = 0, \ldots, n$ and knowing the properties of $\bullet_j$ it is obvious that (9) defines an $SE(3)$-equivariant mapping. The proof is analogue to the one given in [4] for the single-channel filter. Algorithm 1 gives a realization of the multichannel harmonic filter, where $\mathbf{N}^j = \mathbf{N}_{\text{Multi}}^j$. In order to become invariant against additive intensity changes we neglect the zero order descriptor $\mathbf{p}^0$. To gain robustness against illumination/contrast changes we use the same normalization as [4]. Therefore, we introduce a special case for $j = 1$ in the for-loop in algorithm 1 from line 3–6, and normalize the first order descriptor $\mathbf{p}^1$ by

$$\mathbf{p}^1(\mathbf{r}) = \frac{1}{\gamma + s_{\text{dev}}(\mathbf{r})} \boldsymbol{\nabla}^1 f(\mathbf{r}) \,. \tag{10}$$

Here, $\gamma \in \mathbb{R}$ is a fixed regularization parameter avoiding zero divisions 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. A possible workflow of the filter is depicted in Fig. 2 (only symmetric products $\bullet$ are computed for $\mathbf{N}^j = \mathbf{N}_{\text{Multi}}^j$).

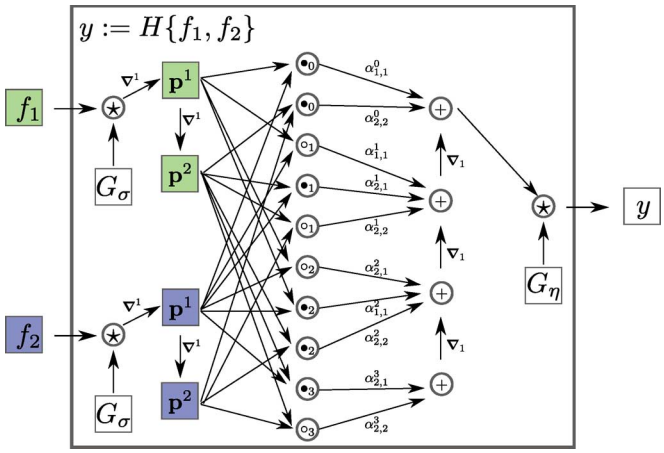



Fig. 2. Workflow of the harmonic filter for two channels: green and blue. The input is expanded up to a degree of $n = 2$. The star $\star$ denotes a convolution of the incoming images. The plus denotes an addition of incoming channels. $\nabla$ denotes the spherical derivative. $\circ$ and $\bullet$ denote the spherical coupling of incoming channels.

**Algorithm 1** Multichannel Harmonic Filter Algorithm

**Input:** $f_1, f_2 : \mathbb{R}^3 \to \mathcal{T}_0, i = 1, 2$

**Output:** $y : \mathbb{R}^3 \to \mathbb{R}, y := H\{f_1, f_2\}$

// compute projections onto harmonic basis

1: Convolve $\mathbf{p_1}^0 := G_\sigma * f_1$

2: Convolve $\mathbf{p_2}^0 := G_\sigma * f_2$

3: **for** $j = 1 : n$ **do**

4: $\mathbf{p_1}^j = \boldsymbol{\nabla}^1 \mathbf{p_1}^{j-1}$

5: $\mathbf{p_2}^j = \boldsymbol{\nabla}^1 \mathbf{p_2}^{j-1}$

6: **end for**

// compute harmonic descriptors

7: Initialize $\mathbf{y}^n := 0 \in \mathcal{T}_n$

8: **for** $j = n : -1 : 1$ **do**

9: $\mathbf{y}^{j-1} = \boldsymbol{\nabla}_1 \left(\mathbf{y}^j + \mathbf{N}^j \left[\mathbf{p_1}^0, \ldots, \mathbf{p_1}^n, \mathbf{p_2}^0, \ldots, \mathbf{p_2}^n\right]\right)$

10: **end for**

11: Let $y := \mathbf{y}^0 + \mathbf{N}^0 \left[\mathbf{p_1}^0, \ldots, \mathbf{p_1}^n, \mathbf{p_2}^0, \ldots, \mathbf{p_2}^n\right]$

// collect the contribution of all voters

12: Convolve $y := G_\eta * y$

A second possibility is to extend $\mathbf{N}^j_{\text{Multi}}$ by using antisymmetric products. This extension makes only sense for multichannel images, as we will see. For example, let $j = j_1 = j_2 = 1$ and $\mathbf{v} = \mathbf{S}\mathbf{b}$, $\mathbf{w} = \mathbf{S}\mathbf{c} \in V_1$ and $\mathbf{b}, \mathbf{c} \in \mathbb{R}^3$ it holds

$$(\mathbf{v} \circ_1 \mathbf{w}) = -\frac{\mathbf{i}}{\sqrt{2}} \mathbf{S}\,(\mathbf{b} \times \mathbf{c})$$

where $\times$ denotes the cross product. When we think of $\mathbf{v} = \boldsymbol{\nabla}^1 f(\mathbf{r})$ and $\mathbf{w} = \boldsymbol{\nabla}^1 g(\mathbf{r})$ being the spherical gradient at position $\mathbf{r}$ of two scalar fields $f$ and $g$ [actually we have $\boldsymbol{\nabla}^1 f(\mathbf{r}) = \mathbf{S}\,\nabla f(\mathbf{r})$ and $\boldsymbol{\nabla}^1 g(\mathbf{r}) = \mathbf{S}\,\nabla g(\mathbf{r})$], then $\mathbf{v} \circ_1 \mathbf{w}$ computes the cross product of their gradients at position $\mathbf{r}$. Obviously, this is only useful when we are in the context of multichannel images, otherwise we have (think of the single channel case) that $(\boldsymbol{\nabla}^1 \mathbf{f}(\mathbf{r})) \circ_1 (\boldsymbol{\nabla}^1 \mathbf{f}(\mathbf{r})) = 0$.

Since the resulting field $\mathbf{f} \circ_j \mathbf{g}$ for $\mathbf{f} \in \mathcal{T}_{j_1}, \mathbf{g} \in \mathcal{T}_{j_1}$ for $j + j_1 + j_2 = \text{odd}$, maps onto elements of $\mathbf{i}V_j$ (see Prop. 1), we have to multiply with $-\mathbf{i}$ in order to get a spherical tensor field that maps onto $V_j$

$$-\mathbf{i}\,(\mathbf{f} \circ_j \mathbf{g}) : \mathbb{R}^3 \to V_j\,.$$

We have then that $-\mathbf{i}\,(\mathbf{f} \circ_j \mathbf{g}) \in \mathcal{T}_j$. Inserting this in our voting function $\mathbf{N}^j_{\text{Multi}}[\cdot]$ we are able to formulate the following extension for multichannel images. We define $\mathbf{N}^j_{\text{Asym}}$ as

$$\mathbf{N}^j_{\text{Asym}}[\cdot] := \sum_{\substack{|j_1 - j_2| \le j \le j_1 + j_2 \\ j_1 + j_2 + j \text{ even} \\ j_1, j_2 \le n}} \alpha^j_{j_1,j_2}\, \mathbf{p}_1^{j_1} \bullet_j \mathbf{p}_2^{j_2} - \sum_{\substack{|j_1 - j_2| \le j \le j_1 + j_2 \\ j_1 + j_2 + j \text{ odd} \\ j_1, j_2 \le n}} \alpha^j_{j_1,j_2}\, \mathbf{i}\,(\mathbf{p}_1^{j_1} \circ_j \mathbf{p}_2^{j_2}) \quad (11)$$

where $\alpha^j_{j_1,j_2} \in \mathbb{R}$ are expansion coefficients. A possible workflow of the filter is depicted in Fig. 2. Algorithm 1 gives a possible realization for $\mathbf{N}^j = \mathbf{N}^j_{\text{Asym}}$.

### *B. Multichannel Filters for Tensor Fields*

The generalization to tensor-valued input and output is indeed very simple. We will give an example for a first order tensor field with a scalar output.

One has to construct a filter of type $H : \mathcal{T}_{j_1} \times \cdots \times \mathcal{T}_{j_1} \to \mathcal{T}_{j_2}$ in the sense of definition 3 (but now for arbitrary $j_1$ and $j_2$). According to the formulation we gave in Section III in order to compute the projection onto the harmonic basis, we can compute the descriptor images for a higher order input $\mathbf{f} \in \mathcal{T}_{j_1}$ as before by $\mathbf{p}^j = \boldsymbol{\nabla}^j (G_\sigma * \mathbf{f})$. The difference is that the resulting descriptor images have higher orders

$$\mathbf{p}^j \in \mathcal{T}_{j+j_1}\,.$$

We also have to adjust the local nonlinear mappings $\mathbf{N}^j[\cdot]$. They have to be of the following type:

$$\mathbf{N}^j : \mathcal{T}_{o+j_1} \times \cdots \times \mathcal{T}_{n+j_1} \to \mathcal{T}_{j+j_1}\,.$$

Actually, we can use the same formulation as in (8) to map the local nonlinearities onto the output.

A vectorial filter with scalar-valued output is given by the following definition.

*Definition 4 ($SE(3)$-Equivariant Vectorial Filter):* A vectorial image filter is a mapping

$$H : \mathcal{T}_1 \times \cdots \times \mathcal{T}_1 \to \mathcal{T}_0\,.$$

We call such a mapping $SE(3)$-equivariant if

$$H\{gf_1, \ldots, gf_k\} = gH\{f_1, \ldots, f_k\}$$

for all $g \in SE(3)$ and $f_i \in \mathcal{T}_1$ for $i = 1, \ldots, k$.

### C. Preprocessing for the Vectorial Harmonic Filter

We used the gradient vector flow field ([15] and [16]) of a scalar-valued function in order to construct an rank-1 tensor field. Given a scalar field $f : \mathbb{R}^3 \to \mathbb{R}$, the gradient vector flow field is defined as the vector field $\mathbf{v} : \mathbb{R}^3 \to \mathbb{R}^3$ that minimizes the energy functional

$$E = \int_{\mathbb{R}^3} \mu \sum_i \left( \frac{\partial \mathbf{v}(\mathbf{r})}{\partial r_i} \right)^2 + \|\nabla f(\mathbf{r})\|^2 \|\mathbf{v}(\mathbf{r}) - \nabla f(\mathbf{r})\|^2 \, d\mathbf{r} \, . \tag{12}$$

Due to the variational formulation, the resulting field gets smooth in regions with low gradient magnitude $\|\nabla f\|$. In regions with large gradient magnitude $\|\nabla f\|$ $\mathbf{v}$ tends to $\nabla f$. The parameter $\mu$ is a regularization parameter and should be set according to the amount of noise presented in the image (see [16] for details). After computing the gradient vector flow we applied the vectorial image filter to the resulting field.

## IV. Experiments

### A. Detection of Dividing Cells in Metaphase of Mitosis

*1) Data:* For our study, the 3D structure of metaphase cells was visualized by staining the DNA material with DAPI (blue channel) and immunofluorescence staining of the protein kinase Aurora-A [17] (green channel). This was performed on colorectal cancer cells lines, which reflect differential Aurora-A protein expression of colorectal cancer tissue specimens [18], [19]. With this, we provide the basis for continuing functional biological studies addressing the question of how differential Aurora-A affects quantitative and qualitative aspects of metaphase in mitotic colorectal cancer cells. All image stacks were recorded with a Carl Zeiss microscope equipped with an ApoTome and PlanApochromat 40x/1.3 oil objective at 1 $\mu$m distance, resulting in a voxel size of 0.16 $\mu$m$\times$ 0.16 $\mu$m$\times$ 1 $\mu$m. For further analysis the data was rescaled to cubic voxels of size 0.5 $\mu$m$\times$ 0.5 $\mu$m$\times$ 0.5 $\mu$m. This was done for practical reasons, such as reducing the actual window size of our basis functions resulting in convolutions with smaller data and smaller features to store in memory. The remaining information after rescaling was sufficient for solving the problem at hand. However, one can easily adapt the method to the real-world voxel size by the following modifications: Our method uses a Gaussian window for windowing the harmonic basis functions and for gathering the contribution of all voters in the final collection step. In order to take the real-world voxel size into account, the Gaussian in grid-coordinates has to be anisotropically squeezed and stretched to obtain an isotropic Gaussian in the real world coordinate system. In a similar way the noncubic voxel size has to be incorporated during the computation of the spherical derivatives via the finite difference scheme by dividing the individual partial derivatives by the corresponding edge length of the voxel.

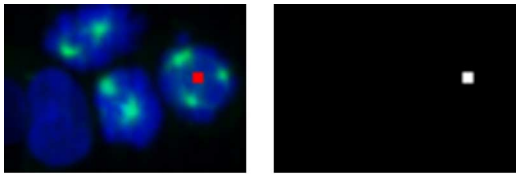
*2) Reference Approaches:* For steerable filters, we used a similar detector to the one introduced in [2]. Since we do not have templates that can be defined in an analytical way, like a line or a surface as used by [2], we determined the detector (template) empirically. We chose the parameters to best fit the spindle centers of mitotic cells in shape and size. Since the steerable filter approach can be interpreted as kind of local voting procedure for the desired template, we did a further collection step as for the harmonic filters. This means we applied a Gaussian of size $\eta$, after the steerable filter algorithm.

For a better comparison with our results we used a second steerable filter for roundish structures on the blue channel and combined it with the previously described steerable filter for the green channel. To be more precise, we computed the gradient magnitude image of the blue channel and detected sphere-like structures. The steerable filter approach itself is an approach per channel, and by combining channels the response is therefore multidimensional and has the dimensionality of the number of channels (the filter response is the Cartesian product of the filter responses of all channels). For decision making one would usually use a classifier (e.g., a support vector machine) which maps the multidimensional filter response onto a binary output (decision). However for the problem at hand, there is a way to combine the multidimensional filter response to a one-dimensional, namely by using a logical AND operation. Since the structure of a mitotic metaphase cell has to be detected in both channels (green: spindle centers, blue: roundish DAPI) and therefore has a high filter response at the same positions in both channels, we combine them by multiplication (logical AND). Although, the filter on the blue channel only has very poor results, the combination is better than only using the green channel.

Furthermore, we used the original (single channel) harmonic filter approach with the same parameters for $\sigma$, $\eta$, $\gamma$ and expansion degree as for our proposed multichannel filter. In fact, we made a little modification, namely training the filter on both, the green and the blue channel, which led to much better classification results than using only one channel.

In addition we applied a standard 3D morphology-based approach: On the blue nuclei channel this approach mainly consists of a shading correction, a gray-valued "fill-holes" and an Otsu-Thresholding. In a postprocessing step a watershed on the inverted distance transform of the binary image was used to cut clustered cell nuclei. On the green spindle channel a thresholding relative to the maximum intensity and a connected component labeling was performed. An analysis of the overlapping regions of the two channels then finds the mitotic cells (the full algorithm is described in the supplements).

*3) Training (Only Harmonic Filters):* We determined the expansion coefficients $\alpha$ of the voting function in a training step. For all harmonic filters, training was conducted on two image stacks containing 17 mitotic cells in metaphase. Since the filter is linear in $\alpha$ we followed a linear least square fit for determining optimal expansion coefficients. Therefore, we provide for all training samples the desired output in form of a label image with ones where we want the filters to have high response, and otherwise zeros where no response is requested. In Fig. 3, this is illustrated at a 2D example. In this example we want to have our filter detecting the cell, which is marked with a red



(a) (b)

Fig. 3. Illustration of how the label images were designed at a 2D example. (a) shows the original image with the cell we want to detect. The wanted cell/position is marked with a red point. (b) shows the corresponding label image. In this case the label image contains a one at the pixel which corresponds to the red point in (a), and other pixels remain zero.

point [see Fig. 3(a)], and not detecting other cells in the image. Furthermore, we also want our filter to have a peak exactly at the marked position. This would then result in a label image as depicted in Fig. 3(b), in which a single pixel is set to one at the center of the cell, and all other pixels remaining zero. Afterward, the expansion coefficients arise by solving the normal equations of the corresponding system of linear equations.

*4) Results:* Our test data set contained 14 image stacks with an overall amount of 91 mitotic cells in metaphase. The steerable filter **SF** was applied to the green channel only. In addition, we used a version with a further collection step of local voters for the steerable filter **SFS**. The steerable filter which combines responses of both channels as mentioned in Section IV-A-II is denoted as **SFMC**.

The single-channel harmonic filter **HF-Single** was trained on both channels. This means that after training the model also incorporates the information and relation of both channels. This is due to the fact that the evolving system of linear equations for computing the filter parameters consists of information of both channels.

The (scalar) multichannel filters **HFM**, **HFM-Asym** using the information of each single channel and the combination of both. **HFM** uses the voting function as described in (9), and **HFM-Asym** the one given by (11). In detail, **HFM** computes the same products as **HF-Single**, plus all possible symmetric products. In addition to **HFM**, **HFM-Asym** computes all possible antisymmetric products. The (vectorial) multichannel filter **HFM-GVF** is in principle the same filter as **HFM**, but this filter uses the gradient vector flow field of each channel as input data. We used a normalized version for all harmonic filters except for the vectorial filter **HFM-GVF**. The normalization was done as proposed in Section III-B. This means we normalized the first order descriptor of each channel as described in (10), and neglected the zero order descriptor (simply the intensities). For **HFM-GVF** we are already invariant against additive intensity changes, because the gradient vector flow field is based on the image gradient (see Section III-D). But we did no further normalization for this type of filter. An overview of all filters can be found in Table I. The color or color combinations indicate the channels and/or products they incorporate.

For $\eta$ we chose half of the average cell diameter (4.5 $\mu$m). We achieved our best results with a local feature size of $\sigma = 1.5\,\mu\mathrm{m}$. Since the intensity values of our data were very stable regarding outliers, we chose to scale the data onto the interval $[0, 1]$ (we

TABLE I
OVERVIEW OF THE DIFFERENT FILTERS USED IN OUR EXPERIMENTS

| | SF(S) | SFMC | HFM | HF-Single | HMF-GVF | HFM-Asym |
|---|---|---|---|---|---|---|
| ● | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| ● | | ✓ | ✓ | ✓ | ✓ | ✓ |
| ●● | | | ✓ | | ✓ | ✓ |

did the same for the steerable filters) and set gamma to the maximum value of the data ($\gamma = 1$) which was found to very robust. Slight variations of $\gamma$ did not influence the result. When $\gamma$ is selected much smaller (e.g., $\gamma = 0.1$), the overall performance drops. For the expansion degree we chose $n = 4$ for **HFM** and **HFM-GVF**. For **HFM-Asym** we only calculated the expansion for $n = 3$. The results for $n = 4$ were worse than for $n = 3$. An explanation for this is, that we compute a lot more products than for the other filters, and the filter looses its ability to generalize and adapts too much to the training data. In this context we also have to mention, that our training data sets do not cover all the intra-class variance of the mitotic cells occurring in our test data set, and one has to be careful to avoid over-fitting in the training step. Fig. 4 shows qualitative results of the filter responses of **SF** and **HFM** on cell line data and should give an impression of the computed filter responses. The presented results of **SF** should also clarify why for **SFS** a further smoothing was performed to gather nearby maxima to one single maximum.

For the evaluation we determined precision and recall. In order to determine these measures we have to determine "number of hits ($\mathrm{N_{Hits}}$)," "number of false alarms ($\mathrm{N_{FalseAlarms}}$)," "number of misses ($\mathrm{N_{Misses}}$)" and calculate

$$\mathrm{precision} = \frac{\mathrm{N_{Hits}}}{\mathrm{N_{Hits}} + \mathrm{N_{FalseAlarms}}}, \quad \mathrm{recall} = \frac{\mathrm{N_{Hits}}}{\mathrm{N_{Hits}} + \mathrm{N_{Misses}}}.$$

For each manually labeled cell we determine all local maxima of the filter response in a certain small radius (3.5 $\mu$m) around the manual label position and mark them as possible detections (see Fig. 5) for this cell. Then for each cell, we consider all possible detections and mark only the one with the highest filter response as possible hit for this cell and mark all others as possible false alarm, as well as all other local maxima which are not in a precision radius around a manual label. The values for $\mathrm{N_{Hits}}$, $\mathrm{N_{Misses}}$, and $\mathrm{N_{FalseAlarms}}$ are then determined by setting a threshold for the filter response (decision boundary for classification). This means $\mathrm{N_{Hits}}$ are possible hits with a value higher than the threshold, $\mathrm{N_{FalseAlarms}}$ are possible false alarms with a value higher than the threshold, and $\mathrm{N_{Misses}}$ are all labels with no hit.

The precision/recall graph for a varying threshold is depicted in Fig. 6. All results were calculated as described above with a radius of 3.5 $\mu$m around the given label positions.

The morphology-based approach reaches good results on the cell line data. A recall of 96.7% at a precision of 84.6% was reached when the parameters were tuned for high recall (see supplements for details).

### *B. Detection of a Specific Structure*

For the task of mitosis detection as discussed above, it is important to have a model which neglects higher frequencies and is able to generalize. However, our approach can also be used to

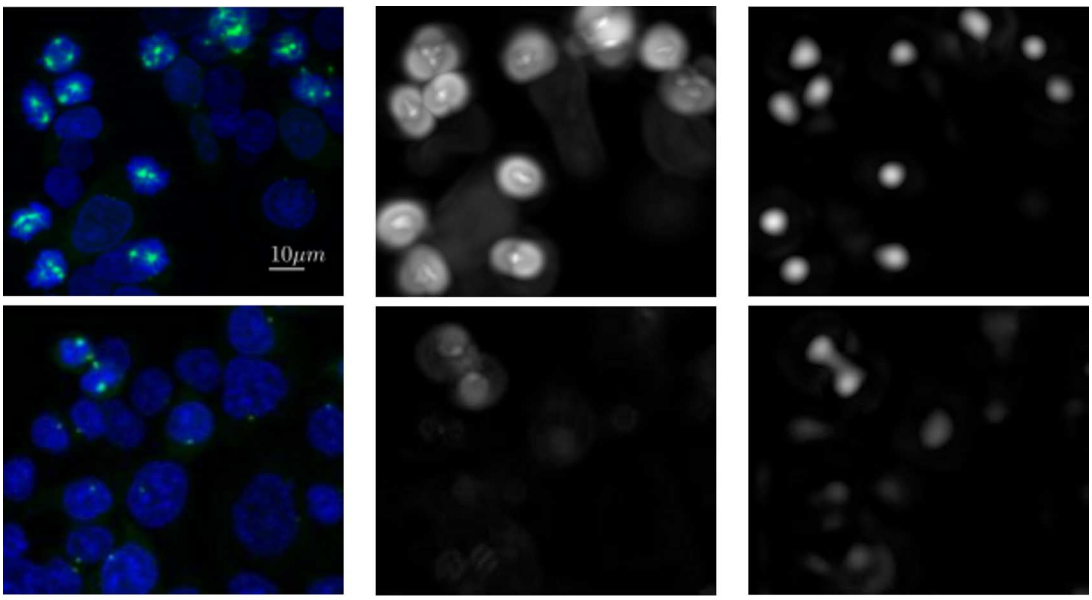



Fig. 4. Qualitative results of (b) a steerable and (c) a harmonic filter on (a) a cell line 3D test image of colorectal cancer. (a) Maximum intensity projection of the test image; (b) and (c) show the corresponding filter response. (a) Original. (b) SF. (c) HFM.

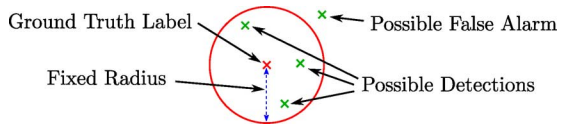



Fig. 5. An illustration of how we determined the values for precision and recall. Of all possible detections (local maxima in the filter response), we mark only the one with the highest filter response as possible hit. A possible false alarm is a local maximum not within the precision radius of a ground truth label. After setting a threshold for the filter response we decide whether a possible hit is counted as hit or not (has to be higher than the threshold), and if a possible false alarm is counted as false alarm or not (has to be higher than the threshold). If a ground truth label has no hit than it is counted as miss.

distinguish objects which have a (well-defined) specific structure from all other objects, e.g., a mitotic cell in metaphase with exactly three blobs in the green channel and a "ball-shaped" structure in the blue channel. Fig. 7(a) shows a training (upper) and a test (lower) data set, each of them containing a mitotic cell with this specific structure. For the training step, we labeled only the cell we want to detect. The others are considered as background. The filter response after training is depicted in Fig. 7(b) (upper). Fig. 8 shows the first few detections sorted in descending order (from left to right) by the filter response value. Following a max-margin approach, we determined our class boundary for decision making as value between the lowest positive detection and the highest negative detection. In our case this results in a value of 0.955 (see the first two detections from left in Fig. 8) for decision making. The center of the red circle is the location of the local maximum in the thresholded filter response. Applying the so gained model to the test data set [Fig. 7(a) lower], we get the response depicted in Fig. 7(b) (lower). After applying the threshold determined in the training step to our test data set (decision making), we are able to detect only the cell with exactly three blobs in the green and a ball-shaped structure in the blue channel. The result is shown in Fig. 7(c) (lower), also

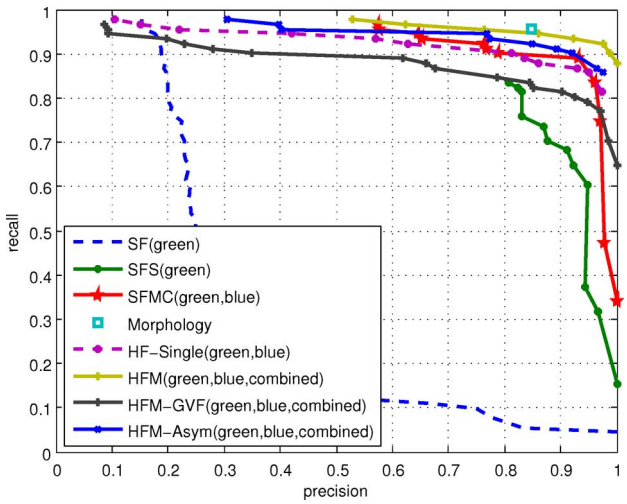



Fig. 6. Precision/Recall graph of the different filter algorithms for the mitosis detection task on cell line data. A detection was counted as positive if it was in a radius of 3.5 $\mu$m around the given label position. Explanation of acronyms in the legend of the plot: SF(green) = "steerable filter green channel;" SFS(green) = "steerable filter green channel with smoothing;" SFMC(green,blue) = "steerable filter green/blue channel with smoothing;" HF-single(green,blue) = "single-channel harmonic filter green/blue channel; " HFM(green,blue,combined) = "multichannel harmonic filter green/blue channel and combined with symmetric products;" HFM-GVF(green,blue,combined) = "multichannel harmonic filter on the gradient vector flow field green/blue channel and combined with symmetric products;" HFM-Asym(green,blue,combined) = "multichannel harmonic filter green/blue channel and combined with symmetric and antisymmetric products."

with a circle around the detected position. A detailed listing of the first few detections of the test data set with the highest filter response values and the corresponding classification results is depicted in Fig. 9.

For gaining such a specificity, one has to expand to higher bands [$n$ in (6)], and as an implication of that compute more products in the voting function [(9) and (11)]. In this case an

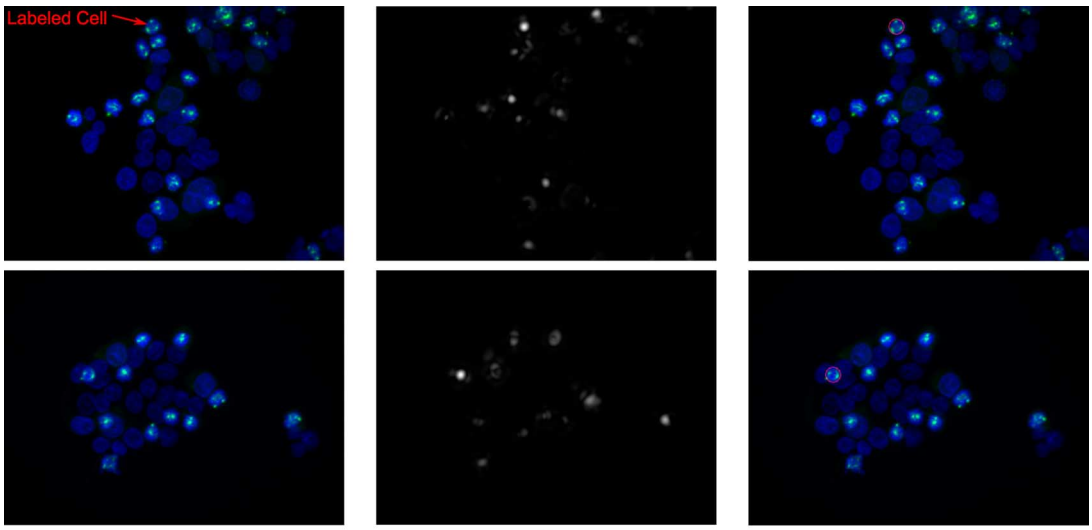
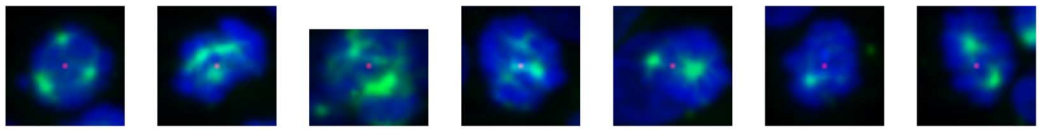
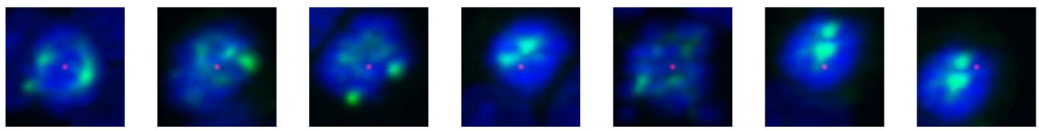



Fig. 7. Results of the experiment where we wanted to detect cells with exactly three blobs in the green channel and a roundish structure in the blue channel. The upper row shows the training data set, and the lower shows the test data set. (a) Original data set of training and test. (b) Corresponding filter response after training and after classification respectively. (c) Locations of the cells marked with a red circle around the detected positions after thresholding the filter response. (a) Original (upper: training/lower: test). (b) Filter response. (c) Detected positions after thresholding.



| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Label | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| Filter Response | 1.13171 | 0.78362 | 0.756323 | 0.566023 | 0.412747 | 0.370805 | 0.359218 |

Fig. 8. Results after training: the first few detections are depicted sorted in descending order (from left to right) by the filter response value. The first row shows a maximum intensity projection of the cropped region around the detection. The detection itself is marked as a red point. The "Label" row shows the corresponding ground truth label which was used during training. In the last row the corresponding filter response value is listed for each detection.



| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Filter Response | 1.35368 | 0.795895 | 0.549038 | 0.416676 | 0.342027 | 0.322867 | 0.316429 |
| Decision | 1 | 0 | 0 | 0 | 0 | 0 | 0 |

Fig. 9. Results after applying the filter to the test data set: the first few detections are depicted sorted in descending order (from left to right) by the filter response value. The first row shows a maximum intensity projection of the cropped region around the detection. The detection itself is marked as a red point. The "Filter Response" row lists the corresponding response value of the detection. The last row shows the class label after decision making.

expansion up to band $n = 5$ was sufficient to detect only cells that have the predefined specific structure.

### C. Experiments on Tissue

We did only few experiments on tissue samples and therefore we have only qualitative results. Fig. 10(c) shows the filter response of the multichannel harmonic filter **HFM** applied on a human tissue 3D test image [Fig. 10(a)]. The filter was trained in a previous step on a human tissue sample [the training sample in this case was the lowermost data set shown in Fig. 11(a)]. As comparison we applied **SFS** [Fig. 10(b)] to the test image.

On the tissue data the morphology-based approach completely failed. The blue nuclei staining in tissue is largely inhomogeneous and the nuclei are densely packed. So neither by the Otsu-Method, nor by manual adjustment a suitable threshold can be found that delivers homogenous regions for the nuclei. Furthermore, the distance-transform based cutting does not work satisfactorily on such densely packed nuclei. In the green spindle channel it was also not possible to segment the requested spindle centers from the highly cluttered background by a threshold approach.

### D. Experiments Concerning the Portability of the Model

Another issue we want to address is, whether it is possible to use a model which was gained by training on cell lines recorded for example by a fluorescence microscope and classify on cell lines recorded by a confocal laser microscope or even classify cells in human tissue.

*1) Training on Cell Lines and Classify in Tissue:* For this experiment, we used the same model as for the classification of cell lines (see Section IV-A) to classify cells in tissue. Fig. 11 shows the results of this experiment. For the harmonic filter

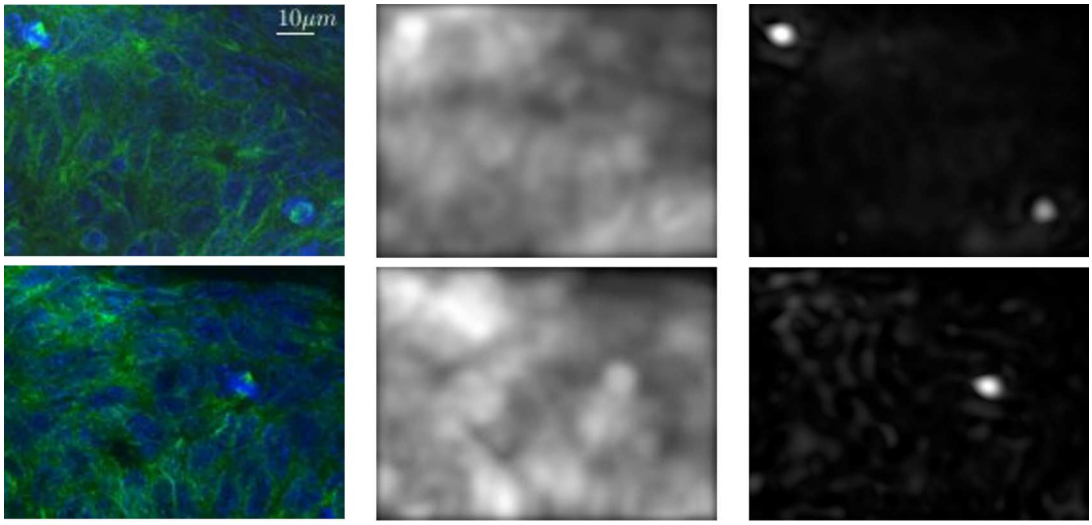



Fig. 10. Qualitative results of (b) steerable and (c) harmonic filter on (a) a human tissue 3D test image of colorectal cancer. (a) Maximum intensity projection of the test image; (b) and (c) show the corresponding filter response. (a) Original. (b) SFS. (c) HFM.

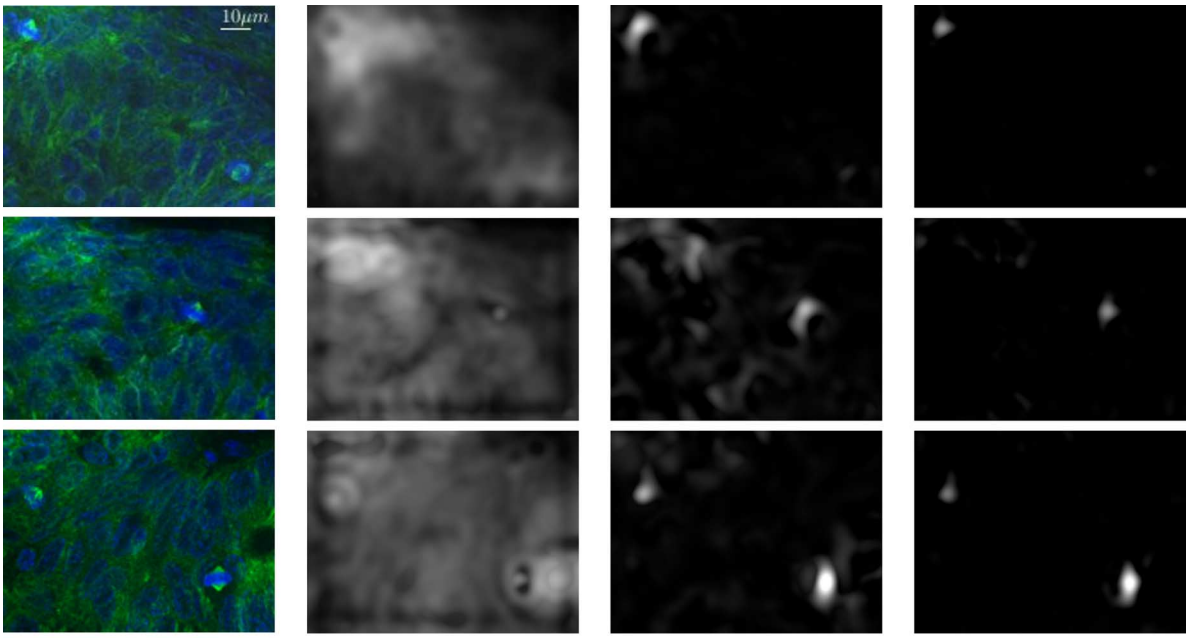



Fig. 11. Results of the classification on tissue data sets with a previous gained model of cell line cells. (a) Original test images. (b) Results of HFM. (c) Results of HFM-GVF. (d) Results of HFM-GVF but with higher expansion in training. (a) Original. (b) HFM. (c) HFM-GVF. (d) HFM-GVF (higher band).

**HFM**, which was trained on the raw image stacks of cell lines, we can clearly see that it is not possible to detect the mitotic cells. Although we have little peaks in the filter response [at least for the lower two data sets in Fig. 11(b)], a thresholding of these responses would result in a very low precision. But if we have a look at the filter responses we get from the **HFM-GVF** [Fig. 11(c)], we see that after using the gradient vector flow as a preprocessing step the results look much better. After thresholding, we are able to detect all cells except the lower right of the first data set.

The explanation for the missing detection of the cell in the topmost data set shown in Fig. 11(a) is due to the fact that the cell is only partially recorded. Fig. 12 shows the corresponding cells of this data set in a shaded volume rendering. One can see that the cell is located at the border of the data set, and only half of the cell is recorded.

We can improve our results by expanding to higher bands in training, as one can see in Fig. 11(d). We expanded one band higher than in Fig. 11(c) from $n = 4$ to $n = 5$.

*2) Changing the Microscope:* In our last experiment, we classified cell line data recorded with a confocal laser scanning microscope. The data sets contained an overall amount of 12 mitotic cells. We used the model of **HFM** from the mitosis detection experiment to classify all data sets. By setting the threshold

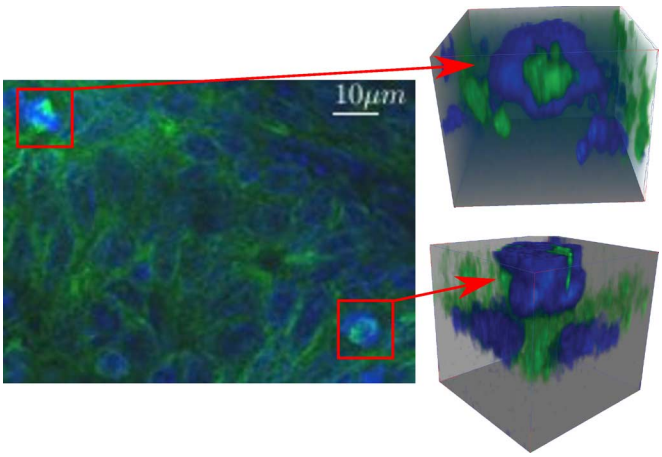



Fig. 12. Shaded volume rendering of two mitotic cells of a tissue data set of colorectal cancer. The upper left cell was fully recorded as the lower right cell was only recorded half. The cutting plane can be clearly seen in the shaded volume rendering.

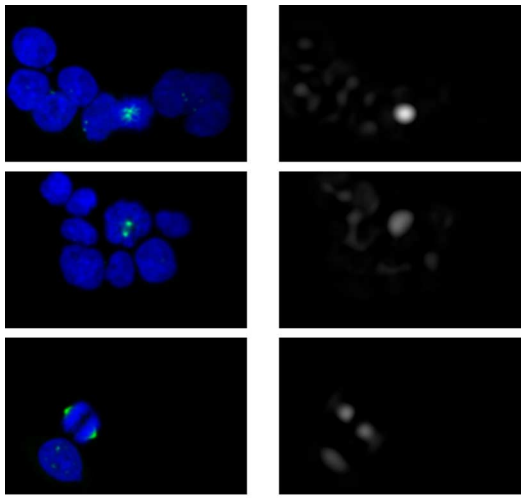



Fig. 13. Results of the classification of cell line data recorded with confocal laser scanning microscope with a model trained on cell line data recorded with fluorescence microscope. (a) Original. (b) Filter response.

to end up with a precision of 100%, we were able to detect 10 out of 12 mitotic cells. The two cells which were not detected are depicted in the lower two rows of Fig. 13. For the lowermost the explanation is quite easy. The cell is not in metaphase anymore (a false label was assigned), but all our training samples were. As one can see in the corresponding filter response, there are two small peaks at each part of the cell. The diameter of middle cell is larger (blue channel) than by the other mitotic cells. As reference in the uppermost row, a representative example for the other mitotic cells is depicted.

## V. Conclusion

We have proposed a new approach for the design of $SE(3)$-equivariant nonlinear multichannel image filters suitable for generic feature/object detection. Our approach is not limited to one specific task like the detection of mitoses. It has the ability to be adapted to other problems due to its learning scheme, and therefore many other applications are possible. Furthermore, we have shown that our filter approach performs very well on more challenging tasks like classification of tissue cells or training a model with cell line data sets and classify cells in tissue data sets.

In addition, we gave suggestions for creating $SE(3)$-equivariant voting functions, but other voting functions are possible, as long as they obey the equi-variance property. Our experiments on cell line data have shown that our proposed method leads to significantly better results than a steerable filter or a single-channel harmonic filter approach for the mitosis detection task. On human tissue samples, the qualitative results of our multichannel approach outperformed the steerable filter approach, as well as the morphology based approach which failed completely on tissue.

## References

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

[2] F. Aguet, M. Jacob, and M. Unser, "Three-dimensional feature detection using optimal steerable filters," in *Proc. IEEE Int. Conf. Image Process. (ICIP 2005)*, Sep. 2005, vol. 2, pp. II–1158–61.

[3] G. Sommer, M. Michaelis, and R. Herpers, "The SVD approach for steerable filter design," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS'98)* , May-Jun. 3, 1998, vol. 5, pp. 349–353.

[4] M. Reisert, Harmonic filerts in 3D—Theory and applications Albert-Ludwigs-Univ. Freiburg, Tech. Rep. 1, 2009.

[5] M. Reisert and H. Burkhardt, "Harmonic filters for generic feature detection in 3d," in *Pattern Recognition*, J. Denzler, G. Notni, and H. Süße, Eds. New York: Springer, 2009, vol. 5748, Lecture Notes in Computer Science, pp. 131–140.

[6] J. Schulz, T. Schmidt, O. Ronneberger, H. Burkhardt, T. Pasternak, A. Dovzhenko, and K. Palme, "Fast scalar and vectorial grayscale based invariant features for 3d cell nuclei localization and classification," in *Proceedings of the 28th Pattern Recognition Symposium of the German Association for Pattern Recognition (DAGM 2006)*. Berlin, Germany: Springer, 2006, Lecture Notes Computer Science.

[7] B. Cyganek, "Object detection in multi-channel and multi-scale images based on the structural tensor," in *Computer Analysis of Images and Patterns*. Berlin, Germany: Springer, 2005, vol. 3691/2005, Lecture Notes in Computer Science, pp. 570–578.

[8] P. A. Viola and M. J. Jones, "Robust real-time face detection," *Int. J. Comput. Vis.*, vol. 57, no. 2, pp. 137–154, 2004.

[9] M. Reisert and H. Burkhardt, "Spherical tensor calculus for local adaptive filtering," in *Tensors in Image Processing and Computer Vision*, S. Aja-Fernández, R. de Luis García, D. Tao, and X. Li, Eds. New York: Springer, 2009, Advances in Pattern Recognition.

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

[11] P. Meraldi and E. T. Nigg, "The centrosome cycle," *FEBS Lett.*, vol. 521, no. 1-3, pp. 9–13, Jun. 2002.

[12] M. Rose, *Elementary Theory of Angular Momentum*. Mineola, NY: Dover, 1995.

[13] M. Tinkham, *Group Theory in Quantum Mechanics*. Mineola, NY: Dover, 2004.

[14] M. Reisert, Spherical Derivatives for Steerable Filtering in 3D Univ. Freiburg, IIF-LMB, Comput. Sci. Dept., Tech. Rep. Internal Rep. 3/07, 2007.

[15] C. Xu and J. L. Prince, "Gradient vector flow: A new external force for snakes," in *IEEE Proc. Conf. Comput. Vis. Pattern Recognit.*, 1997, pp. 66–71.

[16] C. Xu and J. Prince, "Snakes, shapes, and gradient vector flow," *IEEE Trans. Image Process.*, vol. 7, no. 3, pp. 359–369, Mar. 1998.

[17] O. Gautschi, J. Heighway, P. Mack, P. Purnell, P. J. Lara, and D. Gandara, "Aurora kinases as anticancer drug targets," *Clin. Cancer Res.*, vol. 14, no. 5, pp. 1639–1648, Mar. 2008.

[18] S. Lassmann, R. Weis, F. Makowiec, J. Roth, M. Danciu, U. Hopt, and M. Werner, "Array CGH identifies distinct DNA copy number profiles of oncogenes and tumor suppressor genes in chromosomal- and microsatellite-unstable sporadic colorectal carcinomas," *J. Mol. Med.*, vol. 85, no. 3, pp. 293–304, Mar. 2007.

[19] S. Lassmann, M. Danciu, M. Müller, R. Weis, F. Makowiec, J. Schulte-Mönting, U. T. Hopt, and M. Werner, "Aurora a is differentially expressed and regulated in chromosomal and microsatellite instable sporadic colorectal cancers," *Mod. Pathol.*, vol. 22, no. 10, pp. 1385–1397, Oct. 2009.