

# A New 3D Orientation Steerable Filter

Weichuan Yu<sup>1</sup>, Kostas Daniilidis<sup>2</sup>, Gerald Sommer<sup>1</sup>

<sup>1</sup>Institut für Informatik  
Christian-Albrechts-Universität  
Preußerstraße 1–9  
D-24105 Kiel, Germany  
{wy,gs}@ks.informatik.uni-kiel.de

<sup>2</sup>GRASP Laboratory  
University of Pennsylvania  
3401 Walnut Street  
Philadelphia, PA 19104-6228, USA  
kostas@grip.cis.upenn.edu

**Abstract.** *In this paper we present a new filter based on Gaussian functions for the extraction of local 3D orientation information. Compared with current 3D steerability approaches our method achieves higher orientation resolution with lower complexity. This property enables us to solve challenging problems like complex surface analysis and multiple motion estimation. This new method decomposes a sphere with a set of overlapping basis filters which are isotropic in the feature space. We study the problem of non-uniform distribution of the spherical coordinates and discuss the application of a weighting compensation function in the computation of the 3D orientation signature. Comparisons show that our method is more efficient and robust than the 3D Hough transform.*

## 1 Introduction

In general filtering there is a conflict between performance and complexity. For example, in the orientation analysis we prefer filters having fine orientation resolution. But for this fine orientation resolution we have to consider the enormous computational complexity while constructing or rotating such filters. In order to attenuate this conflict the concept of steerability was introduced [5]. A filter  $F(\mathbf{x})$  with  $\mathbf{x} \in \mathbb{R}^n$  is referred to as a steerable filter if its deformed versions  $F_\psi(\mathbf{x})$  can be expressed as [9]

$$F_\psi(\mathbf{x}) = \sum_{k=1}^N b_k(\psi) A_k(\mathbf{x}), \quad (1)$$

where  $\psi$  denotes the deformation parameter. Here we refer to  $A_k(\mathbf{x})$  and  $b_k(\psi)$  as basis filters and interpolation functions, respectively. By using steerable filters, the responses of a given family of filters  $F_\psi(\mathbf{x})$  with  $\psi \in \mathbb{R}$  are expressed as a linear combination of *only*  $N$  basis filter responses.

While many 2D steerable filters have been applied in image processing and low level computer vision (e.g. [10, 14]), there were only a few approaches studying 3D steerability [5, 2]. Freeman and Adelson [5] were the first who introduced the concept of steerability into 3D filtering. They interpolated derivatives of 3D Gaussian functions with a set of basis filters. These basis filters are rotated copies

of the original filter. The corresponding interpolation functions are trigonometric functions of orientation parameters. Andersson designed another 3D steerable filter in the frequency domain, whose basis filters are rotated copies of the steered filter as well [2]. The drawback of both approaches is that they do not provide fine orientation resolution due to large spatial support along the angular direction.

We may overcome this drawback using filters with elongated shape [10]. One possible solution is to generalize the steerable wedge filter [14] from 2D space to 3D space, which is not yet implemented according to current literature. But current steerability is based on the global decomposition principle [16], which suffers from the consequences of the uncertainty principle. For example, we usually use  $b_k(\psi) = e^{j\omega_k\psi}$  in current steerability approaches (equation (1)). This is equivalent to sampling the angular frequency with Dirac functions. According to the well known uncertainty principle, we cannot localize one signal both in the spatial domain and in the spectral domain exactly at the same time. If we use one Dirac sampling function to localize one spectral component of the signal exactly in the frequency domain, we will no more be able to localize this component in the spatial domain. As a compensation, we need a lot of spectral Dirac impulses to increase the spatial localization ability. Correspondingly, we need a huge number of basis filters to achieve high orientation resolution, while these basis filters usually have wide supports which accentuate the computational burden.

The trade-off between spatial- and spectral-localization can be optimized only by using functions with Gaussian shape, since they achieve the lower bound in the uncertainty principle [4]. Based on this motivation, we present in this paper a new 3D steerable filter using angular Gaussian functions to achieve high orientation resolution.

Before the steerability was introduced, Bigün *et al.* connected the orientation analysis with symmetry detection using the principal axis analysis (PAA) in the mechanical engineering [6]. But their method is only suitable to detect *one* dominant orientation, since PAA only provides a set of basis vectors which are **always** orthogonal. In other words, the orientation resolution of PAA is not sufficient to solve non-orthogonal multiple orientation problem. Actually, this is the reason why we introduce steerable filters.

This paper is organized as follows: In section 2 we present the new filter and its responses of 3D planes in detail. Then we compare our filter with current 3D steerable filters and the 3D Hough transform in section 3. In section 4 we display synthetic and real application examples. In the end we conclude this paper with some discussions.

## 2 Local 3D Orientation Analysis

### 2.1 Filter Shape

In order to analyze 3D orientation naturally, we first compute a spherical mapping:  $I(x, y, z) \rightarrow I(r, \theta, \phi)$ , where  $r = \sqrt{x^2 + y^2 + z^2}$ ,  $\theta = \arctan(\frac{y}{x})$ ,  $\phi = \arctan(\frac{z}{\sqrt{x^2 + y^2}})$  (see figure 1). Since we are interested in orientation information, we build an orientation signature  $S(\theta, \phi)$  from  $I(r, \theta, \phi)$ . In order to have

fine orientation resolution, we introduce *conic kernels* as basis filters in our steerable filter. A conic kernel centered at  $(\theta_i, \phi_j)$  reads

$$B_{(\theta_i, \phi_j)}(r, \theta, \phi) := \frac{1}{\mathcal{N}_{R_{\min}, R_{\max}}^{(\theta_i, \phi_j)}(r)} G_0^{(\theta_i, \phi_j)}(\theta, \phi), \quad (2)$$

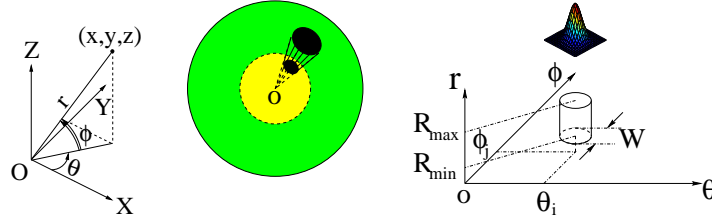
where  $\mathcal{N}_{R_{\min}, R_{\max}}^{(\theta_i, \phi_j)}(r)$  is a weighting function along the radial direction and it is independent of the angular part of the filter. We will come back to the design of  $\mathcal{N}$  later. The angular part of this basis filter is a 2D Gaussian function in the orientation space coordinated with  $(\theta, \phi)$

$$G_0^{(\theta_i, \phi_j)}(\theta, \phi) := \frac{1}{2\pi\sigma^2} e^{-\frac{(\mathcal{D}(\theta, \theta_i))^2 + (\phi - \phi_j)^2}{2\sigma^2}}, \quad (3)$$

with  $\sigma$  denotes the scale of the 2D Gaussian function. Since the angles along the  $\theta$  direction are periodic, we define a  $\mathcal{D}(\cdot)$  to represent the minimal circular difference between  $\theta$  and  $\theta_i$  ( $\theta, \theta_i \in [0, 2\pi]$ )

$$\mathcal{D}(\theta, \theta_i) := \min(|\theta - \theta_i|, |\theta - \theta_i - 2\pi|, |\theta - \theta_i + 2\pi|). \quad (4)$$

Theoretically, a Gaussian function is not compactly supported. In practice we only consider the part of  $G_0^{(\theta_i, \phi_j)}(\theta, \phi)$  inside the circular mask with a diameter  $W$ , as shown in figure 1.

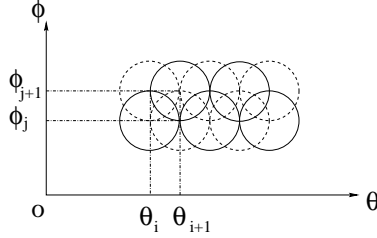


**Fig. 1.** A conic kernel centered at  $(\theta_i, \phi_j)$  with radial boundaries  $R_{\min}$  and  $R_{\max}$ . **Left:** The definition of the spherical coordinates. **Middle:** The filter kernel in the 3D Cartesian coordinates. **Right:** The filter kernel with  $\theta$ ,  $\phi$  and  $r$  as coordinates. It turns into a cylinder. In the  $(\theta, \phi)$  plane the circular mask with a diameter  $W$  is weighted by a 2D Gaussian function, as shown above the cylinder.

After applying this conic kernel on  $I(r, \theta, \phi)$ , we obtain a basis filter response as a local sample located at  $(\theta_i, \phi_j)$

$$A_{(\theta_i, \phi_j)} := \sum_{\{(\theta, \phi) | \sqrt{(\theta - \theta_i)^2 + (\phi - \phi_j)^2} \leq \frac{W}{2}\}} \sum_{r=R_{\min}}^{R_{\max}} G_0^{(\theta_i, \phi_j)}(\theta, \phi) \frac{I(r, \theta, \phi)}{\mathcal{N}_{R_{\min}, R_{\max}}^{(\theta_i, \phi_j)}(r)}. \quad (5)$$

Now let us consider the distribution of basis kernels in the  $(\theta, \phi)$  plane. It is known that a spherical surface forms a rectangular region in the  $(\theta, \phi)$  plane. For this rectangular region it is impossible to have a tessellation with circular cells. Instead, we may overlap neighboring basis kernels to cover the whole rectangular region, as shown in figure 2. In this arrangement we observe that this rectangular region is periodic along the  $\theta$  direction and is mirror-symmetric about the boundary along the  $\phi$  direction. These periodic and mirror-symmetric properties help to solve the boundary problem.



**Fig. 2.** The distribution of conic kernels in the  $(\theta, \phi)$  plane. The horizontal or vertical distance between two neighboring masks is equal to the radius of one mask.

In order to obtain the orientation signature  $S(\theta, \phi)$  from a set of samples  $A_{(\theta_i, \phi_j)}$ , we use 2D Gaussian functions with local support  $G_0^{(\theta_i, \phi_j)}(\theta, \phi)$  again as interpolation functions. Thus, the orientation signature reads

$$S(\theta, \phi) := \sum_{\theta_i} \sum_{\phi_j} A_{(\theta_i, \phi_j)} G_0^{(\theta_i, \phi_j)}(\theta, \phi). \quad (6)$$

So far, we define an analytic model of 3D orientation analysis based on Gaussian functions.

## 2.2 Filter Responses of 3D Planes

For motion estimation we are interested in filter responses of 3D planes (see section 4 for detail). In the 3D Cartesian coordinates a plane passing through the origin with a unit normal vector  $\mathbf{n} = (n_1, n_2, n_3)^T$  reads

$$xn_1 + yn_2 + zn_3 = 0. \quad (7)$$

After converting the equation into spherical coordinates we obtain an equation with variables  $\theta$  and  $\phi$

$$\cos(\phi) \cos(\phi_n) \cos(\theta - \theta_n) + \sin(\phi) \sin(\phi_n) = 0, \quad (8)$$

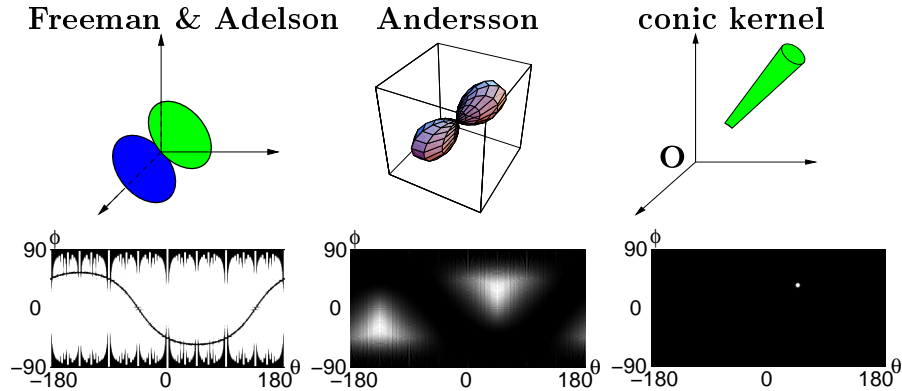
where  $\theta_n$  and  $\phi_n$  denote the desired orientation parameters of the normal vector  $\mathbf{n}$ . This equation describes a periodic curve in the  $(\theta, \phi)$  space (figure 4). In practice, we obtain a set of points in the  $(\theta, \phi)$  space. Extracting the parameters  $(\theta_n, \phi_n)$  from these points is a standard regression problem. We may apply the least square estimation (LSE) algorithm for a single plane or the expectation-maximization (EM) algorithm [15] for multiple planes.

### 3 Comparisons

#### 3.1 Comparisons with Current 3D Steerable Filters

Current 3D steerability approaches are based on the global decomposition principle using Dirac functions. In contrast, our method is based on the local decomposition principle using Gaussian functions with narrow support. This difference leads our approach to have higher orientation resolution. In figure 3 we show the first derivative of 3D Gaussian function  $G_3$  [5], Andersson’s filter [2], and our filter, respectively. Since the orientation resolution of a filter is inversely proportional to the angular support of this filter, we display the angular supports in figure 3 as well for resolution comparison. We observe that  $G_3$  has a so large angular support that only the gap between its two lobes may be useful. Andersson’s filter has smaller support than  $G_3$ . But the resolution is not yet sufficient. Compared with these two steerable filters our filter has much higher orientation resolution. This claim is confirmed in figure 5 as well.

Our filter needs a little bit more computation than the filter  $G_3$  but much less computation than Andersson’s filter (see [16] for detail due to space limitation). At first sight, our filter is less efficient than  $G_3$ . But it should be noticed that a complexity comparison is only fair, when the corresponding filters have the same orientation resolution. Since  $G_3$  cannot achieve the same orientation resolution that our filter provides, its low complexity does not make sense here.



**Fig. 3. Top Left:** The filter  $G_3$  (redrawn from [8]). **Top Middle:** Andersson’s filter in the frequency domain (redrawn from [2]). Note that the angular support of a filter in the spatial domain is the same as that in the frequency domain, since Fourier transform is an isometric mapping. **Top Right:** Our filter. **Bottom:** The corresponding filter supports in the  $(\theta, \phi)$  space. The irregularity in the support of  $G_3$  with  $\theta \in [-180^\circ, 180^\circ]$ ,  $|\phi| > 40^\circ$  is caused by the discrete representation of the filter kernel.

### 3.2 Comparisons with the 3D Hough Transform

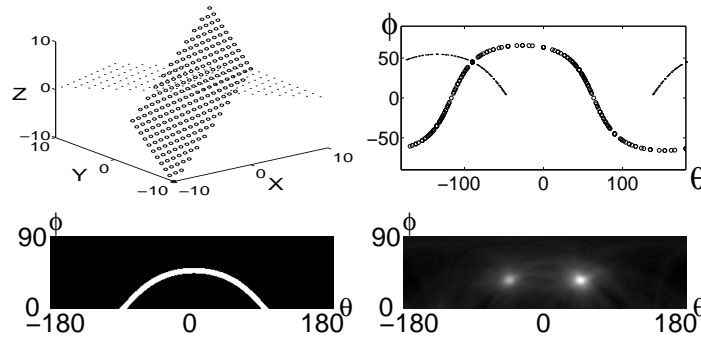
The Hough Transform is a sampling and searching method for parameter extraction. Concretely, for each point  $(I_{ix}, I_{iy}, I_{it})$  we draw all possible planes in the  $(n_1, n_2, n_3)$  space which pass through the origin and satisfy the equation

$$|I_{ix}n_{1j} + I_{iy}n_{2j} + I_{it}n_{3j}| \leq \varepsilon, \quad (9)$$

where  $\varepsilon$  denotes a positive tolerant parameter and  $(n_{1j}, n_{2j}, n_{3j})$  denotes the  $j$ -th plane. After going through all points  $(i = 1, \dots, N)$  we search in the  $(n_1, n_2, n_3)$  space the position with maximal intersection to obtain the desired parameters.

The 3D Hough transform based on equation (9) is equivalent to a 3D filter centered at the origin of the 3D space with the concave disk shape. Correspondingly, the Hough image of a point  $(I_{ix}, I_{iy}, I_{it})$  is equivalent to impulse response of the concave disk filter in 3D space, as shown in figure 4. We observe that the Hough image of a point is similar to our steerable filter response of a 3D plane except that Hough image has no negative  $\phi$  values since we use only normal vectors with  $n_3 > 0$ . Taking into account that our filter response of a plane is composed of a set of filter responses of different points, the Hough image has much larger support than our filter response of a point (figure 3). In other words, our filter samples the orientation space more efficiently than the 3D Hough transform.

Besides, in the Hough transform the search of the second maximal position is generally problematic, since we do not know how to get rid of the neighbors of the first global maximum. This problem is easier to solve in our filter response using zero-crossing analysis on the  $\theta$  axis, since the curves in the  $(\theta, \phi)$  space are periodic and we know that the zero-crossing points on the  $\theta$  axis and the extreme points with maximal  $\phi$  values contain the desired parameters. The reader is referred to [16] for details of this trick due to place limitation.



**Fig. 4. Top:** Two planes and their corresponding curves in the  $(\theta, \phi)$  space applying our filter. **Bottom Left:** Hough image of a point. It is similar to our filter response of a plane. **Bottom Right:** The Hough image of two planes disturbed with noise. We observe two mono-modal distributions. In general we do not know how to search the second maximal position automatically.

## 4 Applications

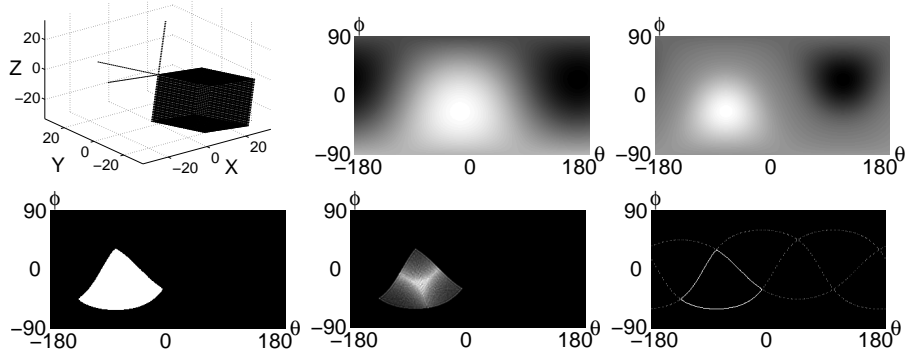
### 4.1 Compensation Issue

Before we present examples, we discuss the design of  $\mathcal{N}$  in equation (2). It is known that the horizontal angle  $\theta$  and the vertical angle  $\phi$  are defined differently in the spherical coordinates. For example, all points with the same  $\theta$  on a spherical surface lie on a great circle of this spherical surface, whereas all points with the same  $\phi$  lie on a small circle. If we divide the whole  $(\theta, \phi)$  space into a grid with equal interval, it turns out that the higher the latitude value is, the denser grid points we have on the spherical surface. This kind of non-uniform distribution was addressed in [7] in detail.

We may compensate this non-uniform distribution by designing  $\mathcal{N}$  as a *normalizing factor* so that the filter response is relatively insensitive to the non-uniform distribution. But this normalizing factor  $\mathcal{N}$  “strengthens” then the outputs of filters with a few points and “suppresses” the outputs of filters with many points. As a result, we are no more able to know the real distribution density of points in the  $(\theta, \phi)$  space, while this density information is desirable in some applications. For example, we use the EM algorithm for multiple motion estimation. The philosophy behind the EM algorithm is that there are more “normal” points than noise and “incorrect” samples and the distribution density works as a weighting function in the parameter regression procedure. If we lose such density information, the estimation result will be much worse. For this reason, we would like to study experimentally the susceptibility of the filter response without normalizing compensation to the non-uniform distribution.

### 4.2 3D Junction Characterization

We begin with an example of 3D junction characterization. In figure 5 we have a cubic with one of its vertices as keypoint. For comparison we apply the steerable filter  $G_3$ , Andersson’s filter, our filter with compensation (setting  $\mathcal{N}$  as the sum of discrete weights in the filter mask), and our filter without compensation (setting  $\mathcal{N}$  as a constant). In the response of  $G_3$ , the location of the maximal value does not have geometrical meaning since the angular support of  $G_3$  is too large to interpret this 3D junction. Andersson’s filter has higher orientation resolution than  $G_3$ . Though the edges of the cubic are blurred, the location of maximal value in the response corresponds to the center of the cubic. Compared with these two steerable filters, our approach provides evidently higher orientation resolution. Besides, a comparison shows that the response of our filter without compensation  $S(\theta, \phi)$  is more sensitive to the non-uniform distribution than that with compensation  $S_a(\theta, \phi)$ . But this susceptibility does not obstruct us from obtaining main structure information in the orientation signature. Thus, we still can use  $S(\theta, \phi)$  for 3D junction characterization.



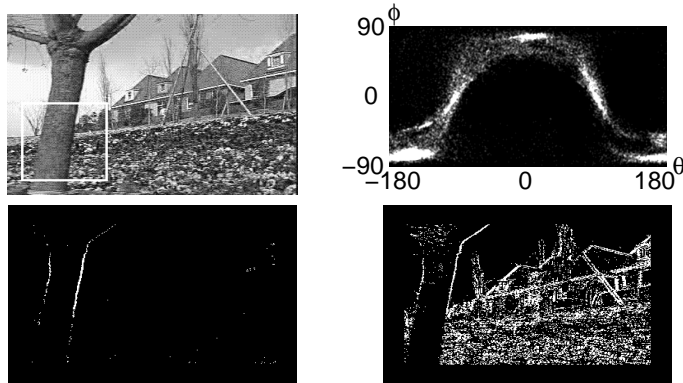
**Fig. 5. Top Left:** 3D plot of a cubic with its three normal vectors at one vertex. **Top Middle:** The response of  $G_3$ . **Top Right:** The response of Andersson's filter. **Bottom Left:** Orientation signature with averaging  $S_a(\theta, \phi)$ . Three surfaces turn out to be three edges in the  $(\theta, \phi)$  space. **Bottom Middle:** Orientation signature without averaging  $S(\theta, \phi)$ . Compared with  $S_a(\theta, \phi)$  it is more sensitive to the non-uniform distribution of points in the  $(\theta, \phi)$  space. However, the edges are still clearly represented. **Bottom Right:** We can extract edge information from  $S(\theta, \phi)$  applying morphological operations. For comparison we also display three surfaces connected with the keypoint using dotted curves. They are consistent with the extracted edges.

### 4.3 Multiple Motion Estimation

It is proven that a single translational motion corresponds to a single plane [1, 8] and multiple motions correspond to multiple planes [13, 3] passing through the origin in the derivative space coordinated with  $(I_x, I_y, I_t)$  or in the frequency space. In both cases, the normal vectors of planes contain the desired motion parameters. Thus, motion estimation turns out to be orientation analysis of planes [1, 6]. For occlusion we have not only multiple planes in the  $(I_x, I_y, I_t)$  space, but also distortions which disturb the orientation estimation [3]. In general the number of distortion points is much less than that of plane components. We can therefore reduce the disturbance of distortions by using the distribution density as weighting function in the estimation, i.e. by using  $S(\theta, \phi)$ .

Figure 6 shows a real example. The flower garden sequence contains a left moving trunk occluding a left moving background. This can be observed as two curves in the orientation signature  $S(\theta, \phi)$ . After applying the EM algorithm based on equation (8) we obtain parameters of two motions. Before our approach, Huang and Chen used the gap of  $G_3$  (see figure 3) to fix the orientation of *one* plane [8]. This was the single one approach using 3D steerable filter for motion estimation. But this method works only for single motion estimation due to the coarse resolution. Observing the spatial coherence on either side of the occlusion boundary, we further use the “warp-and-subtract” technique [11, 3] to segment two motions. The result is shown in figure 6 as well.





**Fig. 6. Top Left:** One frame from the flower garden sequence. The white box shows us the window for motion estimation. **Top Right:** Two curves in the  $S(\theta, \phi)$ . **Bottom:** The segmentation results using the “warp-and-subtract” technique [11, 3]. Here we do not consider the boundary problem in warping. Since there is no difference inside the regions with the aperture problem before and after the warping, we observe only the boundaries of the trunk.

## 5 Discussion

Our original motivation is to improve the orientation resolution of current 3D steerable filters. It is interesting to observe that our approach is related to the 3D orientation histogram (OH). Both methods achieve high orientation resolution and both methods decompose the sphere locally. But there are still differences between them.

- The 3D OH is a discrete approximation of the extended Gaussian image [7]. Our approach provides an analytic model for 3D orientation analysis.
- The 3D OH works on a unit spherical surface. Our approach projects the sphere onto the  $(\theta, \phi)$  space. Though after this non-isometric mapping we lose the rotational symmetry, we gain easier structure representation and post-processing as compensation. For example, on the surface of this paper sheet, using 3D OH we cannot display all parts of a great circle of a sphere, while using our approach we can, though with some deformations.
- The 3D OH is applied for surface analysis of convex objects and it is shift- and scale-invariant. Our 3D filter can be applied not only for surface analysis, but also for volume data analysis. It treats both convex and concave objects. But we must fix the keypoint and the radial boundaries at first.
- The basis cells in the 3D OH are not *isotropic*. Besides, they have either the same round shape or high resolution, but not both simultaneously [7]. Our approach provides *isotropic* cells in the feature space satisfying both criteria simultaneously. But we have to overlap our basis cells to have this property.

For further research we may explore if there exists a dual basis of our non-orthogonal basis filters. The lifting scheme [12] may be helpful in this study.

This may further provide a possible cue to solve the still open problem of tessellation/decomposition of the spherical surface with isotropic cells.

## References

1. E. H. Adelson and J. R. Bergen. Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America*, 1(2):284–299, 1985.
2. M. T. Andersson. *Controllable Multidimensional Filters and Models in Low Level Computer Vision*. PhD thesis, Department of Electrical Engineering, Linköping University, Linköping, Sweden, 1992.
3. W. Yu, K. Daniilidis, S. Beauchemin, and G. Sommer. Detection and characterization of multiple motion points. In *IEEE Conf. Computer Vision and Pattern Recognition*, volume I, pages 171–177, Fort Collins, CO, June 23–25, 1999.
4. J. G. Daugman. Uncertainty relation for resolution in space, spatial frequency and orientation optimized by two-dimensional visual cortical filters. *Journal of the Optical Society of America*, 2(7):1160–1169, 1985.
5. W.T. Freeman and E.H. Adelson. The design and use of steerable filters. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 13:891–906, 1991.
6. J. Bigün, G. H. Granlund and J. Wiklund. Multidimensional orientation estimation with application to texture analysis and optical flow. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 13(8):775–790, 1991.
7. B. K. P. Horn. *Robot Vision*. MIT Press, 1986.
8. Chung-Lin Huang and Yng-Tsang Chen. Motion estimation method using a 3d steerable filter. *Image and Vision Computing*, 13:21–32, 1995.
9. M. Michaelis and G. Sommer. A Lie group approach to steerable filters. *Pattern Recognition Letters*, 16:1165–1174, 1995.
10. P. Perona. Deformable kernels for early vision. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 17(5):488–499, 1995.
11. M. Irani, B. Rousso, and S. Peleg. Computing occluding and transparent motions. *International Journal of Computer Vision*, 12:5–16, 1994.
12. P. Schröder and W. Sweldens. Spherical wavelets: Efficiently representing functions on the sphere. *Computer Graphics Proceedings (SIGGRAPH 95)*, pages 161–172, 1995.
13. M. Shizawa and K. Mase. A unified computational theory for motion transparency and motion boundaries based on eigenenergy analysis. In *IEEE Conf. Computer Vision and Pattern Recognition*, pages 289–295, Maui, Hawaii, June 3–6, 1991.
14. E. P. Simoncelli and H. Farid. Steerable wedge filters for local orientation analysis. *IEEE Trans. Image Processing*, 5(9):1377–1382, 1996.
15. Y. Weiss and E. H. Adelson. A unified mixture framework for motion segmentation: Incorporating spatial coherence and estimating the number of models. In *IEEE Conf. Computer Vision and Pattern Recognition*, pages 321–326, San Francisco, CA, June 18–20, 1996.
16. W. Yu. *Local Orientation Analysis in Images and Image Sequences Using Steerable Filters*. PhD thesis, Institute of Computer Science, University Kiel, Germany, 2000.