# Junction Characterization Using Polar Pyramid

Weichuan Yu, Kostas Daniilidis, Gerald Sommer

Christian–Albrechts–Universität
Lehrstuhl für Kognitive Systeme
Preußerstraße 1–9, 24105 Kiel
Email: wy@informatik.uni-kiel.de

**Abstract.** In this paper we present a new approach in characterizing gray-value junctions. Due to the multiple intrinsic orientations present in junctions the response of a filter is needed at every orientation. As a rotation of the filter would considerably increase the computational burden alternative techniques like filter steerability have been proposed. Steerability relies in interpolating the response at an arbitrary orientation from the responses of some basis filters. Unfortunately, current steerability approaches suffer from the consequences of the uncertainty principle: In order to achieve high selectivity in orientation they need a huge number of basis filters increasing, thus, the computational complexity.

The new approach presented here achieves a higher orientational selectivity with a lower complexity. We consider the local polar map of the neighborhood of a junction where the new coordinates are the radius and the angle. Finding the gray-value transitions of a junction can be interpreted as 1D edge detection. Hence, the orientational selectivity problem can be attacked by applying a pyramidal scheme. It is well known that it is always possible to reconstruct a signal using the sampling kernel as an interpolation function. Therefore, our approach can also steer the response of a Gaussian derivative to every orientation. The total algorithmic complexity encompasses the small support 2D-filtering for polar mapping and radial smoothing plus an 1D-differentiation.

## 1 Introduction

Junctions of gray-value lines or edges are rare events in images carrying important information for many image processing tasks like point matching in object recognition, point tracking in motion analysis, and attentive coding.

In order to utilize junctions for such tasks we must be able to locate their corresponding keypoints (defined as meeting points of lines or edges), to characterize them by means of signatures and to classify them in junction categories. Regarding keypoint detection and localization the reader is referred to Förstner's study [5] and to the comparison of different operators by Rohr [10, 11]. In this paper we address the problem of junction characterization. The resulting signature can be used for further junction classification.

Junctions are local structures with multiple intrinsic scales and orientations [1]. A signature characterizing such a junction can be only obtained by applying

a filter at different scales and orientations around the keypoint. This results in an enormous computational complexity. In order to attenuate this burden the concept of steerability has been introduced [6, 9]. Denoting the deformation parameter of a filter with $\alpha$ we define a filter $F(\boldsymbol{x})$ with $\boldsymbol{x} \in I\!\!R^n$ as a steerable filter if its deformed versions $F_\alpha(\boldsymbol{x})$ can be expressed as [7]:

$$F_\alpha(\boldsymbol{x}) = \sum_{k=1}^{N} b_k(\alpha) A_k(\boldsymbol{x}) \tag{1}$$

where $A_k(\boldsymbol{x})$ and $b_k(\alpha)$ are called basis filters and interpolation functions, respectively. Applying such a steerable filter on an image $I(\boldsymbol{x})$ yields:

$$\langle F_\alpha(\boldsymbol{x})|I(\boldsymbol{x})\rangle = \sum_{k=1}^{N} b_k(\alpha)\langle A_k(\boldsymbol{x})|I(\boldsymbol{x})\rangle \tag{2}$$

where $\langle\cdot|\cdot\rangle$ is the usual scalar product. In this way we can reduce the computational cost of applying a whole set of different $F_\alpha(\boldsymbol{x})$ to $N$ scalar products.

For junction characterization many kinds of steerable filters are applied like derivatives of the 2D Gaussian filter [6], elongated Gaussian-derivative kernels [9], the double Hermite function [7] and the steerable wedge filter [12]. All these steerable filters adopt trigonometric functions or complex exponentials of $\alpha$ as interpolation functions $b_k(\alpha)$ (equation 1) such that the basis functions are rotated copies of the original filter or can be solved using Fourier theory.

Steerability provides us with a solid mathematical theory. However, the complexity of its implementation remains high. Due to the uncertainty principle the product of orientational resolution of a steerable filter and its orientational bandwidth has a lower bound. Therefore, in order to achieve a high orientational selectivity we have to apply a huge number of basis filters [7, 9] with large spatial support.

In our previous work [4] we have presented a new approach to characterize junctions. It is based on applying rotated copies of a wedge averaging filter and estimating the derivative with respect to the polar angle. This averaging filter is polar separable. Its radial and angular components are rectangle functions expressed by following and shown in figure 1:

$$p(\rho) = \begin{cases} 1 & R_{min} \leq \rho \leq R_{max} \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

with $R_{min}$ and $R_{max}$ as its boundaries in radial direction and

$$g_{\theta_i}(\theta) = \begin{cases} 1 & \mathcal{D}(\theta, \theta_i) \leq \frac{W}{2} \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

where $\theta, \theta_i \in [0, 2\pi]$, $\theta_i$ is the center of the filter and $W$ is the angular width of the mask. Since $\theta$ and $\theta_i$ are circular angles, we define a $\mathcal{D}(\ )$ to represent the minimal absolute value of their difference.

In this paper we will observe this simple averaging filter from a more interesting point of view and extend it into a pyramidal scheme. The averaging filter

is nothing else than the kernel required to conduct a local transformation in polar coordinates (Fig. 1). However, this polar transformation involves a filtering: We smooth and subsample the angular direction and keep just the average of the radial direction. This will allow us to treat oriented structures as structures of a 1D-signal. To treat the orientational selectivity problem we choose a pyramidal representation. The computational burden is much lower than usual Fourier-domain schemes because the filters have smaller supports and part of them act only one-dimensionally. We also show experimentally the superiority of our scheme in many real images.
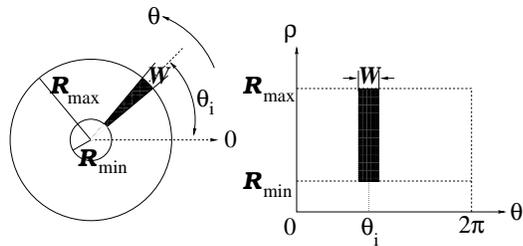


**Fig. 1.** Mask centered at angle $\theta_i$. **Left:** Original mask with keypoint at the center of the circle. **Right:** Mask with $\theta$ and $\rho$ as Cartesian coordinates, $R_{max}$ and $R_{min}$ are radial boundaries of the mask, $W$ is the angular width of the mask.

## 2   The Polar Pyramid

In the following, we will apply an averaging filter in the radial direction since we do no treat spatial scale in this paper. In the angular direction we replace the averaging filter with a Gaussian function:

$$G_{\theta_i}(\theta) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(\mathcal{D}(\theta,\theta_i))^2}{2\sigma^2}} \tag{5}$$

where $\sigma$ is the scale of the Gaussian function, $\theta_i$ and $\mathcal{D}(\ )$ are the same as in equation 4.

We will later see that the choice of a Gaussian in establishing the local polar map has many nice effects. Since we want to use the Gaussian function as a FIR-filter we must cut off its support. It is easy to show that in order to keep the energy of the cut-off area below 1% of the total energy the width of the mask must be at least $5\sigma$. According to the sampling theory we can sample the angular space with an interval of $\sigma$. Thus, even taking the overlapping of neighboring masks into account we need to cover the whole neighborhood only a few times to obtain the required information.

As we already mentioned in the introduction the main motivation of the proposed approach is the appropriate treatment of the orientational selectivity. The local polar map enables us to look at this problem as a 1D gray value transition description. The Fourier-based steerability approaches approximate a filter with a Fourier series expansion with respect to the angle. The more Fourier-coefficients and basis functions they use the more selective they are in discerning

fine gray value transitions in the polar map. However, in order to treat angular transitions of varying scale we introduce a pyramid in the polar domain.

It is well known that one of the most appealing kernels for hierarchical approaches is the Gaussian function [2]. The Gaussian function achieves the lowest bound of the product between spatial support and bandwidth. Burt and Adelson [3] prove that the generating kernel of subsampling can be used as an interpolation function. Moreover, they argue that the interpolation functions can be discrete approximations of Gaussian functions with different scales (figure 2). Thus, the continuous orientational information can be reconstructed from all levels of the polar pyramid by interpolating the pyramid elements $P_j(\theta)$ with Gaussian functions of different scales $G_j(\theta)$:

$$\hat{f}_j(\theta) = \sum_n \delta(\theta - n\triangle\theta)P_j(\theta) * G_j(\theta) \qquad j \in [1, 2, \cdots] \qquad (6)$$

with

$$G_j(\theta) = \frac{1}{\sqrt{2\pi}\sigma_j}e^{-\frac{\theta^2}{2\sigma_j^2}} \qquad j \in [1, 2, \cdots] \qquad (7)$$

where $\hat{f}_j(\theta)$ represents the reconstructed signal from the j-th level.

Because in natural images edges are more important than lines we need to estimate the derivative of $\hat{f}_j(\theta)$ to obtain the required information:

$$D_j(\theta) = | \frac{d}{d\theta}\hat{f}_j(\theta) | \qquad j \in [1, 2, \cdots] \qquad (8)$$

The local maxima in $\hat{f}_j(\theta)$ and $D_j(\theta)$ denote orientations of lines and edges at different levels, respectively.

If we want to build a pure 1D octave Gaussian Pyramid of angles, according to [3] we should have $M2^K + 1$ samples, where $K$ is the number of levels and $M+1$ is the number of samples at the highest level. Taking the periodic property into account we should have $M2^K$ samples as the smoothing outputs. However, usually 360 angular values are used. Observing that $360 = 2^3 \times 3^2 \times 5$ we change the subsampling interval slightly. For instance, for three levels we subsample every two points and for two levels every three.

According to [3] the generating kernel should be normalized, symmetric and unimodal and it should make equal contribution to construct the next higher level. The generating kernels satisfying these constraints for subsampling with factor 2, 3 and 5 can be expressed as following, respectively:

$$S_2 = \frac{1}{16} \left( 1\ 4\ 6\ 4\ 1 \right) \qquad (9)$$

$$S_3 = \frac{1}{264} \left( 3\ 22\ 66\ 82\ 66\ 22\ 3 \right) \qquad (10)$$

$$S_5 = \frac{\left( 1\ 74\ 299\ 725\ 950\ 1022\ 950\ 725\ 299\ 74\ 1 \right)}{5120} \qquad (11)$$
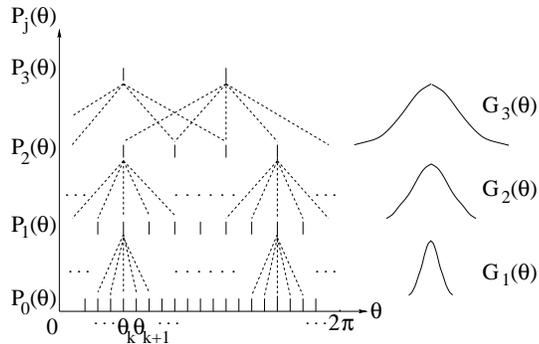
**Fig. 2. Left:** Polar pyramid structure, $P_0(\theta)$ is the output of smoothing Gaussian filter. $P_j(\theta)(j = 1, 2, \cdots)$ are higher levels after subsampling with generating kernels $S_2$ or $S_3$. **Right:** Corresponding interpolation functions at different levels. They are Gaussian functions with different scales.

# 3 Junction Characterization Examples and Discussions

## 3.1 Junctions without Scale Variations

For junctions without orientational scale variations the smoothing outputs $P_0(\theta_i)$ and the absolute values of its derivative $D_0(\theta_i)$ ($\theta_i = 0°, \cdots, 359°$) are adequate for characterizing junctions.

In figure 3 and 4 are examples of synthetic line junctions and edge junctions. The corresponding $P_0(\theta_i)$ and $D_0(\theta_i)$ characterize them correctly. The improvement of replacing the angular averaging filter by a Gaussian smoothing filter is shown in figure 3. The small deviations of orientations in figure 4 are due to the fact that an edge can be only presented by two pixels in the grid, while we can not set the center of a mask between two pixels.

The robustness of our method against keypoint offset is shown in figure 5. A real example is in figure 6. We present results using the steerable wedge filter [12], the averaging method and the polar pyramid scheme. It is clear to see that the polar pyramid reaches the best orientational resolution. In comparison to the averaging method we improve robustness with a moderate increase of complexity which still remains considerably lower than the steerability method.

## 3.2 Solving Orientational Scale Problem

In figure 7 we show how to solve the orientational scale problem by applying the polar pyramid. Each wide line of the junction is best characterized in $P_2(\theta)$ of the third level with an orientational resolution of $2\sigma_3 = 16°$, while at lower or higher levels the junction is not so discernible. In the future work we further need a criterion to judge the quality of reconstructed information analytically.

The result of a child's left eye corner is presented in figure 8. The corner can be regarded as a combination of irregular wide lines and blurred edges disturbed by noise. Our method provides suitable information at different orientational scales for further interpretation.
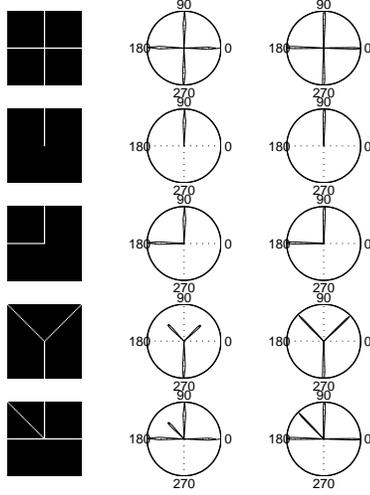
**Fig. 3. Left:** Synthetic line junction images with the size of $65 \times 65$ pixels and line width of 1 pixel. **Middle:** *Averaging outputs* using the rectangle averaging filter defined in equation 4 with $W = 6°$, $R_{min} = 3$, $R_{max} = 15$. The numbers around circles indicate angles in degree. **Right:** Corresponding *Smoothing outputs* $P_0(\theta)$ using the Gaussian smoothing filter with same parameters. Both filters show good performances. But the normalization of the Gaussian filter is better (row 4 and 5).
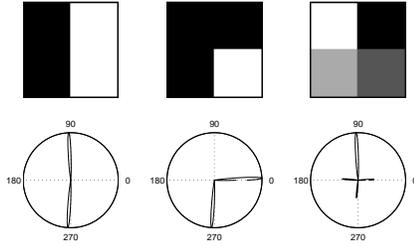


**Fig. 4. Top:** Synthetic edge junctions. **Bottom:** Polar plots of $D_0(\theta)$. The local maxima show the orientation of edges. $R_{min} = 3$, $R_{max} = 15$.
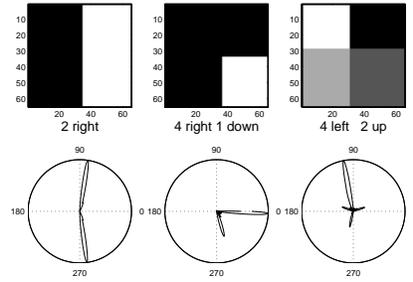
**Fig. 5. Top:** Different deviation of keypoints from centers of masks. **Bottom:** Corresponding $D_0(\theta)$.

## 4   Conclusion

The steering of a filter to different orientations enables the characterization of junctions. To avoid rotating a filter at all orientations the theory of steerability allows the interpolation of a filter response at every orientation from a limited set of basis filters. In Fourier-based steerability approaches it turns out that the number of basis filters is proportional to the angular bandwidth, which has an inverse relationship with the orientational selectivity. In practice, to characterize fine structures we need a huge number of basis filters with very large angular support. A second fundamental fact is that intrinsically the Fourier-based methods involve a local polar transformation which maps a filter rotation to a filter translation [8].

We proposed in this paper the explicit implementation of the polar transformation which allows a further 1D-treatment of the orientation problem just in the angular direction. Instead of using a Fourier series expansion we make use of the pyramidal processing theory and introduce the Gaussian as the interpolating function. Simultaneously, the Gaussian function and its derivative are the generating kernels of the pyramid. The hierarchical treatment enables us to characterize both fine and course transitions in gray-value junctions. Experimental results show the superiority of the hierarchical approach over a classic steerability scheme and a similar polar averaging filter in one scale.

## Acknowledgment

## References

1. D. H. Ballard and L. E. Wixson. Object recognition using steerable filters at multiple scales. In *IEEE Workshop on Qualitative Vision*, pages 2–10, 1993.
2. J. Babaud , A. P. Witkin , M. Baudin and R. O. Duda. Uniqueness of the Gaussian kernel for scale-space filtering. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 8:26–33, 1986.
3. P. J. Burt and E. H. Adelson. The Laplacian pyramid as a compact image code. *IEEE Transactions on Communications*, 31:532–540, 1983.
4. W. Yu, K. Daniilidis and G. Sommer. Rotated wedge averaging method for junction characterization. In *the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'98), Santa Barbara, June 23-25*, 1998.
5. W. Foerstner. A framework for low level feature extraction. In *European Conf. on Computer Vision*, volume II, pages 383–394, Stockholm, Sweden, May 2-6, J.O. Eklundh (Ed.), Springer LNCS 801, 1994.
6. 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.
7. M. Michaelis and G. Sommer. Junction classification by multiple orientation detection. In *European Conf. on Computer Vision*, volume I, pages 101–108, Stockholm, Sweden, May 2-6, J.O. Eklundh (Ed.), Springer LNCS 800, 1994.
8. M. Michaelis and G. Sommer. A Lie group approach to steerable filters. *Pattern Recognition Letters*, 16:1165–1174, 1995.
9. P. Perona. Deformable kernels for early vision. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 17(5):488–499, 1995.
10. K. Rohr. Recognizing corners by fitting parametric models. *International Journal of Computer Vision*, 9(3):213–230, 1992.
11. K. Rohr. On the precision in estimating the location of edges and corners. *Journal of Mathematical Imaging and Vision*, 7:7–22, 1997.
12. E. P. Simoncelli and H. Farid. Steerable wedge filters for local orientation analysis. *IEEE Trans. Image Processing*, 5(9):1377–1382, 1996.
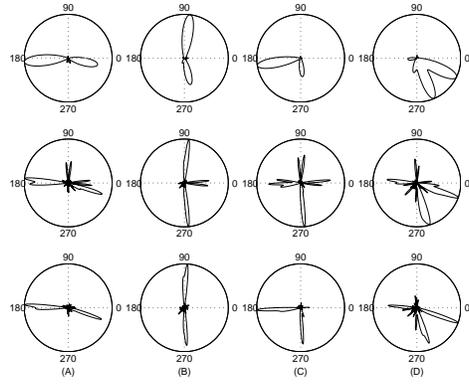
**Fig. 6. Left:** Parkbench with edge junctions. $A$: horizontal edge; $B$: vertical edge; $C$: corner; $D$: 'T' junction. **Right Row 1:** Steerable wedge filter results using 30 basis filters with 19-tap size. **Right Row 2:** Derivative Outputs using angular averaging filter. The averaging filter is sensitive to high frequency component. **Right Row 3:** Derivative Outputs using Gaussian smoothing filter. The result is better than using averaging filter. $R_{min} = 3, R_{max} = 9$.
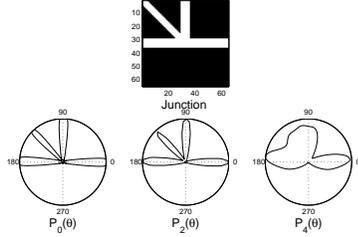


**Fig. 7. Top:** A junction composed of lines with a width of 7 pixels. **Bottom:** orientational signals reconstructed from the first, third and fifth level of the polar pyramid. Each wide line is best characterized with only one maximum at the third level $(P_2(\theta))$, while at even larger scale $(P_4(\theta))$ the structure is hardly discernible.
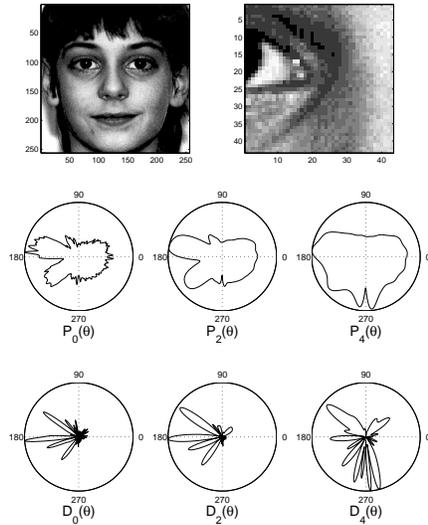


**Fig. 8. Top left:** Face of a child. **Top right:** Its left eye corner in detail. It can be regarded as combination of irregular wide lines and blurred edges disturbed by noise. **Middle:** orientational signals reconstructed from the first, third and fifth level of the polar pyramid. The eyelids are clear to see at the third level as two maxima at $135°$ and $225°$. The local maximum near $180°$ is due the white of the eye. **Bottom:** Corresponding $D_j(\theta), j = 0, 2, 4$.