An accurate and contrast invariant junction detector

Gui-Song Xia
CEREMADE CNRS, Univ. Paris-Dauphine

Julie Delon, Yann Gousseau
LTCI CNRS, TELECOM ParisTech

Abstract

This paper introduces a generic method for the accurate analysis of junctions, relying on a statistical modeling of normalized image gradients. We analyze junctions as local visual events that do not happen by chance under a background model derived from the a-contrario *methodology. The method not only provides thresholds for the detection of junctions, but also enables their accurate characterization, including a precise computation of their type, localization, scale and geometrical configuration. The efficiency of the method is evaluated through various experiments.*

1. Introduction

Local geometrical structures of images, e.g. junctions, play key roles in scene understanding. They have been widely used for various computer vision tasks, including stereo vision, figure/ground separation, segmentation and object recognition. The automatic detection of junctions has been a longstanding and active research topic in image analysis and computer vision. Classical approaches include Harris detectors [1], SUSAN detector [2] and their extensions [3, 4]. Such methods usually neither distinguish different types of junctions nor identify their local geometrical configurations. Such detectors are therefore often used for computing key points but they offer little geometrical information for occlusion analysis or figure/ground separation. These detectors usually rely on the linear scale-space [5] to compute the scale of a junction, implying a loss in precision. Alternatively, junctions can also be detected by grouping edges or contours, see [6], but boundaries are usually imprecise around junctions. Among existing approaches, the most adapted for accurate local junction analysis are template-based ones [7, 8, 9]. However they often heavily depend on the contrast of images, only detect junctions with fixed scale and rely on many sensitive parameters.

This paper is along the axis of template-based models and takes interest in a junction detector meeting the following requirements: (1) contrast invariant detection, meaning that the detection of junctions is independent of changes of image contrast; (2) automatic setting of detection parameters. (3) accurate detection, implying that the position and scale of the junctions should be detected precisely and be clearly related to the local image geometry, which is in strong contrast with the classical detectors used for keypoint detection, such as Harris [1]. We handle these problems by relying on the *a-contrario* methodology [10], in which structures are validated by controlling a number of false detections. A key advantage of this approach is the automatic setting of detection parameters in a way that will prevent excessive detection of junctions in textured areas. The accuracy of detection will be achieved thanks to a competition between junctions using the *number of false alarms*, a quality measure associated with a junction. An extended version of this work is presented in the technical report [11].

2. Geometrical junction representation

This section first introduces a contrast invariant and geometrical representation of a junction and then provides a naive detection algorithm. We consider gray-scale images I on discrete image lattice Ω and denote the gradient as $\nabla I = (I_x, I_y)$. We define the *direction* $\phi(\mathbf{q})$ of a pixel $\mathbf{q}$ as the direction perpendicular to the gradient at $\mathbf{q}$, *i.e.* $\phi(\mathbf{q}) = (\arctan \frac{I_y(\mathbf{q})}{I_x(\mathbf{q})} + \frac{\pi}{2}) modulo (2\pi)$. In what follows, we define the geometrical configuration of a junction in an image I and precise its quantitative measure.

Geometrical junction template A junction $\jmath$ is represented by its center $\mathbf{p}$, its scale $r \in \mathbb{N}$ and a set of M branch directions $\{\theta_1, \dots \theta_M\}$ around $\mathbf{p}$. M is the order of the junction. We consider $M = 2, 3$ and 4, corresponding to corners, T-(or Y-) and X-junctions. An example with $M = 3$ is shown in Figure 1. In a discrete image, the set of directions $\mathcal{D}(r)$ at a given scale r is a collection of $K(r)$ directions, as detailed in Section 4.

For a given scale r, the *branch* of direction θ at location $\mathbf{p}$ is described by a disk sector $S_{\mathbf{p}}(r, \theta)$ with an angular aperture of $2\Delta(r)$, inside which the length of

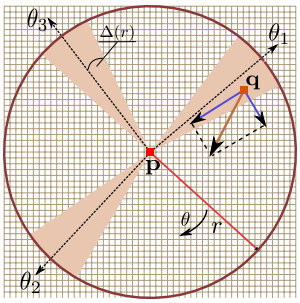

Figure 1: A junction with three branches.

the vector $\overrightarrow{\mathbf{pq}}$ is not larger than r and the angle between $\overrightarrow{\mathbf{pq}}$ and the direction θ is less than $\Delta(r)$. It is natural to require that two branches of a given junction do not intersect. It follows that the angle between two directions in a junction $\jmath$ must be larger than $2\Delta(r)$.

Contrast invariant junction strength We use the image gradient to derive the strength of a junction. For robustness to local contrast changes, we normalize each gradient component I_x and I_y by dividing it by its average gradient magnitude in a small neighborhood of size $n \times n$ (in this paper, n set to 5), and denote it as $\widetilde{\nabla} I$. We define the strength of a branch $S_{\mathbf{p}}(r, \theta)$ as

$$\omega_{\mathbf{p}}(r, \theta) = \sum_{\mathbf{q} \in S_{\mathbf{p}}(r,\theta)} \gamma_{\mathbf{p}}(\mathbf{q}), \tag{1}$$

where $\gamma_{\mathbf{p}}(\mathbf{q}) = \|\widetilde{\nabla} I(\mathbf{q})\| \cdot (|\cos(\phi(\mathbf{q}) - \alpha(\overrightarrow{\mathbf{pq}}))| - |\sin(\phi(\mathbf{q}) - \alpha(\overrightarrow{\mathbf{pq}}))|)$, if the difference between α and ϕ is larger than $\pi/4$, and $\gamma_{\mathbf{p}}(\mathbf{q}) = 0$ otherwise. It measures how well the corresponding angular sector agrees with local orientations. The strength of a junction is defined as the minimal strength of its branches

$$t(\jmath) := \min_{m=1\ldots M} \omega_{\mathbf{p}}(r, \theta_m). \tag{2}$$

Junction proposals So far, for a fixed scale r, we can derive a naive algorithm for junction detection by selecting all the junctions in I with a strength greater than a given threshold T. For a given location $\mathbf{p}$, testing all possible junctions is computationally too heavy, and we only consider directions θ corresponding to local maxima over $[\theta - \Delta(r), \theta + \Delta(r)]$. These local maxima can be efficiently computed through using a non-maximum suppression procedure.

The limitation of this detection algorithm is that the threshold T on the junction strength is fixed regardless of the detection scale and order and of the image size. Setting such a threshold globally is difficult and often leads to over- or under-detections. We now provide a criterion for automatically setting the threshold T, by relying on an *a-contrario* methodology.

3. An a-contrario junction (ACJ) model

The *a-contrario* detection theory, introduced in [10], yields a generic method for validating visual events in images. This section applies this methodology to the detection of junctions.

3.1. Null hypothesis $\mathcal{H}_0$

In the *a-contrario* methodology, meaningful events are those whose probability of occurrence under some null hypothesis $\mathcal{H}_0$ is small enough. In our context, we define the null hypothesis as follows. Denoting $\mathbf{I}$ a random image, $\|\widetilde{\nabla}\mathbf{I}(\mathbf{q})\|$ and $\phi_{\mathbf{I}}(\mathbf{q})$ the random variables corresponding to the contrast and orientation at each pixel $\mathbf{q}$, we say that these variables follow the null hypothesis $\mathcal{H}_0$ if (1) $\forall \mathbf{q} \in \Omega$, $\|\widetilde{\nabla}\mathbf{I}(\mathbf{q})\|$ *follows a Rayleigh distribution with parameter* 1; (2) $\forall \mathbf{q} \in \Omega$, $\phi_{\mathbf{I}}(\mathbf{q})$ *is uniformly distributed over* $[0, 2\pi]$; (3) *the family* $\{\|\widetilde{\nabla}\mathbf{I}(\mathbf{q})\|, \phi_{\mathbf{I}}(\mathbf{q})\}_{\mathbf{q}\in\Omega}$ *is made of independent random variables*.

Remark It is worth noticing that the first assumption is based on the statistics of locally normalized gradient [12, 11]: except on a small neighborhood around 0, the $\|\widetilde{\nabla} I\|$ in natural images closely follows a Rayleigh distribution of parameter 1.

3.2. Distribution of junction strength

Let $\jmath$ be a junction in a random image $\mathbf{I}$ where normalized gradients and directions are assumed to follow $\mathcal{H}_0$. Since branches do not intersect, the strengths $\omega_{\mathbf{p}}(r, \theta_m)$ of the different branches are independent random variables. According to Equation (1) and (2), if we note $\mathbf{t}(\jmath)$ the random variable measuring the strength of $\jmath$, then $\mathbb{P}_{\mathcal{H}_0}[\mathbf{t}(\jmath) \geq t] = \prod_{m=1}^{M} \mathbb{P}_{\mathcal{H}_0}[\omega_{\mathbf{p}}(r, \theta_m) \geq t]$. Next, the strength of each branch, $\omega_{\mathbf{p}}(r, \theta_m)$, is a sum of i.i.d. random variables $\gamma_{\mathbf{p}}(\mathbf{q})$, the distribution of which under $\mathcal{H}_0$ can be written (see [11]) as

$$\mu(z) = \tfrac{1}{2}\delta_0(z) + H(z) \cdot \tfrac{1}{\sqrt{\pi}} e^{-\frac{z^2}{4}} \cdot \mathrm{erfc}(\tfrac{z}{2}) dz,$$

with δ_0 a Dirac mass at 0, $H(z) = 1$, for $z \geq 0$, and $H(z) = 0$ otherwise, and erfc is defined as $\mathrm{erfc}(z) = \frac{2}{\sqrt{\pi}} \int_z^{\infty} e^{-s^2} ds$.

Finally, under the hypothesis $\mathcal{H}_0$, the probability that the random variable $\mathbf{t}(\jmath)$ is larger than a given threshold t is

$$F_{\jmath}(t) := \mathbb{P}_{\mathcal{H}_0}[\mathbf{t}(\jmath) \geq t] = \prod_{m=1}^{M} \int_t^{+\infty} \star_{j=1}^{J(r,\theta_m)} \mu(dz), \tag{3}$$

where $J(r, \theta_m)$ is the size of a sector of orientation θ_m at scale r.

3.3. Meaningful junctions

Following the *a-contrario* methodology, the thresholds are then set by bounding the average number of

false detections under $\mathcal{H}_0$. This number is obtained by multiplying the probability (3) by the number of possible junctions in the image.

Number of tests Let $\mathcal{J}(M)$ be the set of all possible junctions of order M in an image I with N pixels. At a given location $\mathbf{p}$, once the first branch is chosen among the $K(r)$ possible directions, only $K(r)(1-2\frac{\Delta(r)}{\pi})$ possible directions remain for the second one, respecting the rule of no branch intersection. It follows that the size of $\mathcal{J}(M)$ is upper bounded by

$$\#\mathcal{J}(M) = \frac{N}{M!}\sum_{r=r_{min}}^{r_{max}}\prod_{m=0}^{M-1}K(r)(1-2m\frac{\Delta(r)}{\pi}).$$

Number of false alarms (NFA) For a junction $\jmath$ of order M and at scale r in a discrete image I, its number of false alarms is defined as

$$\text{NFA}(\jmath) := \#\mathcal{J}(M)\cdot F_{\jmath}(t(\jmath)). \tag{4}$$

The quantity $\text{NFA}(\jmath)$ is a quality measure of the junction: the smaller it is, the more meaningful the junction $\jmath$. For a given $\epsilon > 0$, we call a junction ϵ-meaningful, if $\text{NFA}(\jmath) \leq \epsilon$, which implies a detection threshold $t(r,\epsilon) := \min\{t;\ F_{\jmath}(t) \leq \frac{\epsilon}{\#\mathcal{J}(M)}\}$.

The value ϵ corresponds to an expected number of false detections in I. For a fixed ϵ, this formula yields a different threshold on $t(\jmath)$ for each value of the scale r. It is important to observe that the expectation of the number of ϵ-meaningful junctions of order M in a discrete random image $\mathbf{I}$ is smaller than ϵ.

3.4. Maximality for junction characterization

In practice, a single junction can be detected at multiple scales and multiple locations with slight differences. Moreover, at the same location, the local geometrical configurations can be interpreted by several junctions with different orders. We remove these redundancies by relying on an *exclusion* principle, called maximality.

We assign to each junction $\jmath$ a neighborhood $\mathcal{N}'_{\jmath}$ around its center $\mathbf{p}$. For a given order M, we only keep the junctions not containing any more meaningful junction in $\mathcal{N}'_{\jmath}$. That is, we only keep junctions with minimal NFA. Observe that the use of the NFA is a keypoint for scale and location selection. In contrast, using the strengths $t(\jmath)$ for the selection would require a well chosen normalization depending on the scale. Observe also that contrarily to classical approaches relying on the linear scale-space, no blur is introduced, which permits a precise localization of junctions. Last, after scale and location selection have been selected by maximality, the classification of junction is achieved by only keeping the locally more complex junction (the one with the largest order).

4. Experimental results

The proposed ACJ junction detector is evaluated on natural images and the performance are compared with the classical Harris detector [1] and the recent "*Pj on gPb*" method [6]. For Harris, the strength is defined as $\det(H) - 0.04\text{Tr}(H)^2$, H being the Harris matrix. For "*Pj on gPb*", the code provided by M. Maire is used. In order to speed up the detection, we use a fast segment detector, the Line Segment Detector (LSD) [13], to preselect junction candidates. The detection threshold λ of the LSD is set to 10^4 for all experiments. The number of possible directions $K(r)$ for junction branches is chosen as $\lfloor 2\pi r \rfloor$, in order to have a precision of roughly one pixel along the circle of radius r. The precision $\Delta(r)$ is set to $\frac{5}{r}$ relying on similar considerations. In all the experiments, the maximum order of junctions is $M = 4$ and the smallest possible radius is $r_{min} = 3$. The maximal scale r_{max} is chosen as 5% of the diagonal of the image.

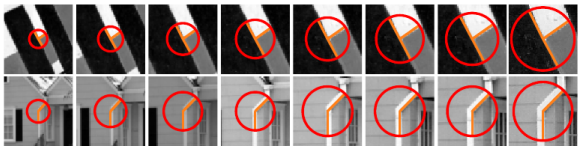

Figure 2: Examples of detected junctions along scale space. Observe the accuracy on simultaneous selection of the scale and location of junctions.

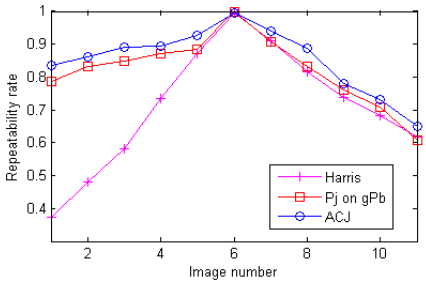

Figure 3: Repeatability rate of different approaches regarding contrast changes.

First, in order to show the accuracy of scale characterization, we run the detection on a sequence of images with different resolutions. An original image I is resized with 8 different zoom factors $\{1, 0.9, 0.8, 0.7, 0.6, 0.5, 0.4, 0.3\}$, using a bilinear interpolation. In Figure 2 are some examples of junctions detected by the proposed algorithm on the sequence. Their detected scales increase linearly with the resolution. A numerical validation of this fact may be found in [11]. To our knowledge, the only related methods permitting such an accurate scale computa-

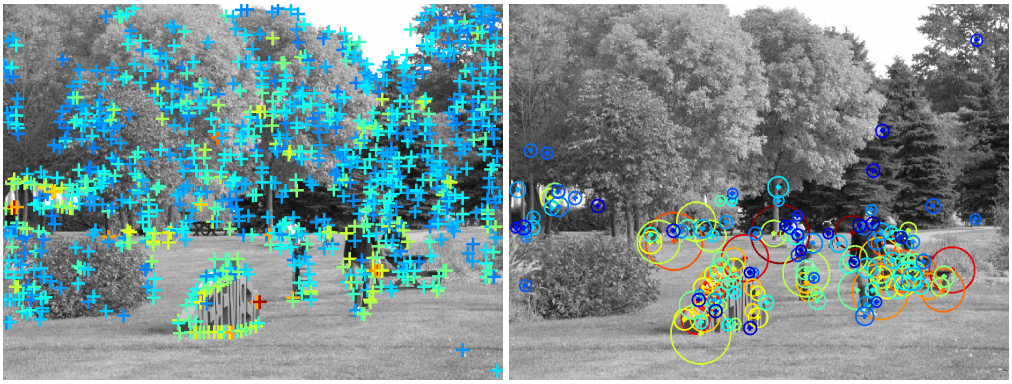

Figure 4: Junction detections on a natural image containing textures. Left: result of the "*Pj* on *gPb*" detector, obtained by thresholding the detector with a value indicated in [6]; Right: result of the proposed ACJ detector. The color of the junction indicates the significance (red for big significance and blue for small one). The results as well as more examples can be found on a dedicated website [14].

tion rely on linear scale-spaces [5]. However, the linear scale-space leads to poor location precisions at large scales. Next, to test the robustness to contrast changes of the ACJ detector, we create sequences of images by applying different gamma corrections (γ takes value in $\{0.25, 0.33, 0.4, 0.5, 0.67, 1, 1.5, 2, 2.5, 3, 4\}$) to an original image. We detect junctions by running the ACJ detector with $\epsilon = 1$, Harris corner detector with a threshold 0.06 on local maxima and "*Pj* on *gPb*" [6]. The repeatability rate of each method on the image sequence is computed. The average curves of repeatability rates on 9 image sequences for the different methods are shown in Figure 3. We can see that ACJ has the best performance. Harris detector known to be highly contrast dependent has the worst performance.

Figure 4 displays the detection results on a natural image with textures. The color of the junction indicates the significance (red corresponds to big significance and blue corresponds to small one). In contrast with "*Pj* on *gPb*" on the left, ACJ detector has much less detections in texture parts, thanks to the a-contrario model. Images used in this section as well as more examples can be found on a dedicated website [14].

5 Conclusion

As a conclusion, we proposed a junction detector with automatic setting of the detection parameter. It is able to control the rate of detection, which is especially visible in textured areas usually proned to yield over-detections. Junctions are accurately characterized through their type, localization and scale. In particular, the method does not rely on the linear scale-space for scale computation, permitting geometric precision. Such an accuracy for junction characterization is potentially useful for depth recovery or motion interpretation.

Acknowledgement This work was partially founded by FUI project CEDCA.

References

[1] C. Harris and M. Stephens, "A combined corner and edge detection," in *Proc. of The Fourth Alvey Vision Conference*, 1988, pp. 147–151.

[2] S. M. Smith and J. M. Brady, "SUSAN: A new approach to low level image processing," *IJCV*, vol. 23, pp. 45–78, 1995.

[3] C. S. Kenney, M. Zuliani, and B. S. Manjunath, "An axiomatic approach to corner detection," in *CVPR*, 2005, pp. 191–197.

[4] E. Rosten, R. Porter, and T. Drummond, "Faster and better: A machine learning approach to corner detection," *IEEE Trans. PAMI*, vol. 32, no. 1, pp. 105 –119, 2010.

[5] T. Lindeberg, "Junction detection with automatic selection of detection scales and localization scales," in *Proc. of ICIP*, 1994, pp. 924–928.

[6] M. Maire, P. Arbelaez, C. Fowlkes, and J. Malik, "Using contours to detect and localize junctions in natural images," in *Proc. of CVPR*, 2008, pp. 1–8.

[7] L. Parida, D. Geiger, and R. A. Hummel, "Junctions: Detection, classification, and reconstruction," *IEEE Trans. PAMI*, vol. 20, no. 7, pp. 687–698, 1998.

[8] M. A. Cazorla and F. Escolano, "Two Bayesian methods for junction classification," *IEEE Trans. Image Processing*, vol. 12, no. 3, pp. 317–327, 2003.

[9] E. D. Sinzinger, "A model-based approach to junction detection using radial energy," *Pattern Recognition*, vol. 41, no. 2, pp. 494 – 505, 2008.

[10] A. Desolneux, L. Moisan, and J.-M. Morel, "Meaningful alignments," *IJCV*, vol. 40, no. 1, pp. 7–23, 2000.

[11] G.-S. Xia, J. Delon, and Y. Gousseau, "Accurate junction detection and characterization in natural images," *TechReport-HAL-00631609*, pp. 1–33, 2011.

[12] D. L. Ruderman and W. Bialek, "Statistics of natural images: Scaling in the woods," *Physical Review Letters*, vol. 73, no. 6, pp. 814–817, 1994.

[13] R. GromponevonGioi, J. Jakubowicz, and J.-M. Morel, "LSD: A fast line segment detector with a false detection control," *IEEE Trans. PAMI*, vol. 32, pp. 722–732, 2010.

[14] G.-S. Xia, http://www.enst.fr/~gousseau/ACJ/acj.html, 2012.