ROBUST 2D LOCATION OF INTEREST POINTS BY ACCUMULATION

*Jean-Philippe Tarel and Rachid Belaroussi**

Université Paris-Est, LEPSiS, IFSTTAR,
58 Boulevard Lefèbvre, F-75015 Paris, France

ABSTRACT

Various interest point and corner definitions were proposed in the past with associated detection algorithms. We propose an intuitive and novel detection algorithm for finding the location of such features in an image. The detection is based on the edges in the original image. Interest points are detected as accumulation points where several edge tangent lines in a neighborhood are crossing. Edge connectivity is not used and thus detected interest points are robust to partial edges, outliers and edge extraction failures at junctions. One advantage of the approach is that detected interest points are not shifted in location when the original image is smoothed compared with other approaches. Experiments performed on Oxford and Cambridge reference databases allow us to show that the proposed detection algorithm performs better than 9 existing interest point detectors in terms of repeatability from multiple camera views.

***Index Terms*—** Interest point, corner, feature point, detection, extraction, accumulation transform, repeatability, edges.

1. INTRODUCTION

Interest points and interest point neighborhoods are highly used features to summarize an image into a very reduced set of information. For instance, interest points are of great importance for object tracking, object recognition, camera calibration and image indexing.

Given an image of a complex scene, it is hard for people to tell where the corners are. The selected positions differ from one person to another. This explains why several corner definitions, applying to images, were proposed in the past. Due to the difficulty to agree on reference corner positions, corner detection algorithms were difficult to rate [1]. Evaluation based on the repeatability of the detected points in different views [2, 3] led to the introduction of interest points and brought reliable results.

Two classes of interest point detectors can be identified: pixel- and edge-based approaches. A recent example of pixel-based algorithm is [4], and an edge-based algorithm is described in [5]. The here proposed algorithm is based on accumulation and thus belongs to the second class. However, contrary to most edge-based algorithms, it does not rely on edge connectivity, so it is more robust to missing edgels (edge elements), outliers and edge extraction failures at junctions. Another advantage lies in its ability to maintain accuracy location even when the original image is smoothed. The proposed algorithm can be used in many tasks, for instance for road sign detection [6].

Fig. 1. Extracted edges, accumulation of tangent line crossings, and detected interest points in red superimposed to the original image. Used parameters are $s = 1$, $g_m = 32$, $d_m = 16$ and $\alpha_m = 0.2$.

After the detailed description of the interest point detection algorithm based on accumulation, a comparative evaluation with state of the art interest point detectors is given, showing the benefit of the proposed algorithm.

2. DETECTION BY ACCUMULATION

2.1. Definition of interest points location

Our idea is to define interest points as points where several tangent lines to the edges are crossing. Our definition is based on the image edges and thus requires edge extraction as a first step. But contrary to many existing methods based on edges, such as [5], we do not want to use the edge connectivity property. Indeed, edge connectivity is often compromised by various perturbations such as noise, occlusion, outliers and smoothing. The second step is an accumulation step over selected pairs of edgels. Third step allows finding the interest point locations and eventually sorting the obtained interest points in decreasing order of strength.

With the previous definition, an interest point is detected at its correct position, even when the corner is smoothed or

*Thanks to the ANR (French National Research Agency) for funding, within the iTowns-MDCO project.

truncated. This is illustrated at the top and bottom right corners of the images in Fig. 1. It is also important to notice that when intensity smoothing is applied on the image, a detected interest point is not affected by a shift in position like in many other approaches. These interesting properties of robustness and accuracy are a consequence of the accumulation of evidence on the crossing of edge tangent lines. In addition to the missing edgels robustness, the process is also robust to outliers as illustrated in the top left corner of the images in Fig. 1. Notice that points not usually considered as corners are also detected as interest points. For instance, in the bottom left interest point of the images in Fig. 1, two interest points are detected as the crossing of horizontal and vertical edges even though the edges do not actually intersect.

2.2. Edge extraction

Whatever the chosen edge extraction algorithm, it must provide a set of edgels. Here, an edgel is a 2D point position P_i in pixels with the gradient vector G_i perpendicular to the edge at P_i. Usually, the edge extraction algorithm is parameterized by the kernel size s of the filter used to reduce the noise on image intensities. Edge detection is applied on the original image after normalization between $[0, 255]$. In practice, we use Deriche edge detector [7] which gives the horizontal and vertical components of the gradient at each pixel position in the original image $I(P)$. Then, a no-maximum suppression is applied to obtain edges. Edgels with a gradient norm lower than a threshold g_m are discarded. This threshold allows removing edges in noisy areas and reducing the number of edgel pairs used in the next step. The output of the first step is thus a set of n edgels (P_i, G_i).

2.3. Accumulation over edgel pairs

To find the points where several tangent lines are crossing, pairs of edgels with indexes i and j are selected. To reduce the set of possibilities and because interest points are defined locally, only pairs of edgels at a distance d_{ij} lower than d_m are selected. Moreover, to avoid detecting interest points along a low curvature edge, a selection is performed with respect to the angle between gradient vector G_i and the opposite of gradient vector G_j. We denote this angle by α_{ij}, see Fig. 2, and it must be within $[-\frac{\pi}{2}, \frac{\pi}{2}]$ for selected edgel pairs. As a consequence, interest points with obtuse angle will not be detected. To take into account the estimation error on gradient angles, the interval $[-\frac{\pi}{2}, \frac{\pi}{2}]$ becomes $[-\frac{\pi}{2} - \alpha_m, \frac{\pi}{2} + \alpha_m]$ where α_m is the assumed maximum error on gradient angles.

For a pair of selected edgels (P_i, G_i) and (P_j, G_j), we compute, when it exists, the crossing point C_{ij} of the two tangent lines of the edges going through P_i and P_j respectively, see Fig. 2. From the equation of the tangent lines, we deduce two equations on the position C_{ij}: $G_i^t(C_{ij} - P_i) = 0$ and $G_j^t(C_{ij} - P_j) = 0$, where ^{t} denotes the transpose operator.

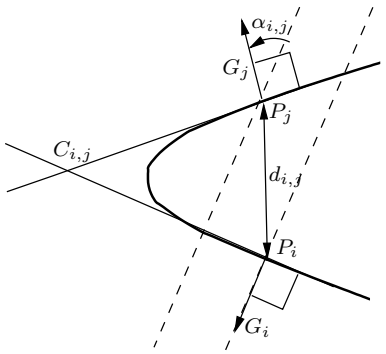

Fig. 2. Accumulation at the crossing $C_{i,j}$ of the two tangent lines going through P_i and P_j with normals G_i and G_j respectively.

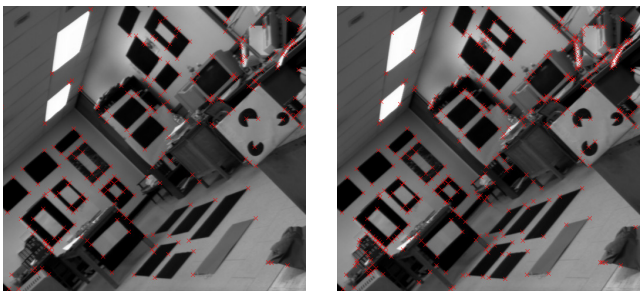

Fig. 3. Detected interest points in red superimposed to the original image when $\alpha_m = 0.2$, left and $\alpha_m = 1.3$, right. Interest points with obtuse angles are not detected in the left image and are detected in the right image. The other parameters are $s = 1$, $g_m = 32$, $d_m = 16$.

The crossing point C_{ij} is thus obtained as:

$$C_{ij} = (G_i \;\; G_j)^{-t} \begin{pmatrix} G_i^t P_i \\ G_j^t P_j \end{pmatrix} \tag{1}$$

The positions of $C_{i,j}$ over the selected pairs of edgels are accumulated in a 2D histogram $h(P)$ the size of which is the same as that of the original image $I(P)$. Each vote at position $C_{i,j}$ can be simply an increment by one. However, to favor edges with strong gradients over edges with low gradients, we use a vote with the value $\sqrt{\|G_i\|\|G_j\|}$. This value is the logarithm average of the gradient norms $\|G_i\|$ and $\|G_j\|$. In summary, the 2D histogram $h(P)$ can be formally described as:

$$h(P) = \sum_{i=1}^{n}\sum_{j=1}^{n} \sqrt{\|G_i\|\|G_j\|}\, \mathbb{1}_{\|P_i - P_j\| < d_m} \mathbb{1}_{|\alpha_{i,j}| < \frac{\pi}{2} + \alpha_m} \mathbb{1}_{P = C_{ij}} \tag{2}$$

where $\mathbb{1}$ is the indicator function (if x is true $\mathbb{1}_x = 1$ else 0), $\alpha_{i,j} = \widehat{G_i, G_j} - \pi$, and C_{ij} is given by (1).

2.4. Local maxima

After the 2D histogram $h(P)$ is obtained, the positions of local maxima give the interest point locations. A local maximum in the 2D histogram h is a pixel position P where $h(P)$ is higher than or equal to the values at its eight neighbors. The detected interest points can be sorted using the values of h at the maximum. The N first sorted points can be selected as the detected interest points. An alternative to sorting is simply to remove the locations where the value of h is smaller than a threshold.

3. EXPERIMENTS

3.1. Parameters selection

The gradient norm threshold g_m used during edge extraction is usually between 10 and 40 gray levels, i.e between 4% and 16% of the maximum intensity of the input image. The higher the value of g_m, the faster the interest point detection algorithm. The scale parameter s used to filter the image during edge extraction is the main parameter of the detection algorithm by accumulation since it specifies the scale at which interest points are detected. In noisy or randomly textured images, a large value of s should be preferred. On the contrary, in clean detailed images, a small value of s should suffice. Usually s is between 0.3 and 3 pixels. The maximum distance d_m used for edgel pairs selection can be used to change the ordering of the detected interest points. But the higher the value of d_m, the higher the number of edgel pairs and thus the slower the interest point detection algorithm. Moreover, d_m must not be too small to allow stable interest points detection first. In practice, d_m is set between 2% and 8% of the input image size. The last parameter α_m does not play an important role and is usually set to 0.2 rad. It is also introduced in the algorithm to allow, when it is necessary, extracting interest points with obtuse angles as well as interest points with acute angles. In the left and right images of Fig. 3, the extracted interest points are shown in red, left with $\alpha_m = 0.2$ rad and right with $\alpha_m = 1.3$ rad. With $\alpha_m = 0.2$, only acute interest points are detected; when with $\alpha_m = 1.3$ both acute and obtuse interest points are detected. The price to pay for the detection of obtuse interest points is an increasing number of low stability interest points along edges with small curvatures. The original image comes from[1], see [1].

3.2. Comparison with other interest point detectors

For the evaluation of the interest point detection algorithm by accumulation in comparison with state of the art interest point detectors, we follow exactly the evaluation scheme described in details in [4], similar in its principle to the one described in [2]. The evaluation is based on the ϵ-repeatability which is defined as the ratio of the number of interest points repeatedly detected in a pair of images (of the same scene) over the number of interest points appearing in both images. ϵ is the maximum error accepted on the interest point location (typically $\epsilon = 5$ pixels). As in [4], two image databases with ground-truth mapping between views are used: the Cambridge database[2] and the Oxford database[3]. The Oxford database consists in 8 sequences of 5 pairs with a homography mapping between two views, whereas the Cambridge database consists in 3 sequences of 210, 182 and 56 pairs with a 3D mapping between two views. For algorithm comparison, the repeatability is plotted versus the number of detected interest points for different algorithms. The algorithms parameters are selected by optimizing the results on the "bark" dataset, then the same values are applied to other datasets. Only the scale s of the smoothing is modified for the Cambridge database due to better registration between views compared with the Oxford database.

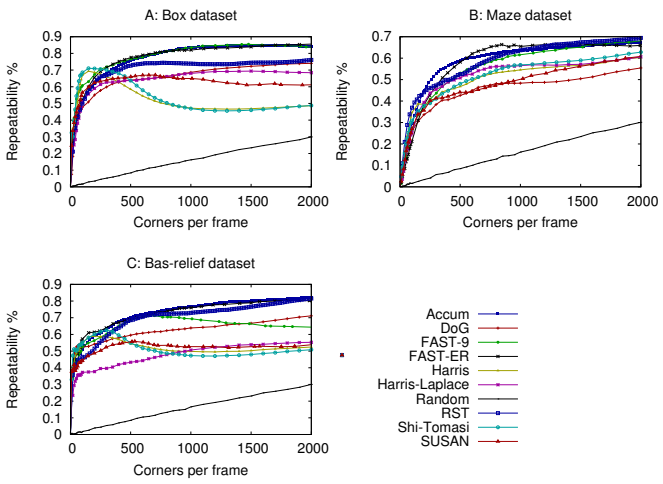

Fig. 4. Comparison of the repeatability curves of different interest point detectors on the 3 sequences of the Cambridge database[2]. Used parameters are $s = \frac{1}{3}$, $g_m = 32$, $d_m = 16$ and $\alpha_m = 0.2$.

The repeatability curves obtained in [4] are shown for Harris, Harris-Laplace, Shi-Tomasi, SUSAN, DoG (used in Sift), FAST-9 and FAST-ER algorithms, along with the curve obtained when interest points are randomly sampled in the image. We added the two blue curves which are for the Radial Symmetry Transform [8] (RST) and the proposed accumulation based interest point detection algorithm (Accum). The results obtained in Fig. 4 on the Cambridge database equal those of the best FAST-9 and FAST-ER algorithms which are based on learning on this database, with the advantage that the algorithm based on accumulation is particularly simple to implement and does not need a learning step.

[1] http://www.ee.surrey.ac.uk/CVSSP/demos/corners/originals.html

[2] http://svr-www.eng.cam.ac.uk/~er258/work/datasets.html

[3] http://www.robots.ox.ac.uk/~vgg/data/data-aff.html

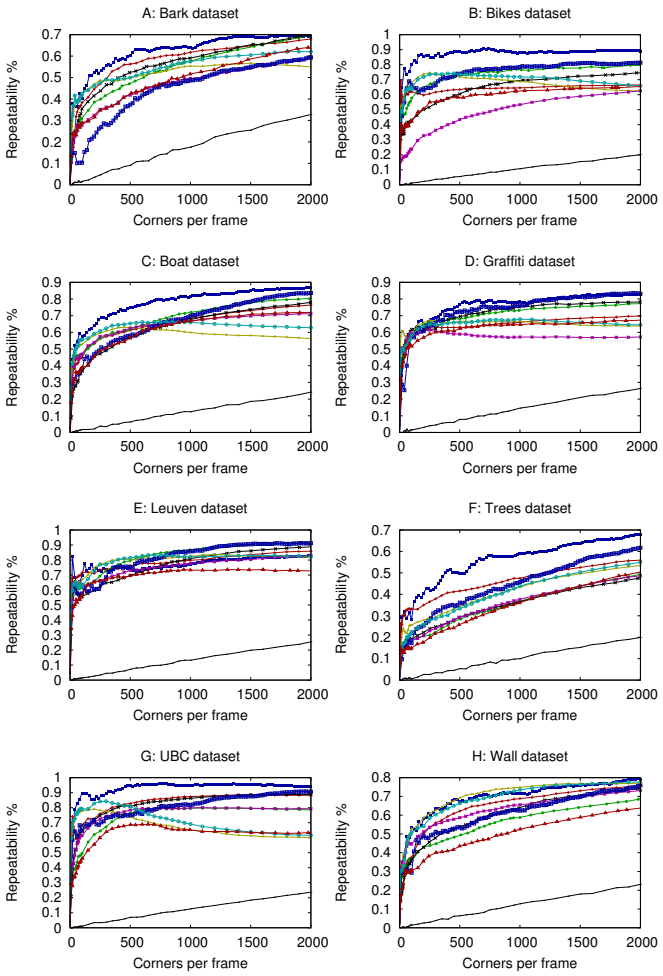

Fig. 5. Comparison of the repeatability curves of different interest point detectors on the 8 sequences of the Oxford database[3]. The curve keys are shown in Fig. 4. Used parameters are $s = 2.5$, $g_m = 32$, $d_m = 16$ and $\alpha_m = 0.2$.

Compared with the other algorithms, ours scores are higher in 6 sequences over the 8 of the Oxford database, see Fig. 5. In the Leuven sequence, where gamma transformations are applied on the image intensities, our results are within the average range of the other results. This indicates a lack of robustness to gamma transformations due to the g_m gradient threshold. This can be fixed by adding a preprocessing step enforcing normalization of the average of the logarithm of the image intensity. In the Wall sequence, where images are taken with various viewing angles, our results are similar to the best results despite a large zoom with a constant smoothing scale s.

Average processing time for interest point detection with sorting is 624 ms on the Oxford images (max size 1000×700) and 417 ms on the Cambridge images (768×576) using a 1.1GHz Intel Core 2 Duo PC. Thus, it is similar in speed to the Harris detector.

4. CONCLUSION

We described an interest point detection algorithm based on the accumulation of evidence on the position of the crossing of pairs of edge tangent lines. This algorithm uses edges but is robust to missing edgels and outliers. In the comparative experiments, our interest point detection algorithm shows enhanced results over 9 state of the art algorithms. We point out that the scale of the filter during the edge extraction is the important parameter related to the scale at which interest points are extracted. We thus plan to extend our algorithm by introducing scale self estimation.

Acknowledgments

Deep thanks to Edward Rosten for providing us, from [4], the curves for 8 of the algorithms which appear in Fig. 4 and 5.

5. REFERENCES

[1] F. Mokhtarian and R. Suomela, "Robust image corner detection through curvature scale space," *IEEE Trans. on PAMI*, vol. 20, pp. 1376–1381, 1998.

[2] C. Schmid, R. Mohr, and C. Bauckhage, "Evaluation of interest point detectors," *International Journal of Computer Vision*, vol. 37, pp. 151–172, 2000.

[3] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," *IEEE Trans. on PAMI*, vol. 27, pp. 1615–1630, 2005.

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

[5] M. Awrangjeb, G. Lu, C. S. Fraser, and M. Ravanbakhsh, "A fast corner detector based on the chord-to-point distance accumulation technique," in *Proc. of Digital Image Computing: Techniques and Applications (DICTA'09)*, Washington, DC, USA, 2009, pp. 519–525.

[6] R. Belaroussi and J.-P. Tarel, "Angle vertex and bisector geometric model for triangular road sign detection," in *Proc. of IEEE Workshop on Applications of Computer Vision (WACV'09)*, Snowbird, Utah, USA, 2009, pp. 577–583.

[7] R. Deriche, "Using Canny's criteria to derive a recursively implemented optimal edge detector," *International Journal of Computer Vision*, vol. 1, no. 2, pp. 167–187, 1987.

[8] G. Loy and A. Zelinsky, "Fast radial symmetry for detecting points of interest," *IEEE Trans. on PAMI*, vol. 25, no. 8, pp. 959 – 973, 2003.