Czech Pattern Recognition Workshop 2000, Tomáš Svoboda (Ed.)
Peršlák, Czech Republic, February 2–4, 2000
Czech Pattern Recognition Society

Feature point detection in multiframe images

Barbara Zitová[1], Jan Flusser[1], Jaroslav Kautsky[2] and Gabriele Peters[3]

[1] UTIA, Academy of Sciences of the Czech Republic
Pod vodárenskou věží 4, Prague 8, 180 00, Czech Republic
[2] Flinders University of South Australia, Adelaide, Australia
[3] Institut for Neuroinformatics, Ruhr-University, Bochum, Germany
E-mail: {zitova, flusser}@utia.cas.cz

Abstract *Feature point (FP) detection is an important preprocessing step in image registration, data fusion, object recognition and in many other tasks. This paper deals with multiframe FP detection, i.e. detection in two or more images of the same scene which are supposed to be blurred, noisy, rotated and shifted with respect to each other. We present a new method invariant under rotation that can handle differently blurred images. Thanks to this, the point sets extracted from different frames have relatively high number of common elements. This property is highly desirable for further multiframe processing. The performance of the method is demonstrated experimentally on satellite images and application on medical data is*

1 Introduction

Detection of feature points (FP) or landmarks is an important step in image processing and computer vision. It provides input information for further operations, such as image registration, image fusion, time-sequence analysis and object recognition. By feature points we understand the points that are easy to identify in the image, such as corners, line intersections, T-junctions, etc.

In this paper, we address a multiframe version of this problem: having two or more images of the same scene, the aim is to detect feature points in each of them. Multiframe FP detection methods must fulfill the condition of repeatability. This property means that the results should not be affected by imaging geometry, radiometric conditions and by additive noise and that the sets of points detected in all frames should be identical. Since the last requirement is not realistic in practice, "maximum overlap" is usually required instead of identity.

In this paper we assume that the individual frames may be rotated and shifted with respect one another, they may have different contrast, they may be degraded by a linear shift-invariant blur and corrupted by additive random noise. Our primary motivation comes from the area of remote sensing, where the registration of images with such kinds of distortions is a very frequent task. Having the FP detection method which works on differently distorted frames and which yields high repetition rate is a fundamental requirement.

2 Present state-of-the-art

Numerous methods for single-frame feature point detection in gray-level images have been published in last two decades. Most of them are known as corner detectors. A survey of basic methods along with a comparison of their localization properties can be found in [9].

Kitchen and Rosenfeld [5] proposed a corner detection scheme based on a differential operator that consists of first and second order partial derivatives of the image $f(x, y)$:

$$K(x,y) = \frac{f_x^2 f_{yy} - 2 f_x f_y f_{xy} + f_y^2 f_{xx}}{f_x^2 + f_y^2}. \qquad (1)$$

$K(x, y)$ represents the curvature of a plane curve perpendicular to the gradient of the image function. Corners are identified as local extrema of this operator.

Brunnström et al. [2] proposed a modified version of the Kitchen and Rosenfeld's corner detector. Their method looks for local extrema of the numerator of $K(x, y)$. In that way, preference is given to the points with high value of the gradient.

Another modification of the Kitchen and Rosenfeld's approach comes from Zuniga and Haralic [13] who detect edges first and then they look for extrema of $K(x, y)$ normalized by the gradient magnitude over edge pixels only.

Beaudet [1] proposed to calculate Hessian determinant

$$|H(x,y)| = f_{xx} f_{yy} - f_{xy}^2 \qquad (2)$$

of the image function and to find corners as local extrema of this determinant.

In Dreschler's and Nagel's approach [3] the local extrema of Gaussian curvature of the image function are identified and corners are localized by interpolation between them.

Unlike the above mentioned methods, the corner detector proposed by Förstner [4] uses first-order derivatives only. Förstner determines corners as local maxima of

$$F(x,y) = \frac{\overline{f_x^2}\,\overline{f_y^2} - (\overline{f_x f_y})^2}{\overline{f_x^2} + \overline{f_y^2}} \qquad (3)$$

where bars denote mean values over some neighborhood of (x, y). Harris' method belongs to the same family of corner detectors as Förstner one. Here, corners are determined as local minima of $1/F(x, y)$. In several comparative studies (see [10] for instance), Harris detector was evaluated as the best corner detector, although it is relatively time-consuming. To reduce its computational cost, Trajkovic and Hedley [10] proposed to calculate the cost function $F(x, y)$ for pixels with high gradient only.

Simple and fast corner detector has been introduced recently by Trajkovic and Hedley [10]. It is based on the idea that the change of image intensity at the corners should be high in all directions. Thus, corners are found as local maxima of minimum change of intensity. Although it is very fast, this detector performs slightly worse than Harris detector because it sometimes gives false responses on the straight lines.

Many of the corner detection methods were developed on the basis of edge detectors but most edge detectors perform poorly on corners, because they assume an edge to be of infinite extend. For this reason Mehrotra [8] developed his half-edge detectors based on the first and second directional derivatives of Gaussian, respectively. Among others, more recently developed methods are an approach of Liu and Tsai [6] which is based on preserving gray and mass moments, a method developed by Xie [12] who combines different cues (edginess, curvature and region dissimilarity) in a cost function to be minimized, and a biologically inspired approach of Würtz and Lourens [11] who utilize a model for end-stopped cells of the visual cortex over several scales and even generalize it to color images.

Most of the above mentioned methods can be used in the multiframe case too, but their repeatability is not sufficient in the case of blurred frames.

3 Description of the proposed method

Our newly proposed method for the detection of feature points uses a parameter approach to handle differently distorted images. Points, which belong to two edges with an angle from the interval $[\pi/2 - d_a, \pi/2 + d_a]$ (d_a is user defined parameter) in between regardless of its orientation are understood here as feature points. The described method is based on this definition.

Information about the number of edges passing through each pixel and about the angle between them is acquired from the number and distribution of local sign changes in the difference between the image function and its local mean values (see (6)).

However, the list of candidates thus produced (Step 5 of the algorithm) usually contains also some undesirable points: points that are not corners but which are close to a straight line and also points which are true corners but with a small variation in gray levels. At first, points closer to a straight line than given threshold are eliminated and then the final choice of the best FP from the list of candidates is done by maximizing the weight function W (5), which quantifies the "significance" of each point. In this way we eliminate false candidates. Furthermore, the requirement not to yield two FP closer to each other than a user-defined distance is incorporated. Finally, the algorithm will produce a user requested number of extracted FP which satisfy the criteria above and maximize the weight function.

More formally, the proposed method is described in the following algorithm.

Algorithm *Find_FP*

1. Inputs:

   f – the image of the size $N \times N$ in which FP should be detected.
   N_{FP} – the desired number of feature points.
   M – the radius of the neighborhood for the mean value computation.
   r – the radius of the neighborhood for computing sign changes.
   d_a – determines the interval, where the angle between FP candidate's edges has to be from.
   s – the minimum allowed distance between FP candidate and a straight line.
   d_s – the maximum allowed curvature divergence for straight line candidates.
   t – the minimum allowed distance between two feature points.

2. Initialize C – zero matrix of the size $N \times N$.

3. Calculate function g of local mean values of f

$$g(i,j) = \frac{1}{\pi M^2} \sum_{\Omega_{i,j,M}} f(k,l), \tag{4}$$

   where $\Omega_{i,j,M}$ denotes a circular neighborhood of (i, j) of the radius M.

4. Calculate the weight function of local variations:

$$W(i,j) = \sum_{\Omega_{i,j,M}} (f(k,l) - g(i,j))^2. \tag{5}$$

5. Detection of FP candidates:

   FOR $i = r + 1$ TO $N - r$
   FOR $j = r + 1$ TO $N - r$
   Construct one pixel thick closed digital circle R of radius r centered at (i, j):

   $$R = \{(i_1, j_1), \cdots, (i_k, j_k)\}$$

   where $i_1 = i$ and $j_1 = j + r$ and next points follow in the clockwise order.
   Calculate the number of sign changes $N_{sc}(i, j)$ in the sequence

$$f(i_1,j_1)-g(i,j), \cdots, f(i_k,j_k)-g(i,j), f(i_1,j_1)-g(i,j) \tag{6}$$

IF $N_{sc}(i,j) = 2$ THEN
Denote the positions of the sign changes as (i_a, j_a) and (i_b, j_b), respectively.
Calculate

$$\alpha_{i,j} = \text{angle}((i_a, j_a), (i,j), (i_b, j_b)).$$

IF $|\alpha_{i,j} - \pi/2| < d_a$ THEN
$C(i,j) = 1$
END_IF
END_IF
END_FOR
END_FOR

6. Elimination of false candidates:
FOR each pixel (i,j) where $C(i,j) = 1$
IF exists pixel (i_f, j_f) such that
the distance of which from (i,j) is less than s,
$N_{sc}(i_f, j_f) = 2$ and $|\alpha_{i_f,j_f} - \pi| < d_s$
THEN
$C(i,j) = 0$
END_IF
END_FOR

7. Selecting feature points:
FOR $m = 1$ TO N_{FP}
Find point (i_0, j_0) as

$$(i_0, j_0) = \arg \max_{i,j:C(i,j)=1} W(i,j).$$

Set $P_m = (i_0, j_0)$.
For each point (i,j) the distance of which from (i_0, j_0) is less than t set $W(i,j) = 0$.
END_FOR

The resulting sequence $P_1, \cdots, P_{N_{FP}}$ contains the coordinates of the detected feature points.

The role of the weight function can be explained as follows. If the candidate is a corner with low contrast between the adjacent regions, its value of W is small. In the case of ideal corner W is high.

If the noise is present in the image, the sequence

$$f(i_1, j_1), \cdots, f(i_k, j_k), f(i_1, j_1) \tag{7}$$

can be smoothed before the sign changes computing in (6).

It can be seen that the described algorithm has a property of rotation invariance. When f' is a rotated version of image f, the functions g' and W' are equal to g and W, respectively, rotated in the same manner. The circle R' contains the same pixels as R but labeled differently. Nevertheless, the number of sign changes N'_{sc} is the same as N_{sc}. Since also $\alpha'_{ij} = \alpha_{ij}$, the set of feature points detected in f' is the same (except for the rotation) as that one found in image f.

During the FP detection several user-defined parameters are used. They allow handling differently blurred and corrupted images, as it is demonstrated in the next Section. This variability is an important feature of the proposed method.

4 Numerical experiments

In this Section, practical capabilities of the proposed FP detection method are demonstrated and a comparison with the classical techniques [5] and [4] is shown. Since the intended major application area is the area of remote sensing, the experiments are performed on satellite images.

A subscene covering the landscape near Prague (Czech capital city) of the size 180×180 pixels was extracted from the SPOT image of the central part of the Czech Republic. This subscene was rotated several times by angles from $\pi/36$ to $\pi/4$ and/or blurred by convolving with square masks of various sizes to simulate degraded multiframe acquisition.

30 feature points were detected in each frame by three different methods: Kitchen and Rosenfeld's, Harris' and ours. In each case we calculated the *success rate Q* that is defined as the number of identical FP detected both in the original and in the degraded/rotated frame. Two FP were assumed to be identical if their positions in both images differ from each other at most by two pixels in each direction.

The results of the experiment are summarized in Table 1. In the first two columns the size of the blurring filter and the angle of rotation is specified for each frame. In the third and fourth columns one can see the success rate achieved by Kitchen and Rosenfeld's and Harris' methods, respectively. The parameter in the fifth column stands for the radius of the neighborhood over which the mean values of the derivatives used in (3) were calculated from. The last four columns present the results achieved by our method: the success rate Q and the values of parameters M (the radius of the neighborhood for mean value computation), r (the radius of the circle for sign changes analysis) and I (the interval, where the angle between FP candidate's edges has to be from), respectively. In each individual case, the parameter values listed in Table 1 in Harris' as well as in our method were selected to yield the best success rate.

In Fig. 1, one can see what feature points were detected by each method. On the left-hand side is the original, on the right-hand side is the image blurred by 9×9 averaging mask and rotated by $\pi/9$. The feature points were detected by Kitchen and Rosenfeld's method (top), Harris' method (middle) and our method (bottom). This figure shows the situation corresponding to the last but one row of Table 1.

Analyzing the results of this experiment, we can make the following claims.

- In the case of heavy blur and small rotation our method outperforms the others.
- If the blur is not feature and the rotation angle is about $\pi/4$, then the Harris' method becomes better than ours.
- In all other cases, Harris' and our methods are quite comparable.
- Kitchen and Rosenfeld's algorithm gives the worse success rate in all cases.
- Computational cost of our method is lower than that of Harris' method.

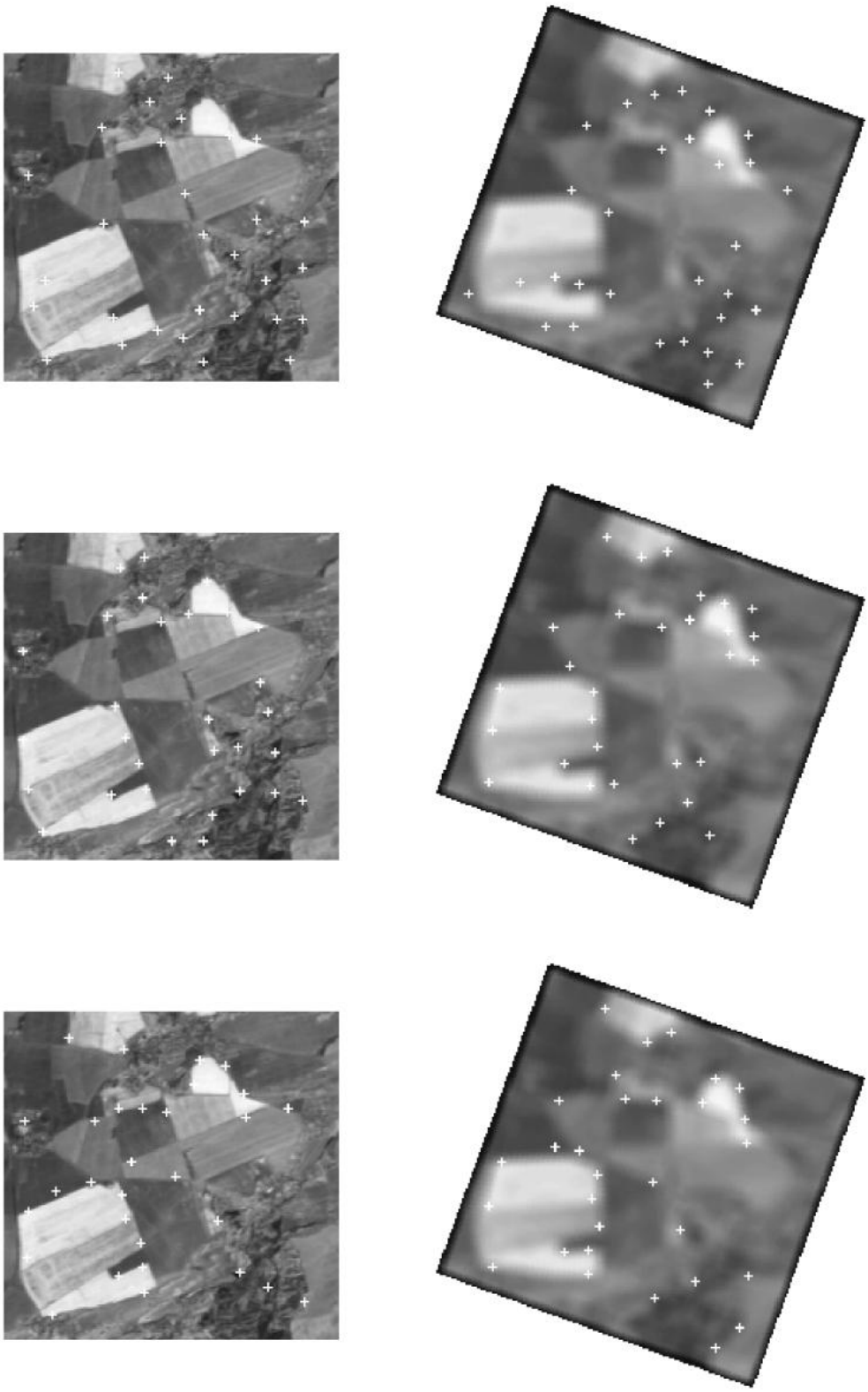

Figure 1: Detection of feature points in two different frames of the same scene: in the original (left) and in the image blurred by 9×9 averaging mask and rotated by $\pi/9$. The feature points were detected by the Kitchen and Rosenfeld's method (top), the Harris' method (middle) and by our method (bottom).

Frame		K + R	Harris		Our method			
Blur	Rotation	Q	Q	h	Q	M	r	I
3×3	–	11	23	6	25	2	4	$34^0 - 146^0$
5×5	–	3	17	9	21	2	4	$34^0 - 146^0$
7×7	–	4	16	9	18	2	4	$34^0 - 146^0$
9×9	–	2	9	9	17	4	8	$6^0 - 174^0$
–	$\pi/8$	17	24	9	25	2	4	$0^0 - 180^0$
3×3	$\pi/8$	13	23	6	24	2	4	$0^0 - 180^0$
5×5	$\pi/8$	8	19	6	18	2	4	$0^0 - 180^0$
7×7	$\pi/8$	3	19	6	17	2	4	$0^0 - 180^0$
7×7	$\pi/4$	5	16	6	14	2	4	$22^0 - 158^0$
9×9	$\pi/36$	3	11	9	19	4	8	$6^0 - 174^0$
9×9	$2\pi/36$	5	14	9	18	4	8	$6^0 - 174^0$
9×9	$4\pi/36$	3	11	9	20	4	8	$6^0 - 174^0$
9×9	$5\pi/36$	3	12	9	17	4	8	$6^0 - 174^0$

Table 1: The results of the FP detection. From left to right: the size of the blurring filter, the rotation angle, the success rate of Kitchen and Rosenfeld's method, the success rate of Harris' method, h – the radius of the neighborhood for calculating the mean values of the derivatives, the success rate of our method, M – the radius of the neighborhood for the mean value computation, r – the radius of the circle for sign changes analysis, I – the interval, where the angle between FP candidate's edges has to be from.

5 Application

Medical imagery application area of image registration and in consequence automatic feature points detection has been extensively studied during last years due to increasing number and availability of different sensor types. For image-based registration of medical images, an extrinsic or an intrinsic approach can be used. Extrinsic methods, relying on artificial objects introduced to the patient space (inside or outside of the body), are usually fast and can be automatized. But not in all cases these often invasive and time-demanding for preparation methods can be applied. Intrinsic solutions proceed patient-generated image content either entire or just limited set of features (points, segmented objects) [7]. Here, our proposed method was used for feature points detection on nuclear MR images of the human brain. In Fig. 2 the 157th axial slice (left image) and its blurred version (right image) are shown together with 30 detected feature points. Blur was introduced by averaging mask 7×7 to simulate on the image pair the difference between images acquired by sensors with different resolution. Detected FP sets reached required 50% overlap, which is sufficient for prospective image registration procedure.

6 Conclusion

In this paper we proposed a novel method for detection of feature points – corners with high local contrast. The method works in two stages: all possible candidates are found first and then the desirable number of resulting feature points is selected among them by maximizing the weight function.

Although the method can be applied to any image, it is particularly devoted to FP detection in blurred images because it provides high consistence. We compared the performance of the method with two classical corner detectors. The number of identical points detected in different frames of the same scene served as a success rate. Our method was shown to be superior if at least one of the frames is heavily blurred and to be comparable with Harris' detector in most other cases except negligible or small blur and big rotation. Moreover, our method is much more computationally efficient.

6.1 Acknowledgment

This work has been partially supported by the grants No. 102/96/1694 and No. 102/00/1711 of the Grant Agency of the Czech Republic.

References

[1] P. R. Beaudet. Rotationally invariant image operators. *Proc. Intl. Joint Conference on Pattern Recognition*, pages 579–583, Kyoto, Japan, 1978.

[2] K. Brunnström, T. Lindeberg, and J. O. Eklundh. Active detection and classification of junctions. *Proc. 2nd European Conf. Computer Vision*, LNCS 588:701–709, St. Margherita, Italy, 1992.

[3] L. Dreschler and H. Nagel. Volumetric model and 3-D trajectory of a moving car derived from monocular TV-frame sequence of a street scene. *Proc. Int. Joint Conf. Artificial Intelligence*, pages 692–697, Vancouver, Canada, 1981.

[4] W. Förstner. A feature based correspondence algorithm for image matching. *Intl. Arch. Photogramm. Remote Sensing*, 26:150–166, 1986.

[5] L. Kitchen and A. Rosenfeld. Gray level corner detection. *Pattern Recognition Letters*, 1:95–102, 1982.

[6] S.-T. Liu and W.-H. Tsai. Moment-preserving corner detection. *Pattern Recognition*, 23:441–460, 1990.

[7] J. A. Maintz and M. A. Viergever. A survey of medical image registration. *Medical Image Analysis*, 2(1):1–36, 1998.

[8] R. Mehrotra. Corner detection. *Pattern Recognition*, 23:1223–1233, 1990.

[9] K. Rohr. Localization properties of direct corner detectors. *Journal of Mathematical Imaging and Vision*, 4:139–150, 1994.

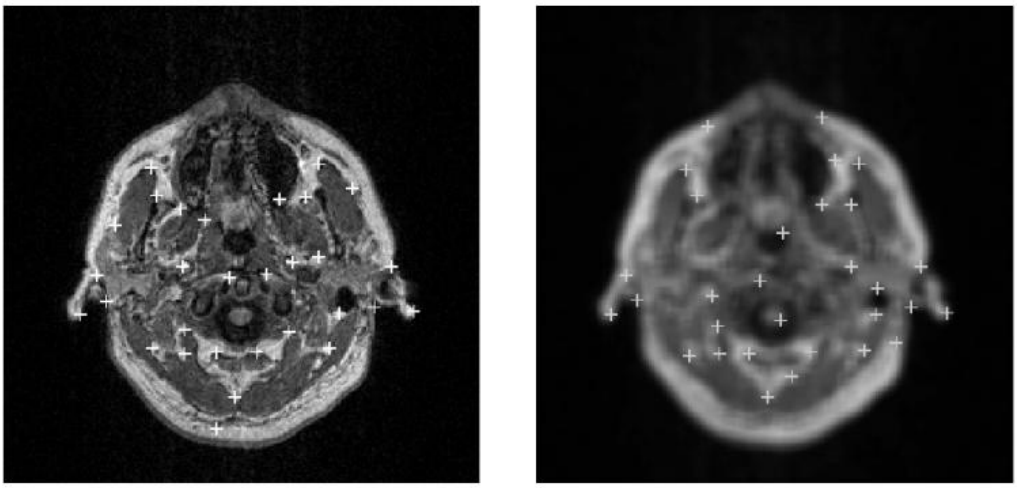

Figure 2: MR image of the brain; the 157th axial slice (left: original image, right: blurred image by 7×7 averaging mask) with 30 detected feature points.

[10] M. Trajkovic and M. Hedley. Fast corner detection. *Image and Vision Computing*, 16:75–87, 1998.

[11] R. P. Würtz and T. Lourens. Corner detection in color images by multiscale combination of end-stopped cortical cells. *Proc. Intl. Conf. Artificial Neural Networks ICANN'97*, LNCS 1327:901–906, Springer, 1997.

[12] X. Xie. Corner detection by a cost minimization approach. *Pattern Recognition*, 26:1235–1243, 1993.

[13] O. A. Zuniga and R. Haralick. Corner detection using the facet model. *Proc. IEEE Conf. Computer Vision Pattern Recognition*, pages 30–37, 1983.