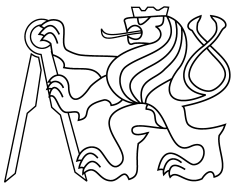
CENTER FOR
MACHINE PERCEPTION



CZECH TECHNICAL
UNIVERSITY

REPRINT

Stable Wave Detector of blobs in images

Jan Dupač
Václav Hlaváč

dupac@dupac@rsdynamics.com, hlavac@fel.cvut.cz

RS Dynamics s.r.o., 149 00 Prague 4, Starochodovská 1359, Czech Republic
http://www.rsdynamics.com

Czech Technical University, Faculty of Electrical Engineering
Department for Cybernetics, Center for Machine Perception
121 35 Prague 2, Karlovo náměstí 13, Czech Republic
http://cmp.felk.cvut.cz

Dupac J., Hlavac V.: Stable Wave Detector of blobs in images, Proceedings of the 28th DAGM (German Pattern Recognition Society) Symposium, Springer, Berlin, Held in Berlin, Germany on September 12-14, 2006, LNCS Volume 4174, pages 760-769, ISBN 3-540-44412-2

Available at
ftp://cmp.felk.cvut.cz/pub/cmp/articles/hlavac/
Dupac-DAGM2006StableWaveReprint.pdf

Center for Machine Perception, Department of Cybernetics
Faculty of Electrical Engineering, Czech Technical University
Technická 2, 166 27 Prague 6, Czech Republic
fax +420 2 2435 7385, phone +420 2 2435 7637, www: http://cmp.felk.cvut.cz

Stable Wave Detector of blobs in images

Jan Dupač[1] and Václav Hlaváč[2]

[1] RS Dynamics s.r.o., 149 00 Prague 4, Starochodovská 1359, Czech Republic
http://www.rsdynamics.com, dupac@rsdynamics.com

[2] Czech Technical University, Faculty of Electrical Engineering
Department for Cybernetics, Center for Machine Perception
121 35 Prague 2, Karlovo náměstí 13, Czech Republic
http://cmp.felk.cvut.cz, hlavac@fel.cvut.cz

Abstract. Stable Wave Detector (SWD) is a new multiscale landmark detector in the intensity image. SWD belongs to a group of interest-point-like operators aiming at detecting repeatedly distinguished entities regardless of their semantics. The speed and the robustness of landmark detection and the precision of landmark localization are main issues. The target landmarks are blobs which correspond to local maxima/minima of intensity (positive and negative peaks). The detector is based on the phase of the first harmonic wave in the moving window. The localization is a result of an integral transformation rather than a derivative. Thus, the blob detector is inherently robust to noise. The SWD provides sub-pixel localization of blobs together with the estimate of its precision, the measure of the strength/significance and the estimate of the size/scale for each blob.

1 Motivation

This work is a part of a larger project leading to a hand-held instrument for on-line 3D positioning of geophysical measurements. The proposed method uses video sequences captured by a calibrated stereo rig. The required precision is in order of 0.1 meter in the area of several hundreds meters. The requirement for online processing and a hand-held battery-powered system induces hard constraints to computational complexity of used algorithms which should simultaneously meet challenging precision requirements. The subpixel measurement is required because enlarging camera resolution leads to quadratically growing amount of data compared to the linear gain in precision.

What are the landmarks to search? We were inspired by the nature. It is known that dogs are short-sighted and cannot see sharply. Nevertheless, they live in the 3D world and can hunt. The first author is short-sighted too, however, even without glasses he can also feel a 3D structure of the world. Another observation concerns fast movements. If something moves very fast then we are unable to recognize details of it. However, we feel its position and speed. These observations are our arguments to seek same integral units as blobs rather than differential structures as corners or edges, i.e., the whole window instead of its corners and its center instead of the border.

2 State-of-the-art

The proposed method can be broadly classified among interest-point-like operators aiming at detecting repeatedly distinguished entities without considering their semantics. These entities are mainly used for matching in different images, stereo, image retrieval, categorization, object recognition, etc. There is a popular group of corner detectors originated in Moravec detector [11] improved by Harris [4]. Fast implementation of Harris detector is provided in [12]. Comparison of interested point detectors is given in [15]. An overview of existing interest point detectors can be found in [9].

Corners are not inherently scale invariant, i.e., a multi-scale Harris detector does not localize the same local structure at the same point in different scale. Mikolajczyk [9] choose the 'correct' scale as maxima of Laplacian-of-Gaussian in the scale space. Deriche [1] approached this problem by fitting the line through the locations at different scales and searched where the series of points converge. Lowe's detector searches maxima of Difference of Gaussian [7] in the scale space.

It was noticed by many that richer structures in images compared to local, 'derivative-based' corners can bring additional benefit. Proposed methods are detectors based on affine normalization around Harris points [9, 14], a detector of 'maximally stable extremal regions' [8], an edge-based region detector [16], a detector based on intensity extrema [17], and the detectors of 'salient regions' based on entropy [5] or wavelet transform [6]. The performance of above region detectors is compared in [3, 10] where the performance to change in viewpoint, scale, illumination, defocus and image compression are considered.

The proposed blob detector was motivated by practical observations and a signal processing theory. Its robustness and speed is derived from basic properties of Fourier transformation, dot product and phase of the first harmonics. A global difference of phase of the first harmonics between two omnidirectional images was used to find their relative orientation [13].

3 Stable Wave Detector

The target patterns in the images are blobs, i.e., local minima or maxima of intensity. In 1D, the blob corresponds to a positive or a negative peak. Let us start our explanation with an experiment in 1D which motivated our approach and gave it the name. The examples of ideal peaks will be shown in Section 3.2. The single scale SWD algorithm will be introduced in Section 3.3. Main components of SWD are described in Sections 3.4 and 3.5. The extension of SWD to multiscale is given in Section 3.6. Section 3.7 describes how to apply SWD to detect blobs in 2D.

3.1 What is Stable Wave?

The word 'wave' comes from Fourier transform which is our basic tool. The term 'stable' says that we seek the wave in the signal (or image) which is stable / well

localized in some sense. The idea of the stable wave originates from a simple experiment in 1D. Let us consider data of length $N > 2T$, $T \geq 4$ containing a peak of width about $T/2$. Then suppose that the peak is located somewhere in the closed interval $[x_0 - T/2, x_0 + T/2]$, where $T < x_0 < (N - T)$. The following iterative algorithm can localize the peak:

1. Define frame F_1, as $[x_0 - T/2, x_0 + T/2 - 1]$.
2. For frame F_i, compute Fourier coefficients a, b and phase $\varphi_i \in (-\pi, \pi]$ of the first harmonic wave (i.e., with a period T).
3. Estimate the peak location relatively to frame F_i as

$$\begin{aligned} x_i &= \tfrac{T}{4}(1 - 2\tfrac{\varphi_i}{\pi})\,;\ \varphi_i < 0\,,\ \text{(maxima)} \\ x_i &= \tfrac{T}{4}(3 - 2\tfrac{\varphi_i}{\pi})\,;\ \varphi_i > 0\,,\ \text{(minima)} \end{aligned} \tag{1}$$

4. Define a new frame F_{i+1} centered around x_i.

Repeat steps 2-4 until the process converges, i.e , until F_{i+1} is the same as F_i (the peak is stable) or diverges $|x_i - x_0| > T/2$ or oscillates $i > T$.

The process typically converges in the first or the second iteration for the ideal peak without noise (Section 3.2). The algorithm works well also for edges for which x_i is a zero crossing instead of an extreme.

An interesting observation is that the precise knowledge of the peak width is not crucial. The algorithm works well up to about an octave below and over the optimal T. The amplitude

$$A = \sqrt{(a^2 + b^2)} \tag{2}$$

can measure the strength of response (suitability) of a period T for a given peak.

3.2 Ideal peaks

Examples of ideal shapes from the SWD point of view are cosine, Gaussian, rectangle and other similar symmetric shapes. The detector localizes peaks of 'ideal shape' with the subsample precision similarly as humans would localize them intuitively when looking at them. For example, the detector finds the maximum of Gaussian or the center of a rectangle.

Figure 1 shows the examples of ideal peaks (solid curves) suitable for SWD with frames of width 16 (line segments bellow the curves). The dotted curves visualize the results of SWD algorithm at variable frame lengths. The dotted curves are cosine waves with the amplitude found by SWD centered at the peak location found by a single scale SWD. The cosine curves are shifted up to have positive values for better visualization. SWD algorithm successfully found peaks even if the frame length was far from the period of the signal or from the double of the peak width.

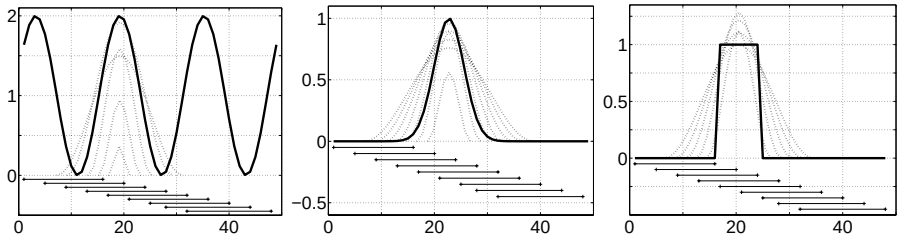


Fig. 1. Examples of ideal peaks (*solid curves*), frames (*line segments*) of length $T = 16$, and stable waves found by SWD at different periods ($T \in \{4, 8, 12, 16, 20, 24, 28, 32\}$, *dotted curves*). a) Sine wave ($T = 16$, $\varphi = 0.69$. b) Gaussian $y = \exp(-(x-22.76)^2/16)$). c) 8-sample-wide rectangle.

3.3 Outline of the single-scale SWD algorithm

The algorithm described in Section 3.1 just motivated SWD. A practical SWD algorithm is sketched in this section for the single-scale. Let us consider data to be a row vector D of length N and the expected width of the peaks be near $T/2$.

The SWD algorithm consists of four forward steps.

1. The input data are divided into n overlapping frames F_i of length T. The frames should overlap more than $T/2$. Efficient choice of the overlap is described in Section 3.4.
2. For each frame F_i, Fourier coefficients a_i, b_i are computed. (Section 3.4)
3. Candidate frames are found as cosine frames meeting Consistent Neighboring Frame (CNFr) criterion (Section 3.5).
4. Having a candidate frame F_i, a subsample location x_i of the peak inside the frame can be found using Equation (1).

3.4 The efficient computation of Fourier coefficients

In the discrete case, Fourier coefficients are computed as dot products $a_i = F_i \cdot S$, $b_i = F_i \cdot C$, where F_i are intensity data in the frame i and

$$S_t = \sin(2\pi(t/T)) , \qquad C_t = \cos(2\pi(t/T)) , \qquad t = 0, \ldots , T-1 .$$

To compute all coefficient a_i, b_i for whole data of the length N, the frame length T, and an arbitrary overlap Ω, the number of multiplication would be

$$\mathcal{O}(N, \Omega) = 2N \frac{T}{T - \Omega} .$$

For example $\mathcal{O}(N, 2/3T) = 6N$, $\mathcal{O}(N, 3/4T) = 8N$, $\mathcal{O}(N, 4/5T) = 10N$.

The choice of the frame overlap is an important decision which influences both speed and precision. Looking at the properties of sine and cosine functions, we

found $3/4T$ to be the best choice. The reason is the well known fact that the sine wave (the first base function) is the cosine wave (the second base function) shifted by a quarter of a period, i.e., $\cos\varphi = \sin(\varphi+\pi/2)$. Exploring this fact, significant amount of computation can be saved for frame shift equal to $T/4 \sim \pi/2$.

Dot product $a_i^1 = F_i \cdot S$ can be written as

$$a_i = \sum_{j=1}^{T} F_i(j)S(j) = D_i^s\left(1,\frac{T}{4}\right) + D_i^s\left(\frac{T}{4},\frac{T}{2}\right) + D_i^s\left(\frac{T}{2},3\frac{T}{4}\right) + D_i^s\left(3\frac{T}{4},T\right),$$

$$b_i = \sum_{j=1}^{T} F_i(j)C(j) = D_i^c\left(1,\frac{T}{4}\right) + D_i^c\left(\frac{T}{4},\frac{T}{2}\right) + D_i^c\left(\frac{T}{2},3\frac{T}{4}\right) + D_i^c\left(3\frac{T}{4},T\right),$$

where D_i^c are dot products of a part of the frame with a part of the cosine wave and similarly D_i^s are dot products of a part of a frame with a part of a sine wave. The two neighboring frames have 3/4T overlap, therefore, six of the eight partial dot products differ only in the signs and it is not necessary to compute them twice. As a result, only $2N$ (instead of $8N$) multiplications are needed for data containing N samples.

3.5 Consistent Neighboring Frame (CNFr) criterion

The principle of the criterion can be illustrated on the example of Gaussian depicted in Figure 1b. The frames of length $T = 16$ and overlap $3/4$ T are shown as line segments below the curves. Table 1 shows the first two Fourier coefficients, phases and amplitudes for the frames of length 16. The frame 4 containing the maximum of the peak is a negative cosine frame ($-b_i > |a_i|$) surrounded by the two sine frames ($|b_i| < |a_i|$). Frames containing rising edge have negative sine coefficients and frames containing falling edge have positive sine coefficients as it could be expected considering the similarity to sine wave. Each cosine frame does not contain a peak. For example falling edge of Gaussian produces cosine frames (frames 7, 8 in Table 1) which do not contain any peak. CNFr criterion can be summarized as:

- A cosine frame F_i contains a stable minimum if
  $b_i > k|a_i|$ & $b_{i+1} > k|a_{i+1}|$ & $a_i > 0$ & $a_{i+1} < 0$ or
  $b_i > k|a_i|$ & $b_{i+1} < k|a_{i+1}|$ & $b_{i-1} < k|a_{i-1}|$ & $a_{i-1} > 0$ & $a_{i+1} < 0$.
- A cosine frame F_i contains a stable maximum if
  $-b_i > k|a_i|$ & $-b_{i+1} > k|a_{i+1}|$ & $a_i < 0$ & $a_{i+1} > 0$ or
  $-b_i > k|a_i|$ & $-b_{i+1} < k|a_{i+1}|$ & $-b_{i-1} < k|a_{i-1}|$ & $a_{i-1} < 0$ & $a_{i+1} > 0$.

The number $k < 1$ (near 1) moderates the criterion on the frame to be considered as cosine. It reduces the chance to loose the candidate due to noise and other perturbations when the both coefficients have similar absolute values. The exact choice of k is not crucial; we used the value 7/8.

Table 1. Gaussian and frames of SWD, a is a sine coefficient, b is a cosine coefficient, φ is the phase relative to the frame.

Frame	1	2	3	4	5	6	7	8	9
a_i	-0.01	-0.12	-0.46	-0.31	0.38	0.42	0.09	0.00	0.00
b_i	0.01	0.11	0.12	-0.38	-0.31	0.25	0.18	0.02	0.00
φ_i	2.17	2.38	2.88	-2.25	-0.70	0.54	1.11	1.34	1.45
Amplitude	0.01	0.16	0.47	0.49	0.49	0.49	0.21	0.02	0.00

3.6 Multi-scale SWD algorithm

A good peak of a certain width has responses for several different SWD periods as can be seen in Figure 1. In the case of a sine wave, the localization precision was better than 0.04 of the sample for $T \in \{8, 12, 16, 20, 24\}$). The strongest response (the amplitude of the stable wave) was at $T = 16$ (i.e., natural period of the signal). Similar results are for Gaussian (the precision was better than 0.03 for $T \in \{8, 12, 16, 20, 24, 28, 32\}$) and for a rectangle (the peak was detected at $T \in \{4, 8, 12, 16, 20, 24, 28, 32\}$, and even zero localization error occurred at $T \in \{12, 16, 20, 24, 28\}$).

Even though SWD algorithm does not require the precise knowledge of the peak width, its estimate must be provided. Such estimate may not be available in many practical situations or the peak width may vary in a wide range. The developed multiscale algorithm searches a hierarchy of peaks at several width levels similarly to the others multiscale detectors [9, 7].

Considering excellent multiscale property of SWD, we can afford 2^n step in scale which is more sparse than 1.4^n used by Mikolajczyk [9], or Lowe[7]. The integer scale allows a more efficient computation compared to a non-integer.

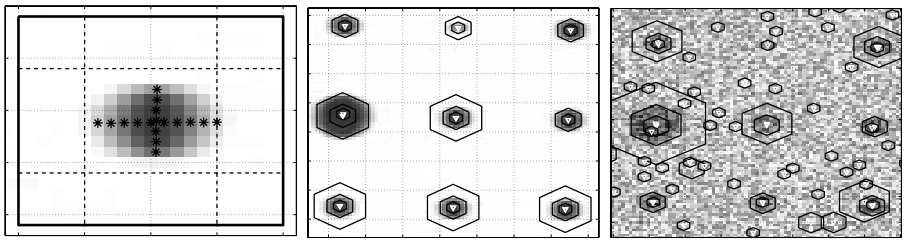
Similarly to [9, 7] we can choose the best scale (or filter out weak peaks) according to the strength of the response measured as the amplitude A, see Equation (2).

The quality of the peak found can be measured as a variance of the locations found at different nearby scales. Good peaks are stable for more than an octave difference of the frame size. We evaluate the stability at approximately a half of an octave lower (T_L) and higher (T_H) than the original scale (T) where the peak was detected. The word 'approximately' means that we always choose an integer period. For example, for the scale with $T = 16$ we choose $T_L = 12$ and $T_H = 24$. We estimate the multiscale stability as $\min(|x_i - x_L|, |x_i - x_H|)$, where x_i is the initial estimate at the period T, x_L is estimated using the frame of length T_L centered around x_i, and x_H is estimated using the frame of length T_H centered around x_i.

3.7 Stable Wave Detector in 2D

Let us observe the example of an ideal blob in 2D in Figure 2a similarly as we observed ideal peaks in 1D. Looking at the intensity along the line (e.g., row

or column of the image) passing a blob, there is a peak. The peak in intensity along the line in the image is a necessary condition for the blob presence. In addition, the peak provides two independent constraints – one for rows, second for columns. As we can see in Figure 2, the results of SWD-1D on the rows passing the blob lie on the line (with subpixel precision). Similarly, the results of SWD-1D on a column passing the blob lie on an another line. The two lines are almost perpendicular. The 2D location of the blob can be estimated as the intersection of these two lines.



(a) (b) (c)

Fig. 2. The image of blobs in a calibration pattern. a) *Black stars* in the middle of the dark patch show the results of SWD-1D along the rows and the columns of the image. b) Results of SWD-2D on the original image. c) Results of SWD-2D on the image with artificially added noise (S/N = 1). The blob locations found by SWD-2D are depicted as *white triangles* and the diameter of *black hexagons* indicates the scale.

More thorough mathematical explanation of this approach can be derived from 2D Fourier Transformation (FT) and phase-based methods of stereo matching. The 2D FT is a vector function giving a phase and an amplitude to a vector $[f_h, f_v]$ (horizontal and vertical frequency). Two independent frequency vectors are needed to obtain a 2D phase information. For better imagination, vector $F_1 = [f, 0]$ corresponds to the wave parallel with axis x. $F_2 = [0, f]$ is the wave parallel to axis y. $F_3 = [f, f]$ is the wave parallel to the line $y = 1 - x$.

Let us look in detail on the integral of Fourier coefficient a for F_1. The base function $S(x, y)$ depends on x only. The integral can be decomposed as

$$a = \int_{x,y} S(x,y)\, I(x,y)\, \mathrm{d}x\, \mathrm{d}y = \int_x s(x)\, i(x)\, \mathrm{d}x\,.$$

$$i(x) = \int_y I(x,y)\, \mathrm{d}y\,.$$

In the discrete case, it means the following. First, sum the intensity I(row, column) over the row to get a function i(column) and then transform function i(column) by 1D FT to get the phase in the horizontal direction.

The phases corresponding to frequency vectors F_1, F_2 can be used to localize the peak similarly as SWD-1D does. However this solution is not good. The

application of CNFr criterion would be difficult. Intuitively, the other problem is that the integrals for Fourier coefficients contain a large neighborhood of the blob. The solid square in Figure 2a shows the optimal square window to detect the blob. The square areas in its corners bring just noise to the integrals.

Our SWD-2D algorithm combines the observation from Figure 2 with the mathematical derivation. In short, we detect the peaks in rows and fit the vertical line v through them. Next, we detect the peaks in columns and fit the horizontal line h through them. The location of the peak is estimated as the intersection of lines h and v. The detailed algorithm is described in [2] due to lack of space.

4 Implementation and experiments

The SWD-2D algorithm has been implemented in MATLAB and tested on synthetic and real data. The experiments are described in detail in [2].

Robustness to noise in 1D case was extensively tested. The results are summarized in Figure 3.

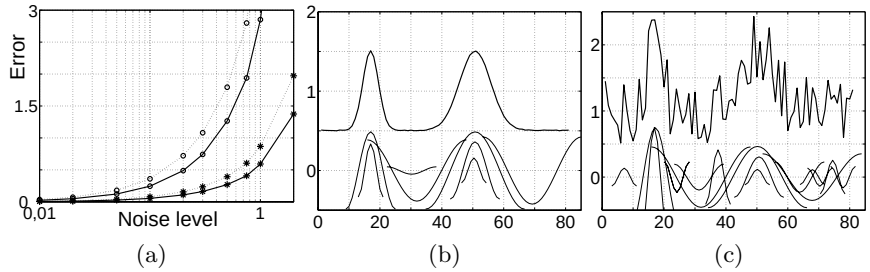
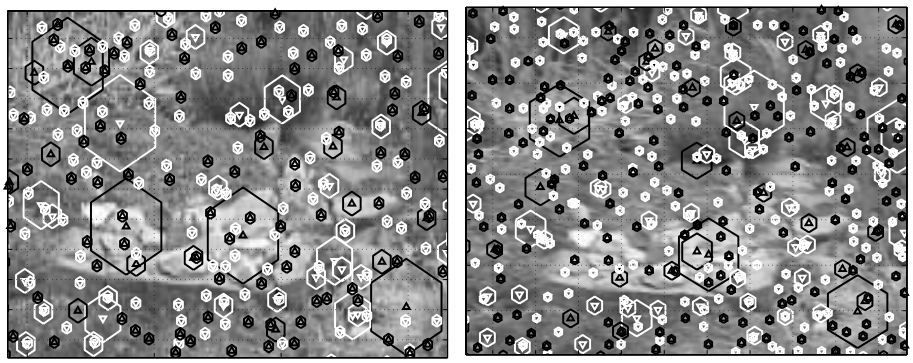


Fig. 3. Robustness of SWD-1D to uniform noise. a) Maximum (*circles*) and average (*stars*) localization error as the function of noise level added to original signal containing two Gaussian peaks (narrow peak – *solid line*, wide peak – *dotted line*). The results come from 100 repeats. b) Example of signal with noise level 0.01. c) Example of signal with noise level 1.

The robustness to noise of WSD-2D was tested on the real image of the camera calibration pattern consisting of black circular dots. The detail of the image is depicted in Figure 2b. The original real image was degraded by the addition of the increasing amount of uniform noise (Figure 2c). The localization proved to be stable to noise level about 1. The achieved precision is below 0.05 of a pixel for original image and below 0.15 of a pixel up to noise level 0.3.

Let us show detected blobs on a real image of a rock garden. The purpose of the experiment is to display visually where blobs are detected. The second aim is to demonstrate the potential of the proposed method for stereo matching. The rock garden is captured from two different view points. The reader can see that many blobs in the first image have a corresponding partner in the second image.



(a) (b)

Fig. 4. Illustration of SWD on the rock garden image. The blob locations found by SWD-2D are depicted as *triangles* and the diameter of *hexagons* indicates the scale. *White* color is used for minima, *black* for maxima. (a) The detail of the first image; (b) The detail of the second image, taken from different place.

5 Conclusions and future work

We have proposed a new blob detector which seems to have several favorable properties for practical applications. It is precise and fast. We believe that it is a right way to our application target – implementing it in the batter-powered hand-held instrument for on-line 3D positioning of geophysical measurements.

The semantic-less interest point-like detectors have had an enormous attention in the computer vision community in last few years. Our thorough comparison to them needs our further attention. We need to move from the MATLAB implementation to C language to be able to perform computation time tests and comparisons with other algorithms and implementations. We also would like to extend our method to cope with color images.

Acknowledgments. The authors were supported by The Czech Ministry of Education under Project 1M0567.

References

1. R. Deriche and G. Giraudon. A computational approach for corner and vertex detection. *IJCV*, 10(2):101–124, 1993.
2. J. Dupač and V. Hlaváč. Stable wave detector for precise and fast detection of blobs in the image. Research Report CTU–CMP–2006–03, Center for Machine Perception, K13133 FEE Czech Technical University, Prague, Czech Republic, April 2006.
3. F. Fraundorfer and H. Bischof. Evaluation of local detectors on non-planar scenes. In *Proc. of 28th Workshop of the Austrian Association for Pattern Recognition*

(AGM/AAPR), pages 125–132, Osterreichische Computer Gesellschaft 3-85403-179-3, Hagenberg, 2004.
4. C. Harris and M. Stephens. A combined corner and edge detector. In *Proc. of 4th Alvey Vision Conference*, pages 147–151, March 1988.
5. T. Kadir, A. Zisserman, and M.Brady. An affine invariant salient region detector. In *Proceedings of the 8th European Conference on Computer Vision*, volume 1 of *LNCS 3021*, pages 228–241, Prague, May 2004. Springer-Verlag.
6. E. Loupias and N. Sebe. Wavelet-based salient points: Applications to image retrieval using color and texture features. In *VISUAL '00: Proceedings of the 4th International Conference on Advances in Visual Information Systems*, pages 223–232, London, UK, 2000. Springer-Verlag.
7. D. G. Lowe. Object recognition from local scale-invariant features. In *Proc. of IEEE International Conference on Computer Vision (ICCV1999)*, pages 1150–1157. IEEE Computer Society, 1999.
8. J. Matas, O. Chum, Urban M., and T. Pajdla. Robust wide baseline stereo from maximally stable extremal regions. In Paul L. Rosin and David Marshall, editors, *Proc. of the British Machine Vision Conference*, volume 1, pages 384–393, London, UK, September 2002. BMVA.
9. K. Mikolajczyk and C. Schmid. Scale & affine invariant point detector. *International Journal of Computer Vision*, 60(1):63–86, 2004.
10. K Mikolajczyk, T Tuytelaars, C Schmid, A Zisserman, J Matas, F Schaffalitzky, T Kadir, and L van Gool. A comparison of affine region detectors. *International Journal of Computer Vision*, 65(7):43 – 72, November 2005.
11. H. P. Moravec. Towards automatic visual obstacle avoidance. In *Proc. of The 5th International Joint Conference on Artificial Intelligence, MIT, Cambridge, Massachusetts*, page 584. IJCAI, August 1977.
12. D. Nistér, O. Naroditsky, and J. R. Bergen. Visual odometry. In *Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'04)*, pages 652–659. IEEE Computer Society, 2004.
13. T. Pajdla and V. Hlaváč. Zero phase representation of panoramic images for image based localization. In Franc Solina and Aleš Leonardis, editors, *Proc. of 8-th International Conference on Computer Analysis of Images and Patterns*, number 1689 in Lecture Notes in Computer Scinence, pages 550–557, Tržaška 25, Ljubljana, Slovenia, September 1999. Springer Verlag.
14. F. Schaffalitzky and A. Zisserman. Multi-view matching for unordered image sets, or 'how do i organize my holiday snaps?'. In *Proceedings of the 7th European Conference on Computer Vision*, volume 1 of *LNCS 2350*, pages 414–431. Springer-Verlag, 2002.
15. C. Schmid, R. Mohr, and C. Bauckhage. Evaluation of interest point detectors. *International Journal of Computer Vision*, 37(2):151–172, 2000.
16. T. Tuytelaars and L. Van Gool. Content-based image retrieval based on local affinely invariant regions. In *International Conference on Visual Information Systems*, pages 493–500, 1999.
17. T. Tuytelaars and L. Van Gool. Matching widely separated views based on affine invariant regions. *International Journal on Computer Vision*, 59(1):61–85, 2004.