

## FEATURE POINT EXTRACTION USING SCALE-SPACE REPRESENTATION

*Y. Abdeljaoued and T. Ebrahimi*

Swiss Federal Institute of Technology (EPFL)  
Signal Processing Institute (ITS)  
CH-1015 Lausanne, Switzerland  
{yousri.abdeljaoued,touradj.ebrahimi}@epfl.ch

### ABSTRACT

An algorithm for feature point extraction is presented. It is based on a scale-space representation of the image as well as a system for tracking across scales. Using synthetic and real images, it is shown that the proposed algorithm produces stable and well-localized feature points estimates – two essential properties for video applications.

### 1. INTRODUCTION

Feature point extraction is an important task in computer vision and image processing. It provides input information for further operations, such as image registration, structure from motion [1], image retrieval [2], and motion estimation. A number of requirements should be fulfilled by any feature point extraction algorithm [3]:

- All true feature points should be detected.
- No false feature points should be detected.
- Feature points should be well localized.
- The algorithm should be robust with respect to noise.
- The algorithm should be efficient.

These requirements can also be expressed as good *repeatability* and high *localization accuracy*. *Repeatability* gives an idea about the geometrical stability of the extracted feature points between different images of a given scene taken under changes in the viewing conditions, such as the camera parameters, and the camera noise. *Localization accuracy* measures whether a feature point is accurately located at the projection of a 3D scene point in the image plane. These two criteria are significant for a large number of applications.

Several algorithms for feature point extraction have been reported in the literature in the last years. They can be broadly divided into two categories: contour-based [4] [3] and intensity based [5] [6] methods. The former category involves first extracting the contours and then searching for

points having maximal curvature along these contours. The latter category works directly at the gray level to detect the feature points. Most of these algorithms did not address the scale problem. Usually, feature points are assumed to be perfectly sharp (i.e. ideal points), and thus appear in the same way at all the scales. However, this assumption is violated by noise, blurring, and quantization. Furthermore, objects, both man-made and natural, often have rounded feature points, which begin to appear at different scales.

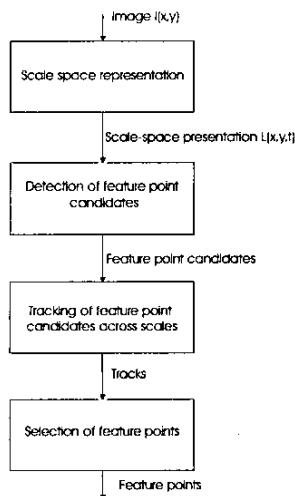
We propose a novel algorithm for feature point extraction. It is based on the scale-space representation. That is a one-parameter family of smoothed images. One of the main problems in this task is the selection of the scale at which the operator for the extraction of feature points is computed. Rather than selecting a single scale, feature points are extracted at different scales, and are then tracked across scales. This tracking method provides an efficient way for combining the information obtained at different scales. This helps to accurately estimate the locations of the feature points as well as to make the extraction robust to noise.

The proposed algorithm is described in Section 2. The results of the evaluation of the proposed algorithm with respect to repeatability and localization accuracy are presented in Section 3.

### 2. PROPOSED ALGORITHM FOR THE EXTRACTION OF FEATURE POINTS

Figure 1 shows a block diagram of the proposed algorithm. First the scale-space representation of the input image is generated. It corresponds to a one parameter family of smoothed images, where fine scale details are successively suppressed. Then, a differential operator is used to detect feature point candidates in the scale space representation. They are equivalent to the maxima of the absolute value of this measure. One of the main issues in multiscale algorithms is how to combine the information obtained at different scales in order to extract the feature points. We propose to track the feature point candidates across scales. In contrast to maxima due to noise which are tracked over few scales, maxima

which originate from feature points are maintained over a large scale range, also called *track lifetime*. We propose to use the track lifetime for the selection of feature points.



**Fig. 1.** Block diagram of the scale-space based algorithm for feature point extraction.

### 2.1. Scale space representation

The scale space representation [7] is defined as follows. Given the intensity function  $I : \mathbb{R}^2 \rightarrow \mathbb{R}$ , its scale-space representation  $L : \mathbb{R}^2 \times \mathbb{R}_+ \rightarrow \mathbb{R}$  is:

$$L(x, y, t) = g(x, y, t) \otimes I(x, y), \quad (1)$$

where  $g : \mathbb{R}^2 \times \mathbb{R}_+ \rightarrow \mathbb{R}$  denotes the Gaussian Kernel with variance (scale)  $t$ :

$$g(x, y, t) = \frac{1}{2\pi t} \exp -\frac{x^2 + y^2}{2t}. \quad (2)$$

### 2.2. Detection of feature point candidates

We selected the following differential operator [6] to extract feature points:

$$\tilde{K}_{norm}(x, y, t) = t^2(L_{xx}L_y^2 + L_{yy}L_x^2 - 2L_{xy}L_xL_y), \quad (3)$$

where  $L_x$  denotes the first derivative of  $L(x, y)$  in  $x$  (analogously for  $L_y$ ) and  $L_{xx}$  denotes the second derivative of  $L(x, y)$  in  $x$  (analogously for  $L_{yy}$  and  $L_{xy}$ ). This choice is motivated by the fact that this normalized operator is suited to be integrated in a multiscale approach for feature point extraction.

The use of the normalized operator allows us to select a set of *feature point candidates* at different scales by applying a thresholding process to the maxima of the scale-space presentation.

### 2.3. Tracking of feature point candidates through scales

An important issue in any multiscale approach for feature point extraction is how to combine the local spatial maxima at different scales in order to obtain an accurate estimate of the location of a feature point. Considering only fine scales, where a large amount of noise can be expected to be present, increases the number of false positives. At coarser scales, the scale-space smoothing removes the noise. However, this smoothing could affect feature point localization and the interference between adjacent feature points.

Rather than selecting a single scale from all scales for feature point extraction [8], we propose to partition the feature point candidates into tracks across scales and to integrate all the locations of the same track using a Kalman filter. The goal of this part is to track the feature point candidates across scales. A tracking system of feature point candidates across scales is designed. In the following, each basic step of the proposed tracking system is explained:

- **State estimation** — By applying the Kalman filter to the feature point candidates starting from the coarse scales down to the fine scales, the localization is improved by removing noise.
- **Data association** — Data association is equivalent to the correspondence between the tracks established at scale  $t$  and the feature point candidates at scale  $t - 1$ . The nearest neighbor filter has been selected to perform data association because it is simple and efficient.
- **Track management** — Track initiation and track deletion are carried out in the track management step. An unassociated feature point candidate is used to initiate a new track, whereas a track which is not updated with a measurement is deleted.

### 2.4. Selection of feature points

The tracks obtained by the tracking system are used as input to this step. Tracks which originate from noise have a short lifetime, whereas tracks which are related to real feature points are stable and have a long lifetime. Therefore, we use the lifetime as criterion for the selection of feature points that are characterized by a long lifetime.

## 3. EVALUATION OF THE SCALE-SPACE BASED ALGORITHM

In this section, we give some experimental results obtained by running the proposed algorithm on synthetic and real images. The state-of-the-art algorithm proposed by Harris [5] serves as the reference. The reason behind this choice is that

this algorithm showed the best results in terms of repeatability and localization according to [9].

### 3.1. Synthetic images

The proposed algorithm is evaluated with respect to the criteria repeatability and localization. A set of synthetic images has been used for the evaluation. As a result of this, the correct feature points, including their positions, are known. This allows us to check whether the extracted feature points are false, missed, or correct detections. Further, it is possible to compute the deviation between the extracted position and the correct position of a feature point.

#### 3.1.1. Repeatability

Let us define the *repeatability rate* as the number of feature points repeated between two consecutive images divided by the total number of feature points. Due to the uncertainties in the estimated position of a feature point, a circle centered at the correct feature point with a radius  $\epsilon$  is defined to model these uncertainties [9]. An extracted feature point which lies inside this circle is considered as a correct detection. The repeatability within a circle of radius  $\epsilon$  is called  $\epsilon$ -repeatability. Figure 2 shows a reference synthetic image containing a set of polygons on a 2D plane. These polygons have sharp feature points as well as slightly rounded feature points. We have added Gaussian noise to this synthetic image in order to model the camera noise.

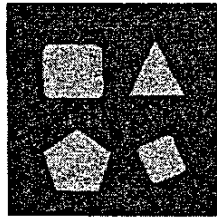


Fig. 2. Reference synthetic image used for evaluation.

Different rotations are applied to this reference synthetic image in order to simulate some camera operations. Such transformations are common between consecutive images in a video sequence. Knowing the correct positions of the feature points in the reference image as well as the geometric transformations, the correct positions of the feature points in the transformed images are computed by applying these transformations to the feature points in the reference image.

The repeatability rate for  $\epsilon = 3$  pixels is displayed in Figure 3 as a function of the rotation angle. We note that the scale-space algorithm is more robust to rotation. This can be explained by the invariance of the differential operator Eq. (3) under affine transformations. Thanks to the use

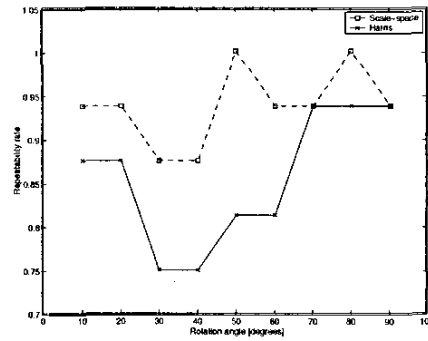


Fig. 3. Repeatability rate with respect to the rotation angle for the Harris algorithm (solid curve) and for the proposed scale-space based algorithm (dashed curve).

of the scale-space presentation, the number of false detections is greatly reduced in the feature points extracted by the scale-space based algorithm.

#### 3.1.2. Localization

The localization criterion measures the accuracy of the position of an extracted feature point compared to the position of the corresponding correct feature point. This criterion can be expressed by the root mean square error (RMSE) given by:

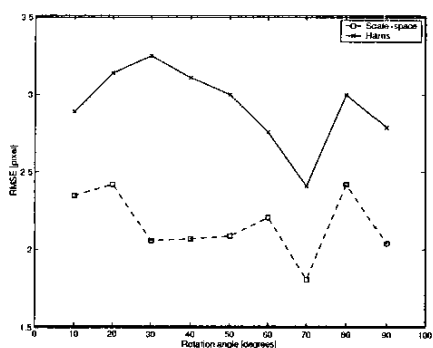
$$RMSE = \sqrt{\frac{1}{N} \sum_{n=1}^N [(x_{n0} - x_n)^2 + (y_{n0} - y_n)^2]}, \quad (4)$$

where  $x_{n0}$  and  $y_{n0}$  are the coordinates of the  $n^{th}$  correct feature point,  $x_n$  and  $y_n$  are the coordinates of the  $n^{th}$  extracted feature point, and  $N$  is the number of extracted feature points.

Figure 4 shows the variation of the RMSE with respect to the rotation angles of the transformed images. The scale-space based algorithm clearly outperforms the Harris algorithm over all rotation angles in terms of localization (i.e. lower RMSE). The scale-space algorithm owes this good performance to the use of the tracking system in the scale-space representation. By applying the Kalman filter, the noise influence is reduced and the position of the detected feature point is estimated more accurately, especially the rounded feature points, which start to appear at high scales.

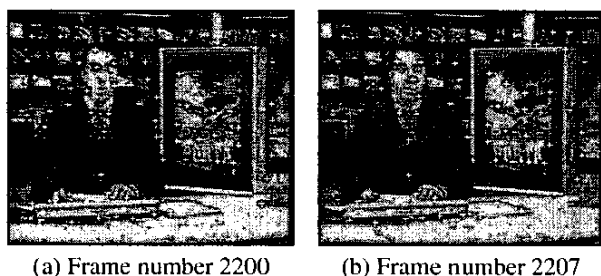
### 3.2. Real images

To illustrate the efficiency of the proposed algorithm on real data, Figure 5 shows an example of good results in terms of repeatability and localization obtained by running the algorithm on real images. More specifically, many feature points are well localized and repeated even after 7 frames.



**Fig. 4.** RMSE vs. rotation angle for the Harris algorithm (solid curve) and for the scale-space based algorithm (dashed curve).

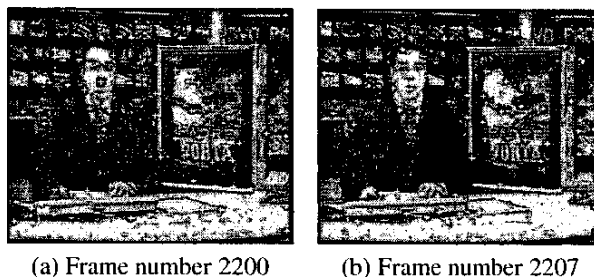
Figure 6 shows the feature points extracted by the Harris algorithm. We note that the feature points are not stable in the different frames. Consequently, the performance of any video characterization technique which relies on the tracking of feature points degrades.



**Fig. 5.** Feature points extracted by the proposed scale-space based algorithm.

#### 4. CONCLUSIONS

We have presented an algorithm for feature point extraction using a scale-space representation as well as a tracking system across scales. Feature point candidates are detected in the scale-space representation. These candidates correspond to local maxima of a differential operator suited for this step. The tracking system allows us to form and maintain tracks across scales. Tracks with an extended lifetime are selected as feature points. The evaluation of this algorithm with respect to repeatability and localization accuracy on synthetic and real images shows that it outperforms the Harris algorithm. A possible extension to this algorithm would be to use a track quality indicator to select the feature points from the whole track set.



**Fig. 6.** Feature points extracted by the Harris algorithm.

#### 5. REFERENCES

- [1] C. Tomasi and T. Kanade, "Shape and motion from image streams under orthography: A factorization method," *Int. Journal of Computer Vision*, vol. 9, no. 2, pp. 137–154, November 1992.
- [2] S. Bhattacharjee, *A Computational Approach to Image Retrieval*, Ph.D. thesis, EPFL, Swiss Federal Institute of Technology, LTS-DE, 1015 Lausanne, 1999.
- [3] F. Mokhtarian and R. Suomela, "Robust image corner detection through curvature scale space," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 20, no. 12, pp. 1376–1381, December 1998.
- [4] R. Deriche and O. Faugeras, "2-D curve matching using high curvature points: Application to stereo vision," in *Proc. 10th International Conf. on Pattern Recognition*, 1990, pp. 240–242.
- [5] C. Harris and M. Stephens, "A combined corner and edge detector," in *Alvey Vision Conf.*, 1988, pp. 147–151.
- [6] T. Lindeberg, *Scale-Space Theory in Computer Vision*, Kluwer, 1994.
- [7] A. K. Witkin, "Scale-space filtering," in *Proc. 8th Int. Joint Conf. Artificial Intelligence*, 1983, pp. 1019–1022.
- [8] T. Lindeberg, "Feature detection with automatic scale selection," *International Journal of Computer Vision*, vol. 30, no. 2, pp. 77–116, 1998.
- [9] C. Schmid, R. Mohr, and C. Bauckhage, "Comparing and evaluating interest points," in *Proc. 6th International Conf. Computer Vision*, 1998, pp. 230–235.