# COMPARISON AND EVALUATION OF FEATURE POINT DETECTORS

V. Rodehorst [a,*] and A. Koschan [b]

[a] Computer Vision & Remote Sensing, Berlin University of Technology, Franklinstr. 28/29, FR 3-1,
D-10587 Berlin, Germany - vr@cs.tu-berlin.de

[b] Imaging, Robotics, and Intelligent Systems, The University of Tennessee, 330 Ferris Hall,
Knoxville, TN 37996-2100, USA - akoschan@utk.edu

**ABSTRACT:**

Low-level feature extraction is the first step in any image analysis procedure and is essential for the performance of stereo vision and object recognition systems. Research concerning the detection of corners, blobs and circular or point like features is particularly rich and many procedures have been proposed in the literature. In this paper, several frequently used methods and some novel ideas are tested and compared. We measure the performance of the detectors under the criteria of their detection and repeatability rate as well as the localization accuracy. We present a short review of the major interest point detectors, propose some improvements and describe the experimental setup used for our comparison. Finally, we determine which detector leads to the best results and show that it satisfies the criteria specified above.

## 1. INTRODUCTION

It has been shown that local features are well suited to automatic image matching and object tracking tasks. Interest operators extract salient image features, which are distinctive in their neighborhood and are reproduced in corresponding images in a similar way. At the same time, interest operators supply one or more characteristics, which can be used during the later image matching. The meaning of the extracted characteristics however depends on the context and therefore interest points do not necessarily correspond to physical corners in the scene.

### 1.1 Requirements

Initially, it is necessary to define the requirements to an optimal interest operator. As criteria for a distinctive matching candidate the characteristics proposed by Haralick and Shapiro (Haralick & Shapiro, 1992) would suit our purpose:

- *Distinctness*: An interest point should stand out clearly against the background and be unique in its neighborhood.
- *Invariance*: The determination should be independent of the geometrical and radiometrical distortions.
- *Stability*: The selection of interest points should be robust to noise and blunders.
- *Uniqueness*: Apart from local distinctiveness an interest point should also possess a global uniqueness, in order to improve the distinction of repetitive patterns.
- *Interpretability*: Interest values should have a significant meaning, so that they can be used for correspondence analysis and higher image interpretation.

These properties make interest points very successful in the context of feature based image matching and the temporal analysis of image sequences. While the characteristics of distinctiveness, invariance and stability define the substantial requirements to an interest operator, the characteristics of the uniqueness and interpretability intensify the meaning of the term 'interesting'.

### 1.2 State-of-the-Art

One of the first interest operators was developed by Moravec (Moravec, 1977). Since then a variety of publications have appeared to this topic. A comprehensive overview of the current methods for the extraction of point features can be found e.g. in Schmid (Schmid et al., 2000). The existing approaches can be divided into three categories:

- *Intensity-based*: The computation of the characteristic for the presence of a salient feature comes directly from the intensity values.
- *Contour-based*: These methods extract an outline and look for places with maximum curvature or make a polygonal approximation of the contour and detect the intersections.
- *Model-based*: By fitting of parametric intensity models to image patterns localization with sub-pixel accuracy can be achieved.

The contour-based methods are critical in the neighborhood of crossings, since edge extraction often produces interruptions or wrong connections in places, where three or more edges meet (s. Figure 10). The model-based variants are limited, depending on the used knowledgebase, to special structures (e.g. L-corners). In practice, the intensity-based methods are most common for general application.

A number of experiments were performed to evaluate interest point detectors. Schmid (Schmid et al., 2000) accomplished a practical comparison of interest operators using the authors' original implementations. The operators of Förstner (Förstner, 1994), Cottier, Heitger, Horaud as well as Harris (Harris & Stephens, 1988) were evaluated quantitatively. It was found that the Harris operator was the most stable of all. In Mikolajczyk (Mikolajczyk & Schmid, 2004) the Harris detector was combined with a *Laplacian-based scale selection* and extended to deal with affine transformations. Hall (Hall et al., 2002) formalized a definition of *saliency* under scale changes and evaluated the Harris, Lindeberg (Lindeberg, 1998) and Harris-Laplacian corner detectors as well.

Johansson (Johansson & Söderberg, 2004) found that the star pattern method and the 4th order tensor perform better than the Harris detector. Köthe (Köthe, 2003) improved the *structure tensor* computation using an increased resolution and non-linear averaging to optimize the localization accuracy.

Lowe (Lowe, 2004) described image feature generation with the *scale invariant feature transform* (SIFT). Mikolajczyk (Mikolajczyk & Schmid, 2003) compare SIFT descriptors, steerable filters, differential and moment invariants, complex filters and cross-correlation for different types of interest points. They observed that SIFT descriptors perform best and steerable filters come second. Sojka (Sojka, 2003) used *Bayesian estimations* to measure the probability that an image area contains a corner candidate. He demonstrated the improvements in contrast to Harris, SUSAN (Smith & Brady, 1997), Deriche-Giraudon (Deriche & Giraudon, 1993), Beaudet, Noble and Kitchen-Rosenfeld corner detector approaches. Zuliani (Zuliani et al., 2004) proposed a unifying description and *mathematical comparison* of the Harris, Noble, Kanade-Lucas-Tomasi and Kenney point detectors.

However, the selection of an optimal procedure remains difficult, since the results substantially depend on the respective implementation (Rodehorst, 2004).

## 2. INTEREST OPERATORS

In order to find the methodical advantages, the two most preferred interest operators in digital photogrammetry and computer vision, the Förstner operator and the Plessey point detector, were implemented under comparable conditions.

### 2.1 Differential Operators

In the following differential operators, which assign the derivative to a function, are described. They are used to determine the magnitude and the direction of intensity value changes in an image. In contrast to the rough approximation by direct gray value differences or to the application of discrete $3 \times 3$ masks, e.g. the classical Sobel operator, adjustable continuous filters are recommended.

Canny (Canny, 1986) described a continuous filter function for stage edges, which is approximated by the first derivative of the Gaussian. The convolution and differentiation are linear shift-invariant operations, to which associative and commutative laws apply. Therefore, the derivative of a smoothed image corresponds to a convolution of the image with the derivative of the Gaussian.

Given a two-dimensional image function $f(x, y)$, the *gradient*

$$\nabla f = \left( f_x, f_y \right)^{\mathrm{T}} \tag{1}$$

is defined by the *spatial derivatives* of *f* in x- and y-direction

$$f_x = f * G_x \quad \text{and} \quad f_y = f * G_y, \tag{2}$$

where * denotes the discrete *convolution*

$$f(x, y) * g(x, y) = \sum_{i=-r}^{r} \sum_{j=-r}^{r} f(x-i, y-j) \cdot g(i, j) \tag{3}$$

of the image *f* with a filter kernel *g* and

$$G_x(x, y) = \frac{\partial G_\sigma(x, y)}{\partial x} = -\frac{x}{2\pi\sigma^4} \cdot \exp\left( -\frac{x^2 + y^2}{2\sigma^2} \right),$$

$$G_y(x, y) = \frac{\partial G_\sigma(x, y)}{\partial y} = -\frac{y}{2\pi\sigma^4} \cdot \exp\left( -\frac{x^2 + y^2}{2\sigma^2} \right) \tag{4}$$
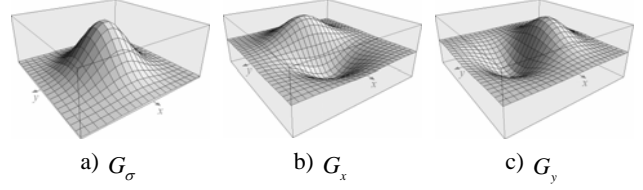


Figure 1. Gaussian filter mask using standard deviation $\sigma = 1.4$ and the two spatial derivatives

denotes spatial derivatives in x- and y-direction of a *Gaussian*

$$G_\sigma(x, y) = \frac{1}{2\pi\sigma^2} \cdot \exp\left( -\frac{x^2 + y^2}{2\sigma^2} \right) \tag{5}$$

with *standard deviation* $\sigma$. An example of the filter kernels is shown in Figure 1.

Deriche's (Deriche, 1990) implementation of Canny's method ensures a constant computation cost using a fixed convolution kernel. In this case, the selection of certain influence areas is achieved by recursive function calls. An extension of Canny's quality criterion around the neighboring edge points shows however that the filter response of the first derivative of a Gaussian is superior to the localization quality of Deriche's operator (Tagare, 1990).

Depending on the influence of the filter, areas outside the image must be accessed for convolution. These *boundary problems* can be eliminated, extending the image borders by the mask radius *r* and mirroring the intensity values. The computation costs for the determination of the derivatives depends on the dimension of the filter kernel. The *radius r* of the filter masks can automatically be estimated as a function of the standard deviation $\sigma$ with $r = \sqrt{2\pi}\sigma$.

The continuous differential operators are steered only by the parameter $\sigma$ within a typical range from 0.5 to 3.0, whereby the smoothing is affected. Using local statistics this parameter can also be estimated (Förstner, 1994). Moreover, for the realization, the operators were *separated*. Using the property

$$\exp\left( -\frac{x^2 + y^2}{2\sigma^2} \right) = \exp\left( -\frac{x^2}{2\sigma^2} \right) \cdot \exp\left( -\frac{y^2}{2\sigma^2} \right) \tag{6}$$

a decomposition of the two-dimensional filter in two linear filters can be achieved. A line-by-line convolution, followed by a column-wise processing, reduces the squared computation costs $O(N^2)$ to a linear measure $O(N)$.

### 2.2 Plessey Point Detector

Harris (Harris & Stephens, 1988) described an improvement of the classical Moravec operator. He solved the problem of the discrete shifts and directions with the help of the autocorrelation function and increased the accuracy of the localization. The idea of autocorrelation exists in a statistical similarity comparison (correlation) of an image window shifted slightly in relation to the original image. The window contains a significant point feature if the similarity for each shift in the neighborhood decreases.

The autocorrelation matrix **A** is computed by summation of the first derivative of the image function *f* over the area $\Omega$ around each image location

$$\mathbf{A}(x, y) = \begin{bmatrix} \sum_{i,j \in \Omega} f_x(i, j)^2 & \sum_{i,j \in \Omega} f_x(i, j) \cdot f_y(i, j) \\ \sum_{i,j \in \Omega} f_x(i, j) \cdot f_y(i, j) & \sum_{i,j \in \Omega} f_y(i, j)^2 \end{bmatrix}, \tag{7}$$

where, in contrast to the original description, the partial derivatives $f_x$ and $f_y$ are determined with the continuous differential operators of section 2.1. In this case the size of $\Omega$ can be automatically determined by the used standard deviation $\sigma$. The matrix **A** finally describes the neighborhood structure in each pixel and has the following characteristics:

- Rank 2: A full rank indicates a salient point
- Rank 1: A singular matrix suggests a straight edge
- Rank 0: The matrix defines a homogeneous area

The Plessey point detector determines the point weight $w$ from the autocorrelation matrix **A** with the *corner response function*

$$w = \det(\mathbf{A}) - k \cdot \operatorname{trace}(\mathbf{A})^2 \cdot \tag{8}$$

In order to receive a separation of the points from edges, the parameter $k$ is selected empirically between 0.04 and 0.06. This yields to positive values at points and to negative values in case of straight edges. The position of an interest point is finally determined by local non-maxima suppression.

## 2.3 Förstner Operator

Förstner (Förstner, 1994; Förstner & Gülch, 1987) also identifies salient points by use of the autocorrelation matrix **A**. First, the derivatives are computed on the smoothed image with the natural scale $\sigma$, and are then summed over a Gaussian window using an artificial scale $\sigma_2$ with the *structure tensor*

$$\mathbf{A} = G_{\sigma_2} * \begin{bmatrix} f_x^2 & f_x \cdot f_y \\ f_x \cdot f_y & f_y^2 \end{bmatrix}, \tag{9}$$

where the indices of the sums over the area $\Omega$ were omitted for simplicity. In our implementation, the continuous filter functions of section 2.1 with only one parameter were used.

Contrary to Harris, Förstner takes the two eigenvalues $\lambda_1$ and $\lambda_2$ of the inverse of **A** as interest value into account. They define axes of an error ellipse. By the computation of their size

$$w = \frac{\lambda_1 \cdot \lambda_2}{\lambda_1 + \lambda_2} = \frac{\det(\mathbf{A})}{\operatorname{trace}(\mathbf{A})}, \quad w > 0 \tag{10}$$

and the form factor for roundness

$$q = 1 - \left(\frac{\lambda_1 - \lambda_2}{\lambda_1 + \lambda_2}\right)^2 = \frac{4 \cdot \det(\mathbf{A})}{\operatorname{trace}(\mathbf{A})^2}, \quad 0 \le q \le 1 \tag{11}$$

the following properties can be derived:

- Small circular ellipses define a salient point
- Elongated error ellipses suggest a straight edge
- Large ellipses mark a homogeneous area

An interest point is present exactly if the given threshold values $w_{min}$ and $q_{min}$ are exceeded. Suitable parameters for this lie within the range

$$w_{min} = (0.5 \ldots 1.5) \cdot \overline{w} \quad \text{and} \quad q_{min} = 0.5 \ldots 0.75, \tag{12}$$

where $\overline{w}$ denotes the mean value of $w$ over the entire image. For illustration, the intermediate results produced by the Förstner operator are shown in Figure 2, and the result of the feature extraction in Figure 11.

## 2.4 SUSAN-2D Operator

For an independent comparison, the SUSAN-2D (Smallest Univalue Segment Assimilation Nucleus) operator proposed by



a) Image of Berlin-Palace      b) Facade section $f$

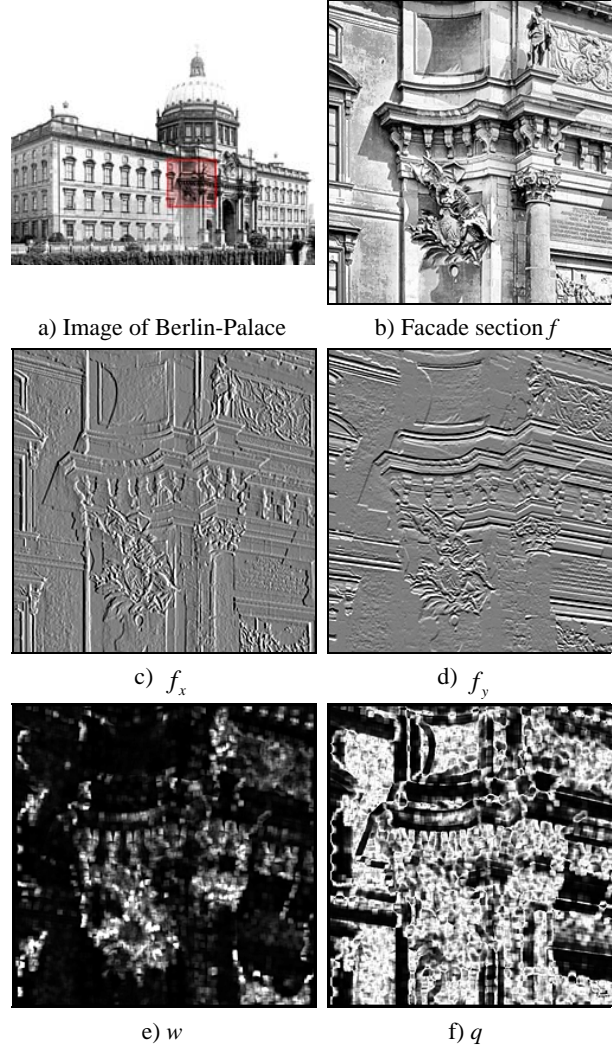c) $f_x$      d) $f_y$

e) $w$      f) $q$

Figure 2. Intermediate results of the Förstner operator

Smith (Smith & Brady, 1997) was used without modifications. The operator is popular due to the freely available source code. It essentially compares the brightness of each pixel within a circular mask with the middle pixel, in order to determine an area, which possesses a similar brightness as the center. In this area the point characteristics are derived by determining the size, the centroid and the $2^{nd}$ order moment.

## 3. COMPARISON AND EVALUATION

Performance evaluation has gained more and more importance in computer vision. In the following the implementations of the interest operators are contrasted with one another. To measure the properties quantitatively, some evaluation criteria must be introduced. We compare the performance of the detectors under the criteria of:

- *Detection Rate*: The detection rate is a measure of the number of true feature points found relative to the number of false detections and missed features.
- *Repeatability Rate*: The points must be obtained independent of varying image conditions. The repeatability rate evaluates the stability under different geometric and radiometric transformations, as well as the effect of noise. The results are shown on natural imagery.

- *Localization Accuracy*: This criterion is most often used to evaluate interest points. The exact position of the features is significant for tasks like camera calibration and 3D reconstruction. The comparison with precise 3D properties is difficult because the camera may introduce a systematic bias. Therefore, the sub-pixel location of features is evaluated using artificial images.

First the efficiency of the three test candidates are compared using standardized reference images. The *detection rate* of the feature extraction for the synthetic test patterns is illustrated in Figure 3. The coloring of the interest points corresponds to the computed point weight with the own implementations.

The advantages of the SUSAN-2D operator are fast computation and outstanding results with unimpaired and unsmoothed images, in particular the own test pattern (see Figure 3.a).

The Plessey point detector reacts sensitively to the sharp raster structures of the image. Unfavorable is also that for the used definition of the point weight, no threshold value could be found which separates the important from the uninteresting points (see Figure 3.b). The Förstner operator shows clearly better results regarding error ellipses as measure for the interest values. The requirement of an approximately round ellipse with $q_{min} = 0.5$, however, suppresses points at straight edges, so that
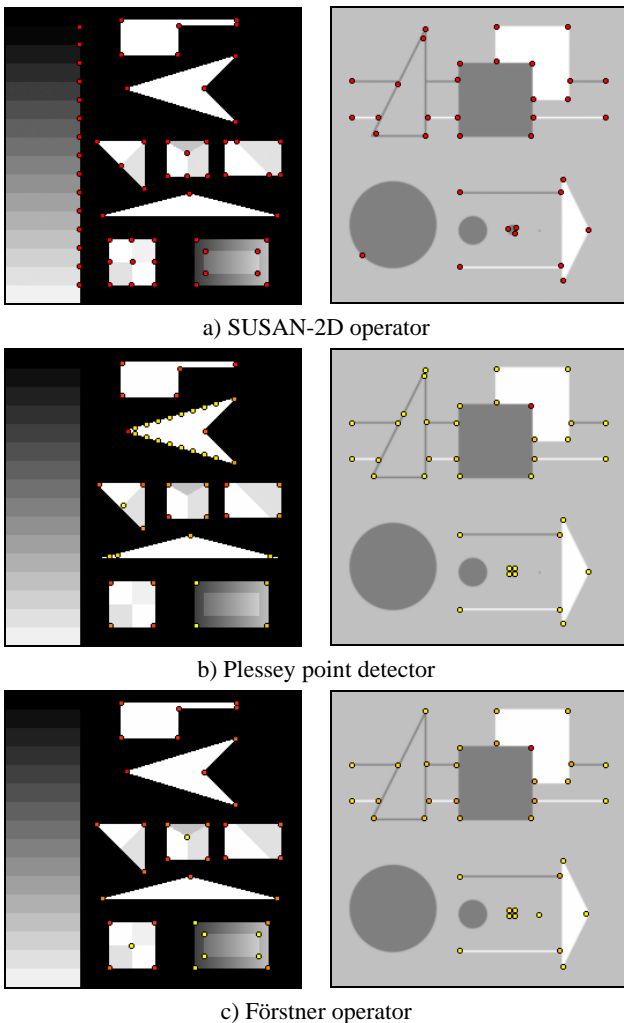


a) SUSAN-2D operator



b) Plessey point detector



c) Förstner operator

Figure 3. Results of the interest operators with the reference image of Smith (left) and Rosenthaler (right)



a) Intensity 25%  b) Intensity 175%  c) Contrast 25%  d) Contrast 175%

e) Scaling 0.4   f) Viewing angle 60°

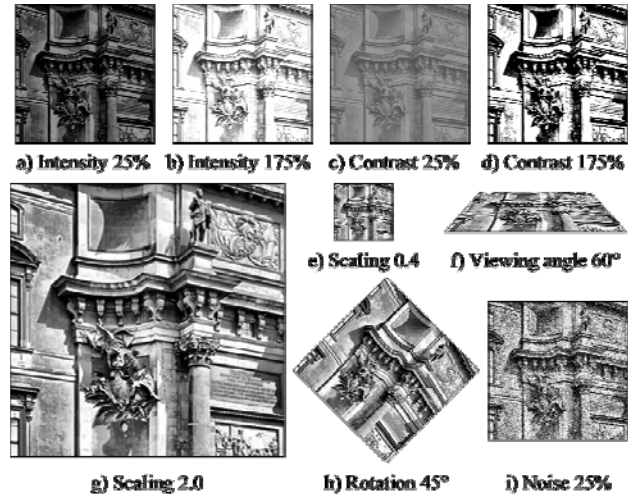g) Scaling 2.0   h) Rotation 45°   i) Noise 25%

Figure 4. Transformed test pattern for the quantitative comparison of the interest operators

fewer features are detected than e.g. with the SUSAN-2D operator (see Figure 3.c).

The synthetic test pattern of Rosenthaler (Rosenthaler et al., 1992) simulates somewhat more realistic conditions, since it is smoothed and contains less contrast. As expected, the two procedures with the robust gradient computation show better results in comparison to the SUSAN-2D operator. However, the experiments carried out, that only the Förstner operator succeeded finding all potential interest points without errors.

This performance rating should be verified by a quantitative comparison, and the stability of the feature extraction should be measured. Therefore the front section of the Berlin-Palace was subjected to different transformations (see Figure 4):

- *Brightness and Contrast*: The radiometric characteristics were modified for the entire picture between 25 and 175 percent
- *Rotation*: Due to symmetry the rotations in the image plane were executed in steps between 0 and 90 degrees
- *Perspective*: Different points of view were simulated by a projective mapping of an image rotated around the horizontal axis between 0 and 60 degrees
- *Scaling*: The dimension of the image area was changed with a factor between 0.4 and 2.0
- *Noise*: The intensity values of the image were disturbed with a Gauss normal distribution up to 25 percent

We simulate the transformations on the computer using the indirect method with bi-cubic resampling. This may introduce interpolation noise and the camera model probably is unrealistic. On the other hand, the homography is exactly known and this approach is simple.

As a criterion for the evaluation, the *repeatability rate* was selected. On the basis of the well-known positions of the interest points in the test patterns, the number of correct features in the transformed images can be counted. The detailed results of this quantitative comparison are shown in Figure 5.

The investigations showed that e.g. the Förstner operator can cope with twice as strong radiometric changes as the Plessey point detector. Moreover, an increase in contrast leads immediate to losses with SUSAN-2D, while the result remains almost constant with the Förstner operator. All three test candidates show pleasingly stable results for rotations in the image plane. However, substantial differences can be observed during the perspective mapping or image scaling.

a) Relative intensity [%]    b) Relative contrast [%]

c) Rotation angle [deg.]    d) View angle [deg.]

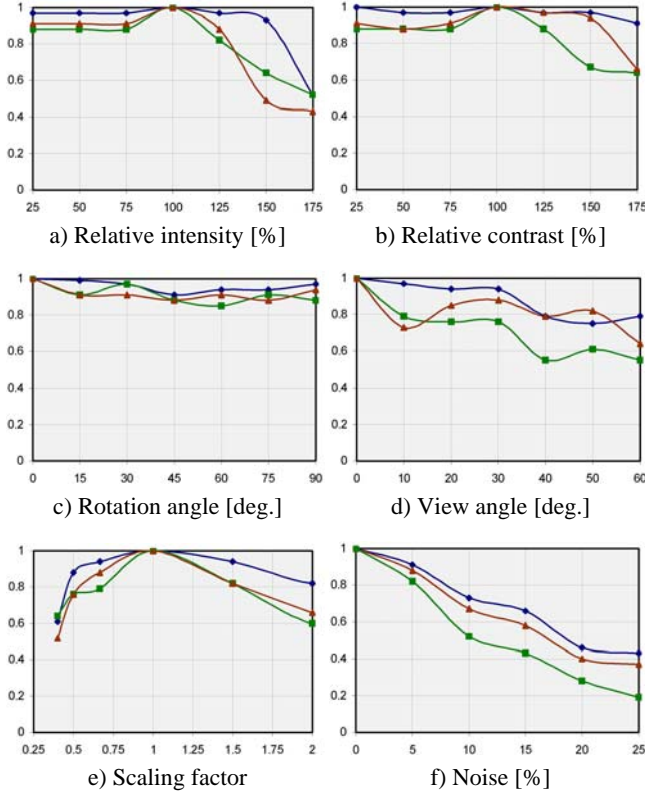e) Scaling factor    f) Noise [%]

Figure 5.    Repeatability rates for the transformed test patterns using the SUSAN-2D operator (■), the Plessey point detector (▲) and the Förstner operator (◆)

Even if the stability of all procedures decreases with image noise significantly, the Förstner operator is clearly superior to the SUSAN-2D operator. In summary the implementation of the Förstner operator shows the best results during the experiments (see Table 2).

## 4.    IMPROVEMENTS AND EXTENSIONS

Since the Förstner operator succeeded as the most efficient method under the requirements of image matching and object tracking, some further improvements are to be suggested in the following.

### 4.1  Sub-pixel Localization with Paraboloid Fitting

Since the accuracy for determining the image coordinates has a crucial influence on the later reconstruction, the localization should take place as precisely as possible.
Therefore the sub-pixel position for point like features is determined with a paraboloid model. On the basis of the integer position of a feature with maximum point weight $w$ its direct neighborhood is regarded and the nine values are normalized to 1 (see Figure 6.a). For modeling a paraboloid the bi-squared function

$$w(x, y) = ax^2 + by^2 + cxy + dx + ey + f \qquad (13)$$

can be used, whose surface is adapted on the normalized point weight.



a) Local maximum of the point weight    b) Adapted surface of the paraboloid
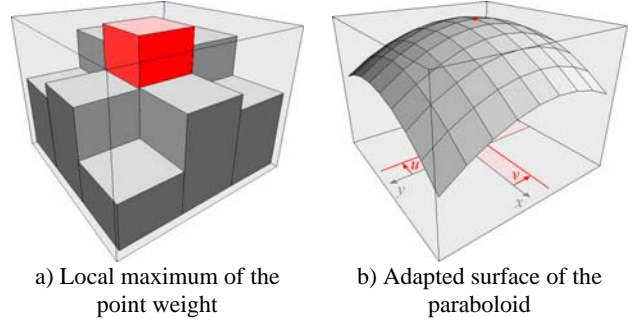
Figure 6.    Spatial representation of a paraboloid fitting to the normalized point weight

In order to determine the coefficients $\mathbf{x} = (a,\ldots,f)^{\mathrm{T}}$ of the paraboloid, a linear set of equations of the form

$$\mathbf{A}\,\mathbf{x} = \mathbf{b} \qquad (14)$$

can be formulated with

$$\mathbf{A} = \begin{bmatrix} x_1^2 & y_1^2 & x_1 y_1 & x_1 & y_1 & 1 \\ x_2^2 & y_2^2 & x_2 y_2 & x_2 & y_2 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_9^2 & y_9^2 & x_9 y_9 & x_9 & y_9 & 1 \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_9 \end{bmatrix}. \qquad (15)$$

If a point with maximum weight at position $(x, y)$ is defined as the origin of a local coordinate system (see Figure 7), the parameters from $\mathbf{A}$ are simplified to a constant design matrix:

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 & -1 & -1 & 1 \\ 0 & 1 & 0 & 0 & -1 & 1 \\ 1 & 1 & -1 & 1 & -1 & 1 \\ 1 & 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & -1 & -1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \qquad (16)$$

The use of the maximum point and its eight neighbors with only six unknowns in the coefficient vector $\mathbf{x}$ leads to an over-determined set of linear equations, which can be solved numerically e.g. with the help of the *Moore-Penrose Pseudo-inverse*

$$\mathbf{x} = \left(\mathbf{A}^{\mathrm{T}}\mathbf{A}\right)^{+} \mathbf{A}^{\mathrm{T}}\,\mathbf{b}\,. \qquad (17)$$

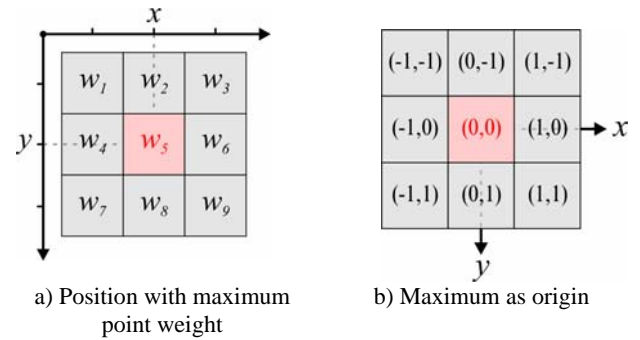After the coefficients of the paraboloid are computed, the position of the maximum can be found (see Figure 6.b).



a) Position with maximum point weight    b) Maximum as origin

Figure 7. Definition of the local coordinate system

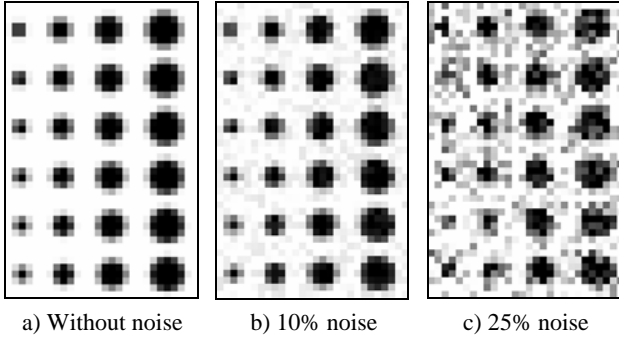a) Without noise    b) 10% noise    c) 25% noise

Figure 8.    Test pattern for the sub-pixel localization under different noise conditions

Therefore, the first partial derivatives

$$w_x = \frac{\partial w}{\partial x} = 2ax + cy + d \quad \text{and} \quad w_y = \frac{\partial w}{\partial y} = 2by + cx + e \quad (18)$$

must to be set to zero. The sub-pixel shift in $x$ and $y$-direction can be solved directly in closed form. After rewriting

$$u = \frac{(2bd - ce)}{(c^2 - 4ab)} \quad \text{and} \quad v = \frac{(2ae - cd)}{(c^2 - 4ab)} \quad (19)$$

applies. Finally the position for point like features can be improved by adding the sub-pixel shift to the integer position $(x + u, y + v)$. In order to be able to evaluate the accuracy of the sub-pixel localization, synthetic test patterns with points of a radius between two and five pixels in different positions were generated (see Figure 8).

By adding noise of a Gauss normal distribution with 10 and 25 percent, realistic conditions can be approximated. Table 1 contains the quantitative results of the sub-pixel localization.

| Noise [%] | Mean error [pixel] | | | | Maximum error [pixel] | | | | Resolution [pixel] | Reliability [pixel] |
|---|---|---|---|---|---|---|---|---|---|---|
| | 2 | 3 | 4 | 5 | 2 | 3 | 4 | 5 | | |
| 0 | 0.029 | 0.030 | 0.024 | 0.023 | 0.052 | 0.045 | 0.045 | 0.059 | **1/37** | **1/19** |
| 10 | 0.058 | 0.038 | 0.043 | 0.037 | 0.116 | 0.104 | 0.115 | 0.175 | **1/22** | **1/7** |
| 25 | 0.147 | 0.148 | 0.116 | 0.135 | 0.462 | 0.668 | 0.424 | 0.481 | **1/7** | **1/2** |

Table 1.    Localization error of the sub-pixel approach as a function of the point size and the noise level

It shows, that under ideal conditions, the position can be determined with an average accuracy of approximately 1/37 pixel and a maximum deviation of 1/19 pixel is not exceeded. In strongly disturbed images with a noise level of 25 percent the mean resolution reduces to 1/7 pixel and the reliability to a half pixel.
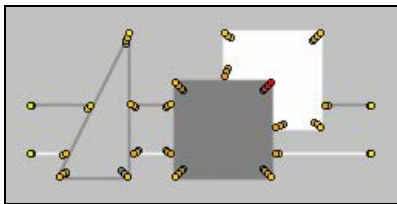


Figure 9.    Systematic localization error with increasing smoothing



a) Image pattern    b) GoG $\sigma = 0.5$    c) GoG $\sigma = 10.0$    d) LoG $\sigma = 10.0$
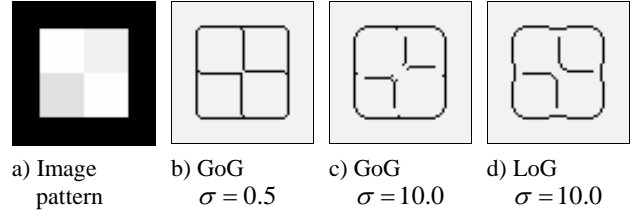
Figure 10. Localization problem by the intersection of edges

However, the use of a window with finite size for autocorrelation causes a systematic localization error for L- or T-corners (see Figure 9). The shift, depending on the standard deviation used for the Gaussian, is described in Wang (Wang & Brady, 1994). Förstner (Förstner & Gülch, 1987) as well as Deriche (Deriche & Giraudon, 1993) suggest a multi-level procedure to increase the accuracy of the determined position using edge intersections.

Nevertheless, classical edge operators that look for the maximum of the first derivative (Gradient of Gaussian, GoG) are less suitable for a correction, since they exhibit the same shifts. LoG operators (Laplacian of Gaussian) that search for zero-crossings of the second derivative make edge intersection difficult by strong rounding of the corners. This behavior is shown in Figure 10 using examples of the Canny edge operator (Canny, 1986) and the Marr and Hildreth operator (Marr & Hildreth, 1980) respectively.

## 4.2 Uniform Point Distribution with Adaptive Threshold

Due to the requirements in image orientation or camera tracking, not all extracted points are needed, so only the most important interest points must be selected.
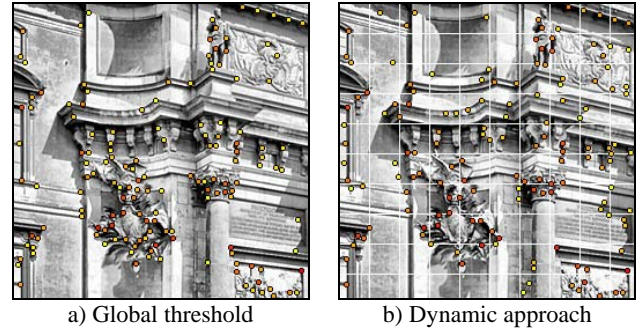


a) Global threshold    b) Dynamic approach

Figure 11. Optimization of the point distribution with an adaptive threshold

When using a global threshold for the entire image, unfavorable point accumulations may occur in structured areas while homogeneous ranges are neglected. An adaptive threshold procedure can be applied to optimize the point distribution. In order to affect the local point density, the entire image is first divided into equal sized segments.

In relation to the overall number of points each segment keeps proportionate a contingent, which is used for the strongest points within each segment. If a contingent remains during the allocation, or if segments contain fewer points than intended in the contingent, these are assigned afterwards to the global strongest points. The optimization of the point distribution with an adaptive threshold is shown in Figure 11. From approx. 1000 extracted point features the most important 150 candidates in 100 segments were selected.
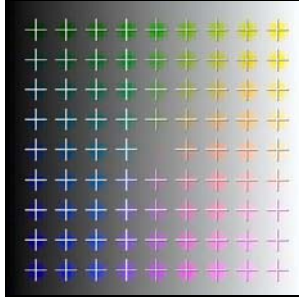
Figure 12. RGB color test pattern with extracted image points

### 4.3 Extension of the Autocorrelation for Color Images

A majority of the current sensor systems acquire color images. By an extension of the interest operator to color images, point features should be recognized by their color, even if they don't differ from the background in brightness. Montesinos (Montesinos et al., 1998) extended the autocorrelation matrix **A** for RGB color images $f = (r, g, b)^{\mathrm{T}}$, by computing the first partial derivative for each color channel. The determination of the autocorrelation matrix is similar to section 2.3 with:

$$\mathbf{A}_{RGB}(x, y) = \begin{bmatrix} \sum\left(r_x^2 + g_x^2 + b_x^2\right) & \sum\left(r_x r_y + g_x g_y + b_x b_y\right) \\ \sum\left(r_x r_y + g_x g_y + b_x b_y\right) & \sum\left(r_y^2 + g_y^2 + b_y^2\right) \end{bmatrix} \quad (20)$$

In order to demonstrate the functionality of the color extension, a synthetic test pattern with colored points was provided (see Figure 12). The color image has a continuous brightness gradient and the points differ only in their color or saturation from the background. Each cross marks a successfully detected color interest point.

## 5. EXPERIMENTAL RESULTS

After the performance of the implemented interest operator was illustrated using synthetic test patterns, this section presents results for real images. The interest points on the historical facade are based exclusively on intensity values. Results for color images are shown for the bust of the Nofretete (see Figure 13.c-d). The photographs were produced with a DV-Camcorder at a resolution of $1020 \times 1360$ pixels. Finally, from the popular Valbonne sequence (Schaffalitzky & Zisserman, 2001) two rotated test images with $768 \times 512$ pixels were selected. The overlapping area only constitutes a small portion of the entire image (see Figure 13.a and b).

| Detector | Detection rate | | Repeatability rate | | | | | | Locali-zation |
|---|---|---|---|---|---|---|---|---|---|
| | Smith | Rosen. | Inten. | Contr. | Rot. | View | Scale | Noise | |
| **SUSAN-2D** | 160* | 78 | 80 | 83 | 92 | 72 | 74 | 54 | 0.34 |
| **Plessey** | 24 | 86 | 79 | 90 | 92 | 82 | 74 | 65 | 0.29 |
| **Förstner** | 100 | 100 | 90 | 97 | 96 | 88 | 86 | 70 | 0.28 |

Table 2. Performance of the interest operators under the criteria of their detection and repeatability rate [in %] as well as the localization accuracy [in pixel] (* resulting from edge features).

## 6. CONCLUSIONS

Feature extraction using interest operators supplies stable point samples, which are suitable as candidates for image matching and object tracking. However, the quality of the interest points



a)                                                b)



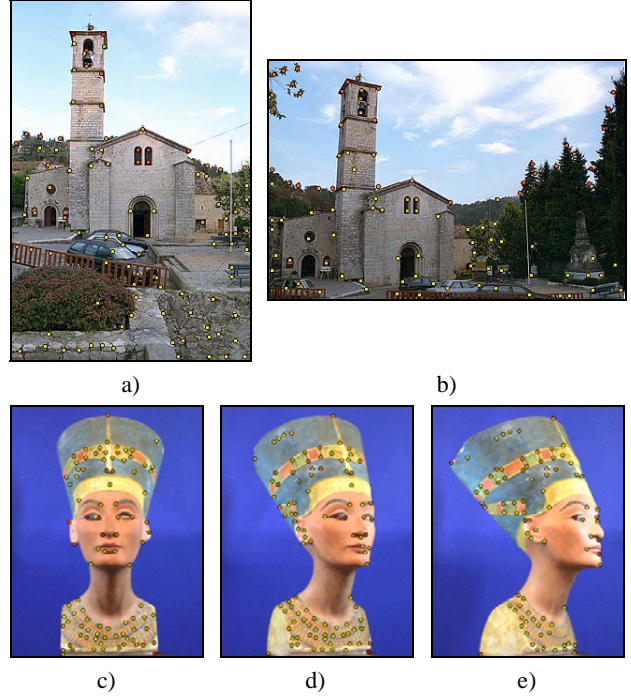c)                    d)                    e)

Figure 13. Extracted interest points for the test images of the Valbonne church, France, and the Nofretete bust

depends on the detector. In addition to the mathematical model, the implementation is of major importance for practical use. The implementation of the robust and well adjustable gradient computation with the continuous differential operators yields to far better results in practice as the use of discrete filter masks or gray value differences. The methodical comparison and the quantitative analysis of the most popular interest operators showed, contrary to Schmid (Schmid et al., 2000), that the implementation of the Förstner operator obtained the best results with regard to distinctness, invariance, stability, uniqueness, and interpretability.

Moreover, it was demonstrated that, for isolated points, the localization can be achieved with sub-pixel accuracy using a paraboloid fitting. In order to increase the stability of orientation procedures and camera tracking tasks, an adaptive threshold was suggested, which optimizes the point distribution in the image. Finally, to evaluate the additional information of color images an appropriate extension of the autocorrelation function for colored point features was described and tested.

However, the systematic localization error should be reduced for L and T-corners, since the shifts are different in the images due to the perspective distortions in close range. First attempts to eliminate the localization error with a hierarchical image pyramid show promising results, which must be examined in detail. Finally, a local statistic would be very helpful for the estimation of the control parameter to steer the automatic computation of interest points.

## 7. REFERENCES

Canny, J., 1986: A computational approach to edge detection, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 8, No. 6, pp. 679-698.

Deriche, R. and Giraudon, G., 1993: A computational approach for corner and vertex detection, *Int. Journal of Computer Vision*, Vol. 10, No. 2, pp. 101-124.

Deriche, R., 1990: Fast algorithms for low-level vision, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 12, pp. 78-87.

Förstner, W., 1994: A framework for low level feature extraction, In: Ecklundh (Eds.): *Proc. 3rd European Conf. on Computer Vision*, LNCS 800, Springer, pp. 383-394.

Förstner, W. and Gülch, E., 1987: A fast operator for detection and precise location of distinct points, corners and circular features, *Proc. Intercommission Conf. on Fast Processing of Photogrammetric Data*, pp. 281-305.

Hall, D. Leibe, B. and Schiele B., 2002: Saliency of Interest Points under Scale Changes, In P.L. Rosin, A.D. Marshall (Eds.): *Proc. of the British Machine Vision Conference*, pp. 646-655.

Haralick, R.M. and Shapiro, L.G., 1992: *Computer and Robot Vision*, Addison-Wesley, Reading, Massachusetts, USA, 630 p.

Harris, C. and Stephens, M., 1988: A combined corner and edge detector, *Proc. 4th Alvey Vision Conf.*, pp. 147-151.

Johansson, B. and Söderberg, R., 2004: A repeatability test for two orientation based interest point detectors, TR LiTH-ISY-R-2606, Dept. of Electrical Engineering, Linköping University, Sweden, 18 p.

Köthe, U., 2003: Integrated Edge and Junction Detection with the Boundary Tensor, *Proc. of 9th Int. Conf. on Computer Vision*, pp. 424-431.

Lindeberg T., 1998: Feature detection with automatic scale selection, *Int. Journal of Computer Vision*, Vol. 30, No. 2, pp. 79-116.

Lowe D.G., 2004: Distinctive image features from scale-invariant keypoints, *International Journal of Computer Vision*, Vol. 60, No. 2, pp. 91-110.

Marr, D. and Hildreth, E., 1980: Theory of edge detection, *Proc. Royal Society of London*, Vol. 207, pp.187-217.

Mikolajczyk, K. and Schmid, C., 2004: Scale & affine invariant interest point detectors, *Int. Journal of Computer Vision*, Vol. 60, No. 1, pp. 63-86.

Mikolajczyk, K. and Schmid, C., 2003: A performance evaluation of local descriptors, *IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 257-263.

Montesinos, P., Gouet, V. and Deriche, R., 1998: Differential invariants for color images, *Proc. 14th Int. Conf. on Pattern Recognition*, IEEE Computer Society Press, pp. 838-840.

Moravec, H.P., 1977: Toward automatic visual obstacle avoidance, *Proc. 5th Int. Joint Conf. Artificial Intelligence*, Cambridge, USA, p. 584.

Rodehorst, V., 2004: *Photogrammetrische 3D-Rekonstruktion im Nahbereich durch Auto-Kalibrierung mit projektiver Geometrie*, PhD. thesis, Wissenschaftlicher Verlag Berlin.

Rosenthaler, L., Heitger, F., Kübler, O. and v.d. Heydt, R., 1992: Detection of general edges and keypoints, In Sandini (Eds.): *Proc. 2nd European Conf. on Computer Vision*, LNCS 588, Springer, pp. 78-86.

Schaffalitzky, F. and Zisserman, A., 2001: Viewpoint invariant texture matching and wide baseline stereo, *Proc. 8th Int. Conf. on Computer Vision*, Vol. 2, IEEE Computer Society Press, pp. 636-643.

Schmid, C., Mohr, R. and Bauckhage, C., 2000: Evaluation of interest point detectors, *Int. Journal of Computer Vision*, Vol. 37, No. 2, pp. 151-172.

Smith, S.M. and Brady, J.M., 1997: SUSAN – a new approach to low level image processing, *Int. Journal of Computer Vision*, Vol. 23, No. 1, pp. 45-78.

Sojka, E., 2003: A new approach to detecting the corners in digital images, *Proc. Int. Conf. Image Processing*, Vol. III, pp. 445-448.

Tagare, H.D. and deFigueiredo, R.J.P., 1990: On the localization performance measure and optimal edge detection, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 12, No. 12, pp. 1186-1190.

Wang, H. and Brady, M. 1994: A practical solution to corner detection, *Proc. Int. Conf. on Image Processing*, Vol. 1, pp. 919-923.

Zuliani, M, Kenney, C. and Manjunath, B.S., 2004: A mathematical comparison of point detectors, *Conf. on Computer Vision and Pattern Recognition Workshop*, Volume 11, pp. 172-178.