HALF-SIFT: High-Accurate Localized Features for SIFT

Kai Cordes, Oliver Müller, Bodo Rosenhahn, Jörn Ostermann
Leibniz Universität Hannover
Institut für Informationsverarbeitung (TNT)
Appelstr. 9A, 30167 Hannover, Germany
{cordes,omueller,rosenhahn,ostermann}@tnt.uni-hannover.de

Abstract

In this paper, the accuracy of feature points in images detected by the Scale Invariant Feature Transform (SIFT) is analyzed. It is shown that there is a systematic error in the feature point localization. The systematic error is caused by the improper subpel and subscale estimation, an interpolation with a parabolic function. To avoid the systematic error, the detection of High-Accurate Localized Features (HALF) is proposed. We present two models for the localization of a feature point in the scale-space, a Gaussian and a Difference of Gaussians based model function. For evaluation, ground truth image data is synthesized to experimentally prove the systematic error of SIFT and to show that the error is eliminated using HALF. Experiments with natural image data show that the proposed methods increase the accuracy of the feature point positions by 13.9% using the Gaussian and by 15.6% using the Difference of Gaussians model.

1. Introduction

The precise detection of feature points in an image is a requirement for many applications in image processing and computer vision. In most cases, feature points are used to establish correpondences between different images containing the same scene captured by a moving camera or several cameras simultaneously. These corresponding points can be used to perform measurements to reconstruct the geometry of the observed scene or to recognize observed objects. In most approaches, the feature point detection and the correspondence analysis steps are divided into two parts independent from each other [5, 9], [1, 6, 8]. However, recent approaches show that the procedure can gain in performance if information of the feature point detection is used for the correspondence analysis [10].

Extensive work has been done on evaluating feature point detectors and descriptors [6, 10, 12, 15, 16, 17]. The criteria used for evaluation are *Information Content* [16], *Repeatability* [10, 16], *Overlap Error* [10], *Recall-Precision* [6, 11], and *Matching Score* [6, 11]. All these criteria deal with the performance of the approaches concerning distinctiveness of feature point descriptors. The evaluation is performed under various changing conditions resulting from a viewpoint change like illumination, rotation, perspective, and scale. For comparison, test data packages are provided. This data is mainly used to evaluate the distinctiveness of the feature descriptors for the correspondence analysis step. While the stability of the correspondence analysis is often analyzed in literature, the accuracies of feature points have rarely been considered.

The main motivation of this work is to analyze the localization accuracy of the SIFT detector. We present a proper evaluation scheme, and we introduce a modified feature point detector with increased accuracy based on our analysis. For estimating the subpel coordinate of a feature point, the standard SIFT detector uses a parabolic interpolation function, although it is known that this is a poor approximation for the gradient signal of an image taken with a camera based on a lens system [4, 14]. It has been shown, that using a parabolic interpolation leads to a systematic error in the localization of a feature point for the Canny- and Harris-detector [13, 14].

The scale of a feature is used to map the detected feature point position from the coordinates of the detected pyramid level to the coordinates of the original image. Hence, the accuracy of the localization is especially important at higher levels in the pyramid [2].

In this paper, the systematic error caused by the improper subpel interpolation technique is shown for the SIFT detector [2, 8] regarding the position and scale parameters of a feature point. To avoid this error, two improved interpolation techniques for the subpel estimation of the SIFT detector are proposed, a Gaussian and a Difference of Gaussians function model.

The contributions of this paper are

- the development of a method for evaluating the accu-

978-1-4244-3993-5/09/$25.00 ©2009 IEEE

racy of a detected feature point using synthetic images,

- experimental verification of a systematic error in the feature point localization of SIFT, and
- an accuracy gain of detected feature points of SIFT.

In the following Section, the SIFT detector is briefly presented. In Section 3, the proposed interpolation techniques are derived. Section 4 shows experimental results using synthetic and real image data. In Section 5, the paper is concluded.

2. The SIFT detector

The **S**cale **I**nvariant **F**eature **T**ransform provides distinctive feature points which can be used to establish correspondences between images with the same scene content. The corresponding points can be identified under various transformations of the image which are caused by a viewpoint change.

The workflow of SIFT using an image as input data is shown in Figure 1. First, feature points are detected that are characterized as extrema in the Difference of Gaussians pyramid of the image. The Difference of Gaussians pyramid is used as an approximation of the Laplacian pyramid, which has been proven to provide stable scale-invariant features. In the next step, the localization is refined by an interpolation of the 27 surrounding grid points with x-, y-, and scale - coordinates. This interpolation is done by fitting a 3D parabolic function to the Difference of Gaussians. It provides subpel and subscale accuracy of the localization of a feature point. In order to apply an orientation parameter to a feature, the main orientation of the surrounding image gradients is estimated. This information is used to ensure rotational invariance of a feature. Finally, a 128 dimensional vector is computed using the surrounding gradients. This vector is called SIFT descriptor. It is used to establish the correspondence to a feature in the other image. Correspondences between two images are found by associating feature points with a minimal distance between their descriptors.

3. Accurate Subpel Localization

In order to estimate the subpel and subscale parameters of a feature, a parabolic interpolation curve is used by the SIFT detector [2, 8] and variants of the SIFT detector [1, 6, 11]. The localization is determined by using an approximation of the Hessian matrix. To improve the interpolation accuracy, the parabolic function is replaced by a function that provides a better approximation of the image gradient signal. Assuming a camera lens system with a Gaussian transfer function and an ideal step edge, the gradient signal can be approximated by a Gaussian function

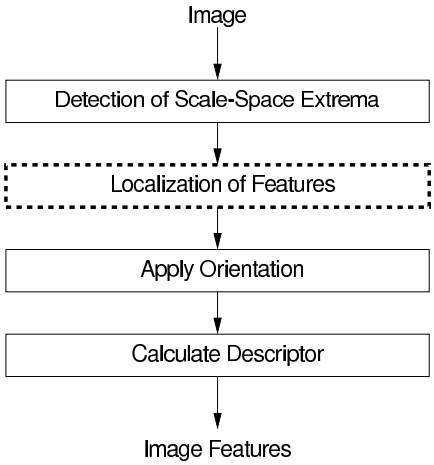

Figure 1. Workflow diagram of the detection of image features with the SIFT detector. The part modified in our approach is marked with a dotted box border.

[3, 4, 14]. The response of a peak distribution leads to a Difference of Gaussians [8].

The scale-space which is used to detect scale invariant features is represented by a Gaussian image pyramid [7]. In the SIFT approach, the first scale of each octave is smoothed with the prior smoothing $\sigma_0 = 1.6$ [8]. The standard deviation σ_i of a scale i can be calculated as

$$\sigma_i = \sigma_0 \cdot k^i = \sigma_0 \cdot (2^{\frac{1}{N}})^i \quad (1)$$

where N is the number of scales per octave and k determines the separation between neighboring scales [8]. Usually, the number of scales per octave is chosen to $N = 3$.

In order to extract an accurate subpel und subscale localization of a feature point, two interpolation functions are introduced, a Gaussian in Section 3.1 and a Difference of Gaussians in Section 3.2. In the following, a model for the interpolation functions is developed. A member of the model function is determined by a parameter vector $\mathbf{p}$. The parameter vector $\mathbf{p}$ is identified through a regression analysis minimizing the distances between the 27 sampling points surrounding a feature point in the Difference of Gaussians pyramid and the model function. For the optimization, the Levenberg-Marquardt (LM) algorithm is used. As pointed out in [14], the regression with LM may not provide stable results, a fact which we can confirm for some points. Further optimization analysis is left for future works.

3.1. Gaussian Interpolation

The interpolation function model using a Gaussian function with the covariance matrix $\Sigma = \begin{pmatrix} a^2 & b \\ b & c^2 \end{pmatrix}$ and its de-

terminant $|\Sigma| = det(\Sigma)$ can be written as

$$G_{\mathbf{x}_G,\Sigma}(\mathbf{x}) = \frac{r_G}{\sqrt{|\Sigma|}} \cdot e^{-\frac{1}{2}((\mathbf{x}-\mathbf{x}_G)^\top \Sigma^{-1}(\mathbf{x}-\mathbf{x}_G))} \quad (2)$$

Here, $\mathbf{x}_G = (x_0, y_0) \in \Re^2_{[-1;1]}$ is the subpel position of the feature point. The parameter vector $\mathbf{p_G} = (x_0, y_0, a, b, c, r_G)$ determines a member of the function model. The resulting solution for the subpel localization of the feature is (x_0, y_0). In order to extract the subscale localization i_0, the function model $G_{\mathbf{x}_G,\Sigma}(\mathbf{x})$ is also applied to the neighboring scales $i-1$ and $i+1$. Then, the maximum in scale direction i_0 is determined by fitting a Gaussian to the amplitudes of the scales $i-1$, i, and $i+1$.

3.2. Difference of Gaussians Interpolation

We assume a Gaussian image signal $G_{\mathbf{x}_D,\Sigma}$ surrounding a feature point as derived in Equation (2). Our model for the Difference of Gaussians function is based on the response of a Difference of Gaussians filter to a Gaussian input

$$\begin{aligned} D_{\mathbf{x}_D,\sigma} &= r_D(G_{\mathbf{x}_D,\Sigma_\sigma} - G_{\mathbf{x}_D,\Sigma_{k\sigma}}) * G_{\mathbf{x}_D,\Sigma} \\ &= r_D(G_{\mathbf{x}_D,\Sigma_\sigma+\Sigma} - G_{\mathbf{x}_D,\Sigma_{k\sigma}+\Sigma}) \end{aligned} \quad (3)$$

with $\Sigma_\sigma = \begin{pmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{pmatrix}$ and the standard deviation σ of the current scale.

Like the Gaussian function (Section 3.1), the Difference of Gaussians function can de described by a six-dimensional parameter vector $\mathbf{p_D} = (x_0, y_0, a, b, c, r_D)$.

In contrast to the Gaussian, the Difference of Gaussians function has more than one local extremum. This increases the probability that the gradient based Levenberg-Marquardt optimization strategy converges to a false local minimum of the cost function with invalid localization parameters (x_0, y_0, i_0). Hence, less feature points are detected with the Difference of Gaussians function model together with a gradient based optimization strategy.

4. Experimental Results

For the evaluation of our approaches, we use the following experimental setup.

- synthetically constructed Gaussian blob feature images (Section 4.1)
- natural image pairs [1] with estimated homographies (Section 4.2.1)
- natural image pairs (Section 4.2.2) using constant extrinsic camera parameters, but varying illumination conditions and varying intrinsic camera parameters

[1] www.robots.ox.ac.uk/~vgg/research/affine/index.html

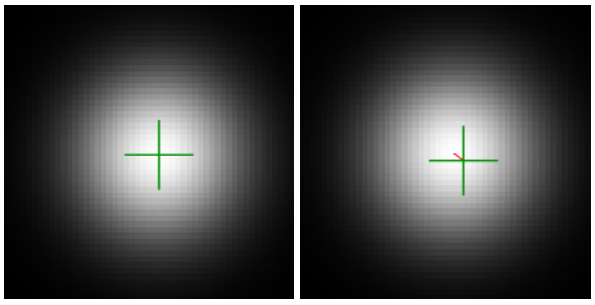

Figure 2. Ground truth feature localization (left) of the constructed Gaussian blob and the feature localization detected by the SIFT detector (right). The difference is indicated by a red line.

To evaluate the accuracy of the standard SIFT and our proposals, images with Gaussian blobs are synthesized. Each of the images provides one scale-space extremum with ground truth subpel und subscale values. As shown in Figure 2, the ground truth position is the center of the Gaussian blob. The scale parameter is determined by the standard deviation σ_g of the Gaussian. With these images, we can directly extract the subpel and subscale localization error of the approaches.

The test image sequences [1] provide image pairs with a camera performing a rotational movement, or observing a planar scene. Thus, all detected feature points in one image can be mapped to the corresponding feature points in the other image by a homography. Estimated [2] homography matrices are also provided with this data. These images are widely accepted in the computer vision comunity, especially for evaluating the distinctiveness of feature descriptors.

Our natural image pairs use constant extrinsic camera parameters. Therefore, the homography mapping between the images equals the identity matrix. The differences between images of a pair are generated by modifying the illumination conditions and varying intrinsic camera parameters.

In the following, we will refer to the three different subpel interpolation methods parabolic of the original SIFT, Gaussian (Section 3.1), and Difference of Gaussians (Section 3.2) with the abbreviations **PARAB**, **GAUSS**, and **DOG**.

For all experiments, default SIFT parameters are used, only the interpolation scheme has been exchanged.

4.1. Synthetic Images

The synthetic images are constructed using Gaussian blobs with a varying localization in x-direction and a varying standard deviation σ_g that determines the standard deviation of the scale $\sigma_i = \frac{1}{\sqrt{k}}\sigma_g$. The differences x_0 in x-direction are within the interval $[-2.0; 2.0]$ with a step

[2] www.robots.ox.ac.uk/~vgg/research/affine/det_eval_files/DataREADME

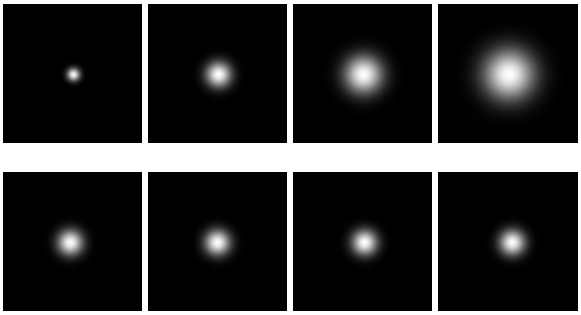

Figure 3. Some examples of the test images with varying standard deviation σ_i (top row, $\sigma_i = 2.0, 4.0, 6.0, 8.0$) and varying subpel position x_0 in x-direction (bottom row, $\sigma_i = 4.0$, $x_0 = -1.5, -0.5, 0.5, 1.5$)

distance of 0.05 pel. The used standard deviations σ_i are within the interval $[1.6; 9.9]$ with a step distance of 0.1 pel. The scale i_0 is determined with Equation (1). As we are using rotationally invariant Gaussians, we can limit the test scenario to the localization parameters x_0 and i_0. The subpel value in y-direction is $y_0 = 0$. With these values, the first two full pyramid octaves are covered. The image size is 64×64. Some examples are shown in Figure 3.

The results of **PARAB** can be seen in Figure 4 for the error of the estimated subpel and subscale localization $\tilde{x}_0$ and $\tilde{i}_0$. The Gaussian blobs are detected in two different pyramid octaves with $N = 3$ scales each. Within an octave, a systematic error $E^x = x_0 - \tilde{x}_0$ and $E^i = i_0 - \tilde{i}_0$ can be observed. With an error $E^x = 0$ on fullpel positions ($x_0 = 0$), its magnitude increases in both directions. The error E^x is periodic and represents the three scales of one octave. From one octave to the next, the structure of the error is similar, only its interval changes. The magnitude of the error increases regarding the x-position. Interestingly, the error E^i in scale decreases. The results of **DOG** can be seen in Figure 5. The position localization error is independent of the subpel and subscale position. The systematic error is eliminated exept for a small error of $E^i_{\mathbf{DOG}} < 0.0576$ pel in scale direction. The localization error of **GAUSS** is very similar to **DOG**. The maximum position errors match those obtained by **DOG**. The scale localization of **GAUSS** is $E^i_{\mathbf{GAUSS}} < 0.062$ pel.

The maximum errors $E^{max\text{-}x}_{\mathbf{PARAB}}$ of **PARAB** and $E^{max\text{-}x}_{\mathbf{DOG}}$ of **DOG** of the x-position for the first eight scales in three octaves are shown in Table 1.

4.2. Natural Image Pairs

For our tests with natural image pairs, we first detect features with all three methods, **PARAB**, **GAUSS**, and **DOG** in an image pair. Then, the default SIFT procedure for correspondence analysis is used to establish corresponding feature points. Mismatches in the correspondence analysis step

i	σ_i	x_0	$E^{max\text{-}x}_{\mathbf{PARAB}}$	$E^{max\text{-}x}_{\mathbf{DOG}}$
1..3	1.6..3.9	$-0.5..0.5$	0.0721	0.0062
4..6	4.0..7.9	$-1.0..1.0$	0.1424	0.0091
7..8	8.0..9.9	$-2.0..2.0$	0.2968	0.0213

Table 1. Maximum errors for the first three octaves of the Gaussian blobs. The maximum scale $i = 8$ is limited by the image boundaries of 64×64.

are eliminated by detecting all correspondences that do not fullfill the *Repeatability Criterion* [10]. This means that a feature point of the first image mapped by the given homography to the second image has a distance of $d < d_{thres}$ to the detected corresponding feature point in the second image. In [10], d_{thres} is set to 1.5 pel. In order to focus on the more accurate points, we decrease the threshold value of the *Repeatability Criterion* to $d_{thres} = 1.0$ pel. This guarantees, that misaligned feature points resulting from an imperfect optimization scheme (see Section 3) do not affect the results.

Feature points provided by the three methods **PARAB**, **GAUSS**, and **DOG** do not only differ in the subpel and subscale parameters. Because of the different localization, there are points rejected by the rejection criteria [8] from one method that are accepted from the other and vice versa. Furthermore, **DOG** tends to detect a smaller amount of features because of the occurence of local extrema as discussed in Section 3.2. To be able to compare the resulting feature points directly, we only use correspondences that are present in both methods.

For the comparison, the feature points are classified by error magnitude into 30 bins within the interval $[0.0; 1.0]$, and their occurence is counted. Thus, the more accurate feature points will appear on the left side of the resulting error histogram.

4.2.1 Image Pairs with Estimated Homographies

The test data provide image pairs with associated homography matrices that map all features of one image to the corresponding features in the other image. The homographies are not exact, but estimated. Since the homographies of the provided images result in a mean mapping error of $E > 0.4$ pel per feature, we only worked with data which provides a sufficiently accurate homography, namely the *ubc* sequence. This fact also indicates that high-accurate SIFT localization has not been in the focus of research so far. The *ubc* stereo pairs are well-suited since the homographies are exact ($\mathtt{H}_{ubc} = \mathtt{E}$). The error histograms for the first pair of *ubc* are shown in Figure 6. Comparing the histogram plots of **PARAB** with our approaches **GAUSS** and **DOG**, a shift to the left of the observed points is obvious. This means that the accuracy of the points increases. The mean error of a

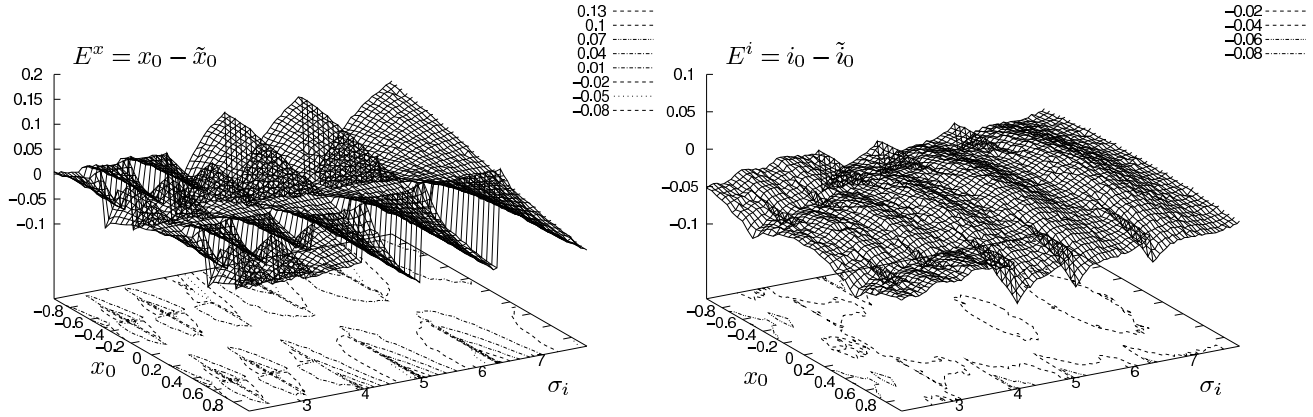

Figure 4. Error distribution of the standard parabolic interpolation **PARAB** of the x-position (left) and scale localization (right) showing the first two octaves. Contour lines are displayed in the ground plane.

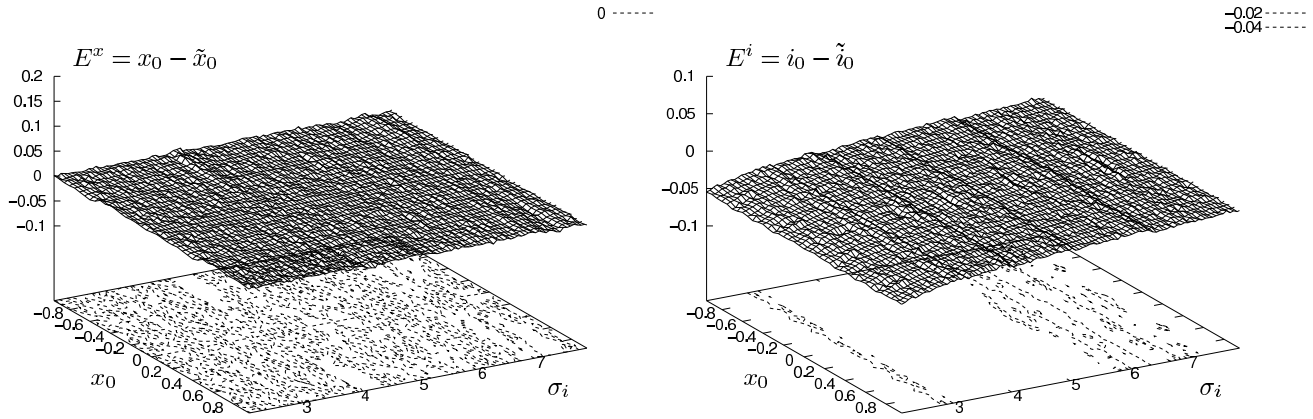

Figure 5. Error distribution of the proposed Difference of Gaussians interpolation **DOG** of the x-position (left) and scale localization (right) showing the first two octaves. Contour lines are displayed in the ground plane.

ubc Pair	Mean Error **PARAB**	Mean Error **GAUSS**	Mean Error **PARAB**	Mean Error **DOG**
1-2	0.1686	0.1452	0.1719	0.1485
1-3	0.2671	0.2471	0.2599	0.2445
1-4	0.3498	0.3262	0.3622	0.3434

Table 2. Results of the first three image pairs of the *ubc* sequence. For the comparison, only features are selected which are detected with both methods. Thus, results for **PARAB** slightly vary within a row.

feature is shown in Table 2. It decreases up to 13.9% using **GAUSS** and up to 13.6% using **DOG** instead of the original **PARAB** as shown in Table 2.

4.2.2 Image Pairs with Illumination Changes

For our test scenario, we chose a Thomson Viper HD TV camera with a resolution of 1920×1080 pixels. A drawing was captured from a distance of 1.50 m in our studio. To get differing images with constant extrinsic camera parameters, the illumination intensity is changed. By varying the illumation, three test scenarios are constructed:

A) all intrinsic camera parameters remain constant which means that the second captured image has a lower dynamic range as shown in Figure 7.

B) the *autoiris*-function of the camera is activated. Thus, the camera compensates the less intense illumination with its aperture. The changes are shown in a part of the captured images in Figure 8.

C) the *automatic gain*-function of the camera is activated which means that the changing illumination circumstances are compensated with the amplifier of the camera. This results in noise added to the image as shown in Figure 9.

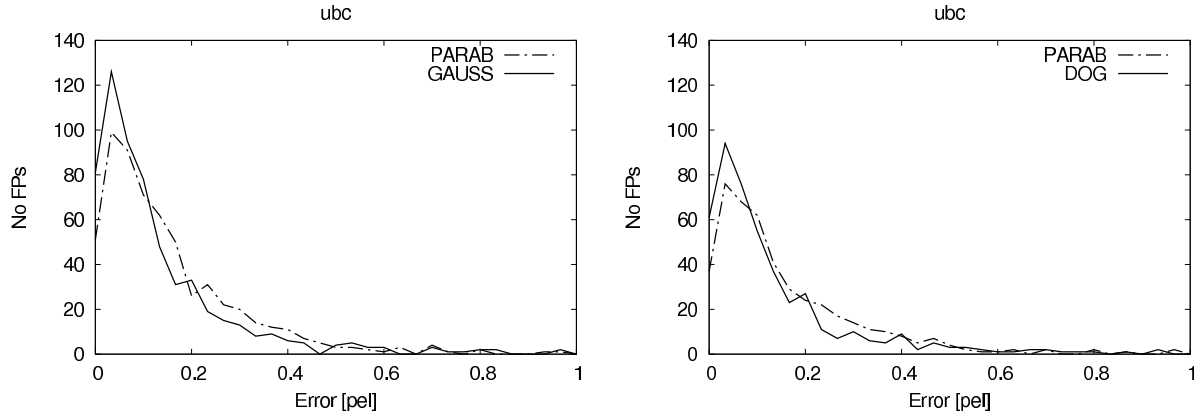

Figure 6. Comparisons of **PARAB** to **GAUSS** and **PARAB** to **DOG** classifying the position error of a feature point by its magnitude for the first image pair of *ubc*. The more accurate feature points appear on the left side of the error histogram.

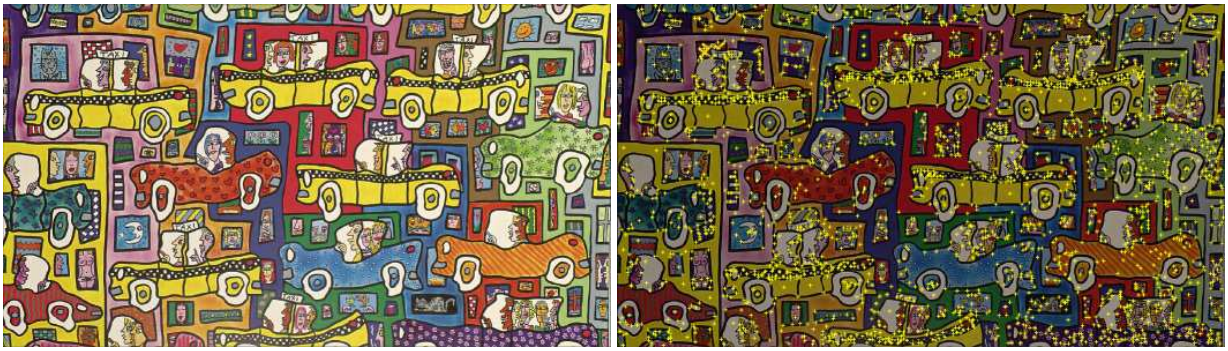

Figure 7. Our natural test image pair from the first test scenario $A^{50\%}$. The illumination intensity changes from 100% (left) to 50% (right). In the right image, the detected feature points are marked in yellow.

Figure 8. Part of the test image pair $B^{50\%}$. The *autoiris*-function of the camera compensates the illumination intensity change from 100% to 50%

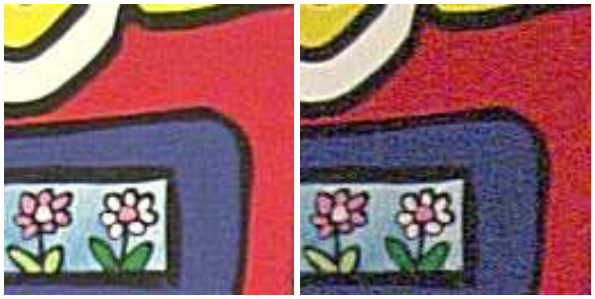

Figure 9. Part of the test image pair $C^{50\%}$. The *automatic gain*-function of the camera compensates the illumination intensity change from 100% to 50%

With each of the scenarios A, B, and C, two image pairs are generated, the first with an illumination intensity change from 100% to 75% and the second pair with a change from 100% to 50%. The image pairs are called $A^{75\%}$, $A^{50\%}$, $B^{75\%}$, $B^{50\%}$, $C^{75\%}$, and $C^{50\%}$. The image pair $A^{50\%}$ with an illumination intensity change from 100% to 50% is shown in Figure 7. The effect of the second scenario $B^{75\%}$ and $B^{50\%}$ is a decreasing depth of field (shown in an image part in Figure 8) while the effect of the third scenario $C^{75\%}$ and $C^{50\%}$ is an increase in noise (shown in an image part in Figure 9).

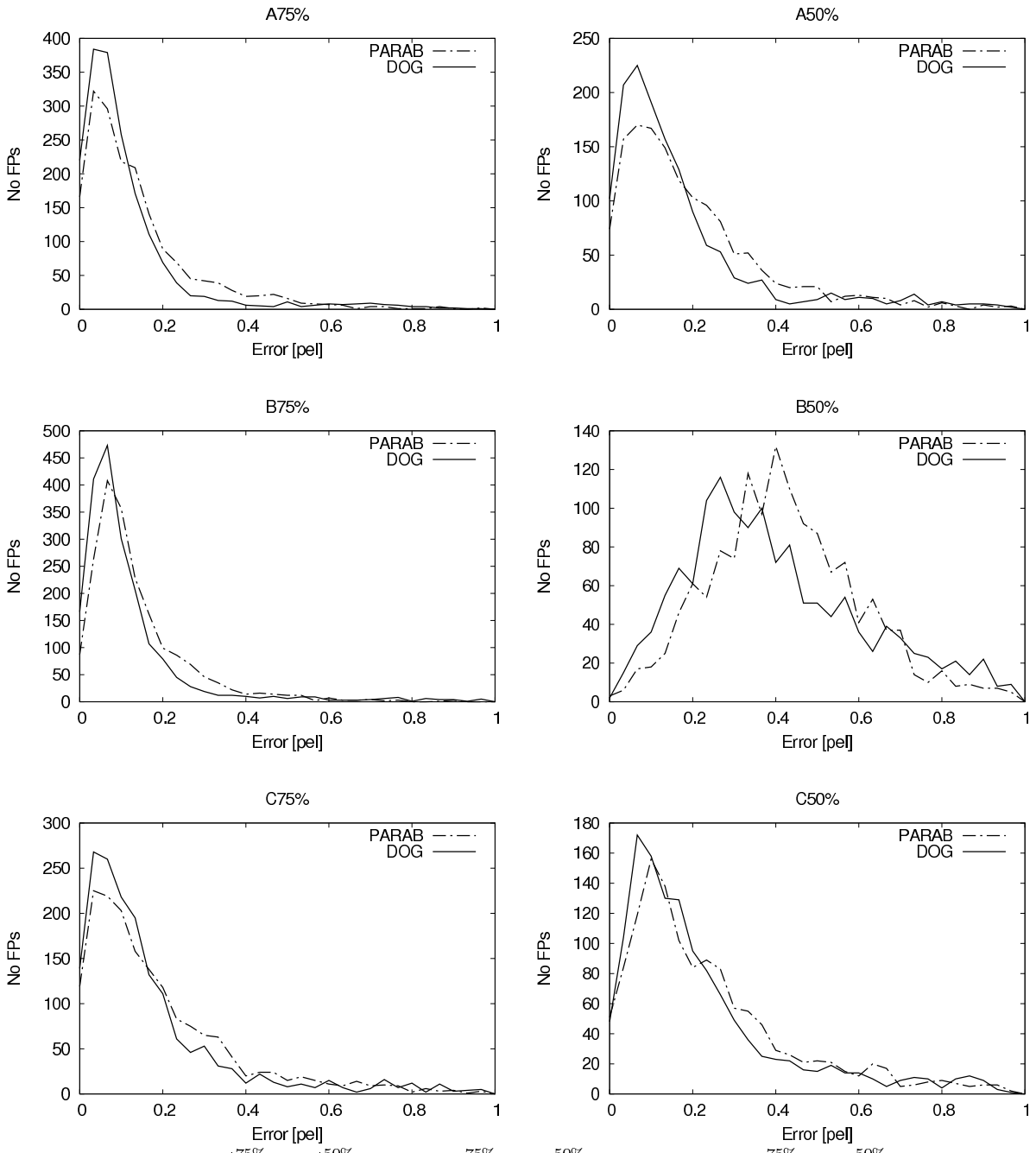

Figure 10. Error histograms for $A^{75\%}$ and $A^{50\%}$ (top row), $B^{75\%}$ and $B^{50\%}$ (middle row) , and $C^{75\%}$ and $C^{50\%}$ (bottom row) comparing **PARAB** with our approach **DOG**

The resulting error histograms are shown in Figure 10 for $A^{75\%}$ and $A^{50\%}$ (top row), $B^{75\%}$ and $B^{50\%}$ (middle row), and $C^{75\%}$ and $C^{50\%}$ (bottom row). In the histograms, all feature points occuring for both methods **PARAB** and **DOG** are classified by their error magnitude into bins within the interval $[0.0, 1.0]$. For **DOG**, there are more features having a smaller error than for **PARAB** in all image pairs. This result is confirmed by Table 3, where the mean errors of the feature points are shown for the comparisons between **PARAB** and **DOG** and for the comparisons between **PARAB** and **GAUSS**. Both methods perform better than the original interpolation scheme of SIFT. The accuracy incrases by 11.0% in maximum for **GAUSS** and by 15.6% in maximum using **DOG**.

scene	Mean Error PARAB	Mean Error GAUSS	Mean Error PARAB	Mean Error DOG
$A^{75\%}$	0.1505	0.1345	0.1573	0.1327
$A^{50\%}$	0.2076	0.1848	0.2127	0.1852
$B^{75\%}$	0.1548	0.1430	0.1549	0.1343
$B^{50\%}$	0.4682	0.4624	0.4393	0.4151
$C^{75\%}$	0.2025	0.1812	0.2042	0.1833
$C^{50\%}$	0.2634	0.2422	0.2570	0.2375

Table 3. Results of the natural test image pairs diplaying the mean error of a correspondence. For the comparison, only features are selected that are detected with both methods. Thus, the results for **PARAB** slightly vary within a row.

5. Conclusions

Using the standard SIFT feature point localization, by fitting a 3D quadratic function to the Difference of Gaussians pyramid leads to a systematic error in the localization of a feature. This is shown in our work by evaluating synthetic images with varying localized extrema in the scale-space. The systematic error can be avoided by using a function model which is adapted to the image signal characteristics. Using a Difference of Gaussians function model eliminates the systematic error.

Experiments with natural images reveal that the position localization error can be reduced by up to 15.6% using the Difference of Gaussians approach. However, the approach tends to detect less feature points because of multiple local extrema. Using a Gaussian function model decreases the localization error up to 13.9%. These results are obtained for accurate feature points ($\leq$ 1 pel). The evaluation of less accurate points is left for future works.

Our approach is designed for the SIFT detector, but can be applied to all feature point detectors which use scale-space extrema as characterization of feature points.

References

[1] H. Bay, T. Tuytelaars, and L. V. Gool. Surf: Speeded up robust features. In *European Conference on Computer Vision (ECCV)*, volume 3951, pages 404–417. Springer, 2006.

[2] M. Brown and D. G. Lowe. Invariant features from interest point groups. In *British Machine Vision Conference (BMVC)*, pages 656–665, 2002.

[3] J. Canny. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 8(6):679–698, 1986.

[4] J. W. Goodman. *Introduction to Fourier Optics*. McGraw-Hill, 1968.

[5] C. Harris and M. Stephens. A combined corner and edge detector. In *4th Alvey Vision Conference*, pages 147–151, 1988.

[6] Y. Ke and R. Sukthankar. Pca-sift: A more distinctive representation for local image descriptors. In *International Conference on Computer Vision and Pattern Recognition (ICCV)*, pages 506–513, 2004.

[7] T. Lindeberg. Feature detection with automatic scale selection. *International Journal of Computer Vision (IJCV)*, 30:79–116, 1998.

[8] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision (IJCV)*, 60(2):91–110, 2004.

[9] B. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In *7th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 674–679, 1981.

[10] K. Mikolajczyk and C. Schmid. Scale & affine invariant interest point detectors. *International Journal of Computer Vision (IJCV)*, 60(1):63–86, 2004.

[11] K. Mikolajczyk and C. Schmid. A performance evaluation of local descriptors. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 27(10):1615–1630, 2005.

[12] K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Schaffalitzky, T. Kadir, and L. V. Gool. A comparison of affine region detectors. *International Journal of Computer Vision (IJCV)*, 65(1-2):43–72, 2005.

[13] P. Mikulastik, H. Broszio, T. Thormählen, and O. Urfalıoğlu. Error analysis of feature based disparity estimation. In *Advances in Image and Video Technology (PSIVT)*, volume 4319, pages 1–12. Springer, 2006.

[14] P. Rockett. The accuracy of sub-pixel localisation in the canny edge detector. In *British Machine Vision Conference (BMVC)*, 1999.

[15] C. Schmid, R. Mohr, and C. Bauckhage. Comparing and evaluating interest points. *IEEE International Conference on Computer Vision (ICCV)*, pages 230–235, 1998.

[16] C. Schmid, R. Mohr, and C. Bauckhage. Evaluation of interest point detectors. *International Journal of Computer Vision (IJCV)*, 37(2):151–172, 2000.

[17] T. Tuytelaars and K. Mikolajczyk. *Local invariant feature detectors: a survey*, volume 3. Foundations and Trends in Computer Graphics and Vision.