Hendrik Schilling*, Maximilian Diebold, Marcel Gutsche, and Bernd Jähne

On the design of a fractal calibration pattern for improved camera calibration

Über den Entwurf eines fraktalen Kalibriermusters für eine verbesserte Kamerakalibrierung

Abstract: Camera calibration, crucial for computer vision tasks, often relies on planar calibration targets to calibrate the camera parameters. This work describes the design of a planar, fractal, self-identifying calibration pattern, which provides a high density of calibration points for a large range of magnification factors. An evaluation on ground truth data shows that the target provides very high accuracy over a wide range of conditions.

Keywords: Camera calibration, high accuracy, passive target, image geometry, image measurement, fiducial marker, self-identifying, image localization.

Zusammenfassung: Eine gute Kamerakalibrierung ist für Computer Vision Aufgaben unentbehrlich und setzt oft auf planare Kalibriermuster, um die Kameraparameter zu ermitteln. Diese Arbeit beschreibt die Entwicklung eines planaren, fraktalen selbst-identifizierenden Kalibriermusters, welches eine hohe Dichte von Kalibrierpunkten für einen weiten Bereich von Abbildungsmaßstäben bereitstellt. Eine Evaluierung auf Ground-truth-Daten demonstriert die hohe Genauigkeit, die sich über einen weiten Parameterbereich erzielen lässt.

Schlüsselwörter: Kamerakalibrierung, hohe Genauigkeit, passives Target, Bildgeometrie, Bildmessung, Referenzpunkte, selbstidentifizierend, Bildlokalisierung.

*Corresponding author: Hendrik Schilling, Heidelberg University, Heidelberg Collaboratory for Image Processing (HCI), Berliner Straße 43, 69120 Heidelberg, Germany, e-mail: Hendrik.Siedelmann@iwr.uni-heidelberg.de
Maximilian Diebold, Marcel Gutsche, Bernd Jähne: Heidelberg University, Heidelberg Collaboratory for Image Processing (HCI), Berliner Straße 43, 69120 Heidelberg, Germany

1 Introduction

Common methods for calibrating camera systems utilize passive targets with checkerboard patterns, which can be fabricated by printing. Such targets are easy to detect by a large range of available software, like the OpenCV library [1].

In computer vision tasks it is already possible to locate image features with an accuracy down to a few hundreds of a pixel, for example in light-field measurements [2]. However, the camera calibration, critical in describing the relation between the world and the observed features, commonly achieves an accuracy not much below a tenth of a pixel, also visible in our evaluation, see Figure 13. Thus a lot of the accuracy cannot be exploited due to the lack of a suitable calibration.

Regarding the calibration target we can find several reasons for the limited accuracy. If a pattern is used without identifying marks, it is impossible to determine pattern coordinates, if parts of the pattern are outside the field of view of the camera. This forces the user to place the calibration pattern away from the borders and therefore induces a bias towards the center of the imaging area. To avoid this it is necessary to use a self-identifying marker pattern [3, 4].

The size of the calibration markers also plays a key role in the accuracy of the localization. If too few calibration points are available per view, then calibration quality suffers due to overfitting. However, if the calibration markers become too small, it becomes impossible to properly detect them. As the optimal size depends on the camera magnification, it is not possible to create a pattern with a single optimal density.

This work introduces a fractal calibration target, published under an open source license [5], which addresses this problem by providing calibration markers at several scales, allowing the adaption to the camera magnification at the time of detection. This is a first step towards a more sophisticated camera calibration, lifting accuracy to the next level suitable for highly accurate measurements.

2 Related work

The following overview is limited to the subject of calibration point localization. For passive targets there are two main types of localization features, checkerboard corners and circular markers.

For the checkerboard localization three main principles have been established. The first method uses line intersection, where lines are refined individually and the intersection of two lines defines the calibration points [3]. The second approach uses functional descriptions of saddle points [6–8], where the local neighborhood is approximated by a 2D polynomial in which the corner point can be derived analytically. In comparison to line intersection, the saddle point method is more suited to smaller neighborhoods and at higher resolutions achieves either no improvements [8] or even worse results [7].

The third method makes use of orthogonal gradients, as implemented for example by OpenCV [1]. It exploits the property that the vector from the desired corner point to any edge pixel is orthogonal to the gradient vector in this pixel. This provides better results compared to line intersection as shown by Atcheson et al. [3]. Note that line intersection also suffers from distortion bias [7].

For circular markers the modeling of the transformation from a perfect circle to the observed image plays the largest role. Mallon et al. [7] evaluate localization performance on real and simulated images, under blur, noise, perspective and distortion. They conclude that circular patterns are dominated by perspective and distortion bias and recommend, for circular patterns, the usage of small (10px) circles and the correction according to the distortion model. Consequently Datta et al. [9] as well as Douxchamps and Chihara [10] use an iterative refinement process which alternates between updating the camera model and refinement of the calibration marker image location. While this gives high quality results this approach inherently links the camera calibration with the marker refinement which makes the process slow and inflexible, when compared to a two step approach where marker refinement and camera calibration are separate.

A possible alternative to these passive methods are active targets, for example using fringe projection or active phase targets. However those come with their own constraints, like requiring an extra device with control and synchronization. Therefore this work focuses on passive targets which can easily be fabricated by printing.

3 Overview

The target introduced in this work relies on a self-identifying fiducial marker pattern, combined with a fractal structure, where a large number of calibration points are recursively inserted at several scales. This keeps the number of calibration points within a fixed bound, independent of the magnification factor. The reasoning and design of the fractal structure is discussed in Section 4.

Markers are identified using a 3×3 binary pattern, see Figure 1, where a black bordered and a white bordered marker are always processed together, which effectively results in 18 bits of payload which are used to encode the 2D marker position on a regular grid.

Detection starts with the identification of the large scale markers, see Section 5.1, combining checkerboard corner detection and the payload for identification, as well as error detection by correlating neighboring marker candidates. In a second step the recursive dots are detected using a recursive strategy which fits 2D Gaussian distributions onto dot candidates, compare Figure 2, whose positions are estimated from the previous scale. The recursive refinement scheme is described in Sections 4.3 and 5.2.
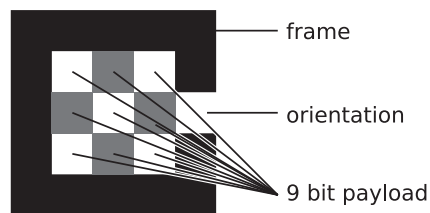

Figure 1: Schematic of a single marker. The constant color border with the hole at one side allows detection and fixes the orientation. A single marker contains a payload of 9 bits which is used for identification.
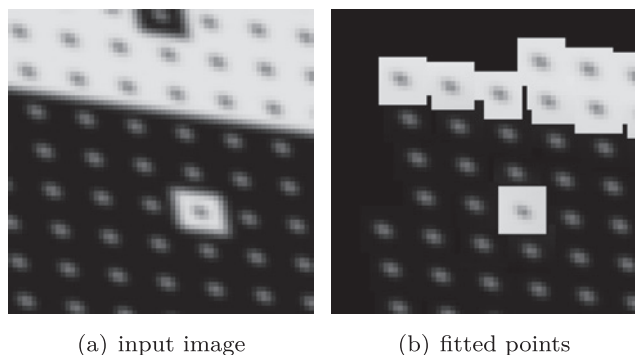

(a) input image   (b) fitted points

Figure 2: An example of the detection performance with small calibration points at the image border. On the left the original input image, on the right the fitted calibration points. Note how the small calibration points can be detected less than 10 pixels from the image border.
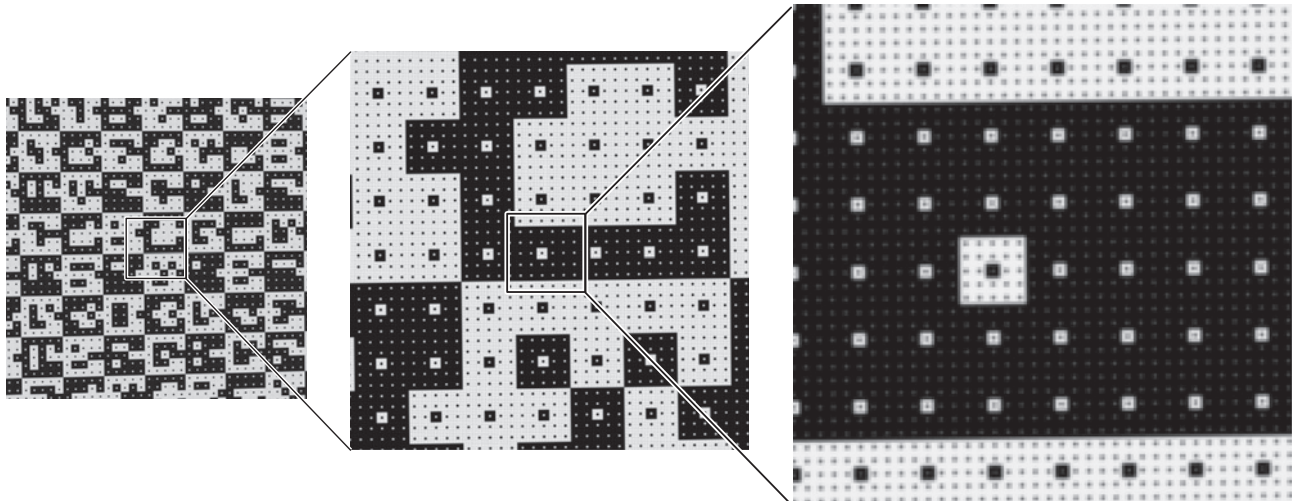
Figure 3: A view of the fractal calibration target. From left to right each step reduces the width of the cut-out to a fifth. Note how individual calibration dots resolve to square features when increasing the magnification.

As is visible in Figure 3, the recursion scheme operates with powers of five. The minimum size for successful fitting is around four pixels, which means that for high quality images we get one calibration point for every 4–20 pixels in each dimension, or between 2500 and 62 500 calibration points for each megapixel of image data.

While the individual calibration points are not necessarily detected with a higher accuracy compared to other calibration targets, the high number of calibration points leads to better calibration results as errors are averaged out. Compared to the OpenCV subpixel refinement we can increase the accuracy of the camera calibration by one order of magnitude, when comparing the calibrated projection against the ground truth projection, as shown in Section 4.2.

4 Design

Regarding the marker size, there are two possible directions. Either few, large, individually highly accurate markers, as used by Douxchamps and Chihara [10], or many small markers of individually low accuracy.

4.1 Why marker size does not matter

We argue that the underlying measurement accuracy for large and small markers is similar, as it depends solely on the quantity and magnitude of image gradients. A large marker could be regarded as a collection of smaller ones, which implies that they should achieve similar accuracy.

Indeed the subpixel marker locations are constrained by the individual image gradients. If the same quantity of gradients is used, where each gradient has the same quality (signal to noise ratio), then it does not really matter whether those gradients are spread over few, large markers or several smaller ones. However the size of the individual markers still has practical implications on the way detection can be performed and where markers can be placed. Note that this underlying accuracy is not reflected by the RMS scores of calibration methods, as those only give the error between the fitted model and the detected markers. Larger markers have already averaged out more of the gradient noise due to the larger marker sizes, where small markers provide a higher ratio of data samples to parameters in the overconstrained minimization problem of the calibration method, which will also lead to a reduction of the influence of gradient noise.

4.2 Why marker size matters

The localization of larger markers is dominated by perspective and radial distortion [7], which require a more complex localization approach for accurate results. While perspective bias is easy to compensate for, even without doing a full calibration, to correct distortion bias the marker localization has to incorporate the radial distortion [7, 9, 10], which ties marker refinement directly into the calibration. Thus it is impossible to separate marker detection and camera calibration, which also makes it difficult to adapt new calibration methods to such a combined calibration scheme. This means that small features

are better suited for a generic calibration pattern, as the individual calibration features are not affected much by bias.

In Section 4.1 we argue that the underlying accuracy for larger and small markers is basically the same. However, there are two ways in which smaller markers can provide accuracy improvements over large markers. A single marker has to be fully visible to provide localization. But the larger the size of an individual marker, the larger the area that is missing, if a part of the marker is obstructed by the image border. Hence the contribution of edges and corners is reduced for larger markers, which is problematic for camera calibration, as those areas often exhibit the strongest distortion. The result is, that larger markers show a more pronounced center bias, where the quality of the calibration is better in the center of the imaging area than in the corners, while small markers can be detected close to the image borders, see Figure 2, which reduces center bias.

Lastly, as the number of calibration points decreases due to increased marker size, the probability of overfitting increases due to reduced ratio of calibration points to calibration parameters.

4.3 Fractal layout

For the reasons exposed in Section 4, we try to minimize the calibration point size. Of course there is a minimal size for calibration points, before detection becomes impossible, because individual calibration points blend together. However, a pattern which is optimal in one given situation can become completely useless if the magnification is decreased. As the effective scale at which target features are projected onto the image sensor changes with distance, angle and radial distortion, it would therefore be necessary to capture a range of targets, each optimal at a different scale, and later fuse the calibration information from multiple targets.

Our solution to this tedious multi-target process is the adoption of a fractal scheme, which operates with multiple scales of calibration points. By using a target which includes calibration points at multiple scales, it is possible to get optimal calibration point density at several magnification ratios. Figure 3 shows the calibration pattern at several magnifications.

Detection is initially performed only for the coarsest layer, then runs on the next higher resolution layer, and so on. When some calibration points cannot be matched at one scale, the distances between fitted calibration points are examined to decide whether detection of the next finer scale should be attempted. While this does not guarantee

an, in practice hard to define, optimal density of calibration points, it keeps the density within a fixed bound, depending on the effective image resolution and the deterioration by external sources, like noise, blur and other aberrations.

The used calibration points are individual square dots. Due to the fractal nature of the pattern the finest resolvable calibration points are always at the pixel scale where several independent degradations, like physical pixel aperture, lens aberrations, diffraction, etc., render a calibration point as a Gaussian-like 2D distribution. When a calibration point becomes large enough that the non-Gaussian characteristics come to bear, it is already possible to resolve the next recursion layer, making more complex refinement methods unnecessary. In addition, this elegantly avoids the bias problems of large markers, as the fractal nature of the calibration pattern always allows the detection of relatively small markers, with a size close to the resolution limit.

A new layer of calibration points is generated from a coarser layer by resizing the pattern with a factor of 5, using nearest neighbor interpolation. In the scaled pattern, new calibration points are inserted at every fifth pixel by inverting the respective pixel, see Figure 3 for a visualization.

5 Implementation

The detection of the fractal pattern is split into two parts, the actual detection using the payload for identification, and the fractal refinement. Detection takes place in scale space, using an x-corner detector followed by brute force iteration of possible markers using neighboring corner candidates. For each set of four corners, the containing image area is projected onto a marker template using a perspective transform, to assess whether the corner selection represents a marker candidate.

5.1 Identification

It is important that parts of the pattern are allowed to be out of view, to give the freedom to cover the whole imaging area with the calibration target, and place calibration points close to edges and corners. Our method builds on square blocks with uniform borders and a payload of 9 bit in the center, see Figure 1. The orientation is fixed using an opening in the border and markers are arranged in a checkerboard manner, compare Figures 3 and 5. This means that borders are either black or white. The coding

scheme combines the payload of black and white markers to provide 18 bits of payload for addressing and always requires the detection of at least two neighboring black and white markers. The address is encoded using an XOR mask to provide a bit of randomization and to keep the overall distribution of bits more uniform. In contrast to other fiducial marker systems, like for example CALTag [3], the payload within the markers does not provide error detection or correction. Instead the address of neighboring markers is compared to check whether they are consistent, which provides very robust error detection but no error correc-

(a) step 0 guess   (b) step 0 fits

(c) step 1 guess   (d) step 1 fits

(e) step 2 guess   (f) step 2 fits

(g) diff between (f) and (h)   (h) input image

Figure 4: Visualization of the fractal refinement scheme. From top to bottom, the refinement from coarse to fine is shown, with candidates on the left and the actual fits on the right. As edges are ignored they are simply estimated from the marker size for visualization purposes, while the 2D distributions (right side) are a correct visualization of the estimated fit. Calibration point positions at layer n are estimated by a perspective transform from known calibration points of layer $n - 1$. These estimation candidates are shown as circles in (a, c, e). Then, individual 2D distributions are fitted to these estimates. The resulting fits are visualized in (b, d, f). This process is repeated until calibration point size is too small for a successful estimate. For reference, (h) shows the input image and (g) shows the difference between the final fit and the input, multiplied by 10. Not that all calibration points which were successfully detected in (f) show no discernible error in (g) (black areas).

tion. This means the marker grid can have a maximum size of 512×512 markers which results in 262144 markers overall. For more details on the identification scheme please see [11].

5.2 Fractal refinement

The fractal refinement starts after markers have successfully been detected. Expected positions of the first layer of calibration points are estimated from their respective marker using a perspective transform, compare Figure 4(a, c, e). Those positions are then refined using a least squares solver which fits rotated 2D anisotropic Gaussian distributions with a linear gradient as background, compare Figure 4. Individual pixels are used as samples in the fit, with a weighting relative to the initially estimated position. This step is repeated once more to improve accuracy by using the updated center position.

The recursion works by repeating this procedure, see Figure 4, until the size of the calibration points becomes too small for a meaningful fit, which is less than 4 pixels across. Various heuristics are used to verify the validity of the fit, like rejecting calibration points for which either the residual of the fit is too large, or for which the expected accuracy is low due to low contrast, steep background gradient, too small or too large width of the Gaussian, or due to saturation.

5.3 Bayer pattern mode

Because individual pixels are directly used as data points in the fit, it is possible to leave out individual samples. For color images acquired using a color filter array, demosaicing can be avoided by processing colors independently of each other, simply leaving out the pixels which belong to a different color channel.

6 Evaluation

The evaluation is based on rendered images corrupted with Gaussian blur, additive Gaussian noise, uneven illumination, reduced contrast and radial distortion. We compare classic checkerboard detection with subpixel refinement using the implementation from OpenCV, with our fractal calibration pattern. The detected calibration points are then used for either a full camera calibration, again using the implementation from OpenCV, or alternatively to
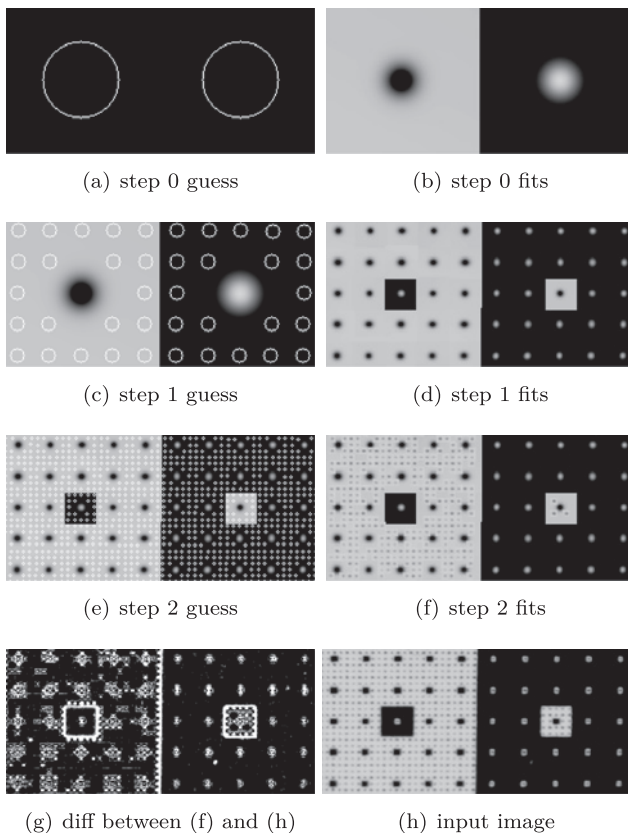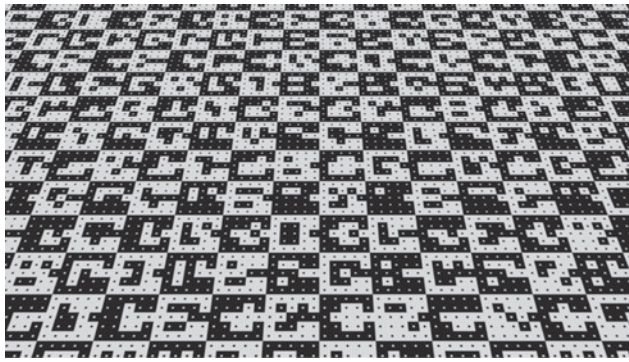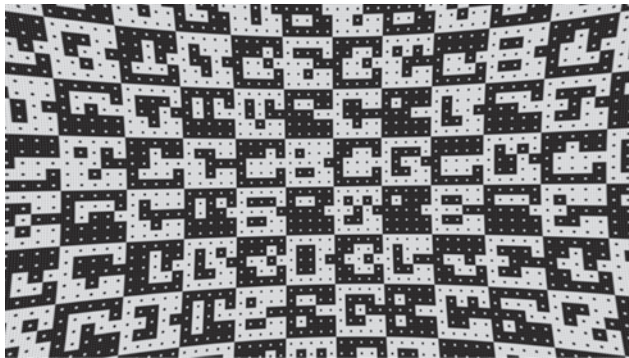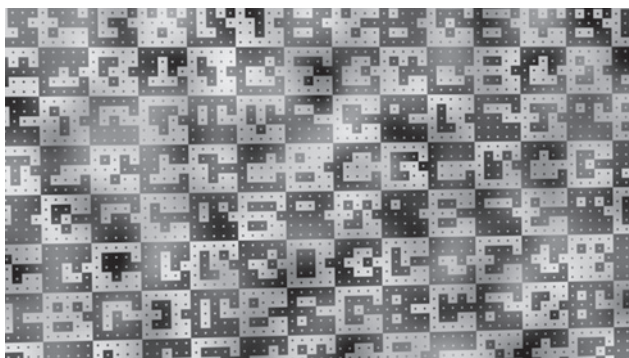
(a) view angle

(b) radial distortion

(c) uneven illumination

Figure 5: Examples of the applied degradations. The view angle images were rendered from different viewpoints, the radial distortion images were interpolated from images with the linear resolution increased by a factor of four. The uneven illumination images were generated by addition and multiplication with very low resolution noise images.

solve only for the camera pose under known camera intrinsics (Perspective-n-Point).

6.1 Ground truth data

The evaluation images were rendered with the cycles renderer from Blender [12] using 1024 samples per pixel, with

Gaussian anti aliasing at a resolution of 1920 x 1080. For distortion simulation the images were rendered with the resolution increased by a factor of four and scaled down after performing the distortion in order to reduce the influence of the interpolation. Figure 5 shows some examples of the calibration target, filtered with the described degradations.

6.2 Error metric

In the literature cited in Section 2, there are two principal methods of evaluating calibration patterns, the root mean square (RMS) error of the calibration model and the RMS error of calibration points against ground truth, called true pixel error (TPE) by Douxchamps and Chihara [10]. The calibration RMS is a poor choice, as it will not detect bias but simply incorporate it into the model. The TPE works well for example for the evaluation of different refinement methods of corner points, but it cannot be used for a meaningful comparison of calibration targets where the number of calibration points is not roughly the same. The reason is that more calibration points of similar localization performance obviously lead to better calibration results which is not reflected in the TPE.

A further method, which is often seen, is to compare triangulated world coordinates of calibration points or the extrinsics of the calibrated cameras. While these are valid metrics for some applications, they will also miss problematic areas like the corners of the image, if there were no calibration points present for that area. This is especially problematic, because the calibrated model will be bad specifically in those areas where no calibration points were acquired.

However, with ground truth data, it is possible to compare the projection defined by the fitted camera model with the known ground truth projection. This gives an error measure which relates to the actual camera calibration and incorporates the error distribution and bias of the used calibration pattern, as well as averaging from multiple markers. We compare the camera models in image space, by projecting each pixel into the scene, at the depth of the target, using the ground truth model, and then projecting them back into image space with the calibrated model. The root mean squared difference between the original pixel coordinate and the projection of the calibrated model defines the error metric, which we call Groundtruth Pixel Error (GPE):

$$\text{GPE}(z) = \sqrt{\frac{\sum_i (i - c_{model}(c_{GT}^{-1}(i, z)))^2}{\#p}}. \qquad (1)$$

The GPE is calculated using the root mean square error between every pixel coordinate i and the results of projecting this pixel into the world using the depth z, the known ground truth projection c_{GT} and reprojecting with the evaluated model c_{model}.

For comparison, when $n \in (I, T)$ iterates all target to image correspondences (I, T) detected by the calibration pattern, the RMS is calculated as:

$$\text{RMS} = \sqrt{\frac{\sum_n \left(I_n - c_{model}(T_n)\right)^2}{\#n}}, \tag{2}$$

which shows how the RMS is merely the residual between model and (detected) calibration points, while the TPE

$$\text{TPE} = \sqrt{\frac{\sum_n \left(I_n - c_{GT}(T_n)\right)^2}{\#n}}, \tag{3}$$

shows the error between ground truth and detection, but again only for the detected image points.

The GPE measures an error in pixel units, stating the localization accuracy of the tested pattern in the context of the tested calibration model. The metric gives a score averaged over the whole image, including the edges and corners, which are the most problematic areas in the context of camera calibration, both because they exhibit the strongest distortion and because it is difficult to place calibration markers on the edges of the image.

For this evaluation the z-coordinates are calculated from the known target position, using the depth of the first view. Alternatively the metric can also be calculated over a depth volume, using a number of depth slices for the reprojection. However the results then depend more on how good the calibration method can estimate the camera model from a limited number of slices through this volume, than from the marker accuracy. Therefore in the following the GPE is always measured at the depth of the calibration target. If multiple views were tested the depth from the first target is used.

6.3 Comparison

For reference, a checkerboard with 12×6 squares is detected using the checkerboard detection from OpenCV with subpixel refinement over an area of 21 pixels. When evaluating the performance under blur and noise we noticed that under most circumstances the refinement was dramatically improved by performing a Gaussian blur with a sigma of 2 before refinement, which is therefore used in the whole comparison.

6.4 Results

In most cases the results of the fractal target surpass the accuracy of the checkerboard pattern by one to two orders of magnitude, for example in Figures 6 and 7. In general the calibration is very robust to noise, reduced contrast, uneven illumination and radial distortion. Two areas where the method is not as robust and where it is eventually surpassed by the checkerboard detection are detection under shallow angles and under strong blur. The reasons and possible enhancements are discussed in Section 7.
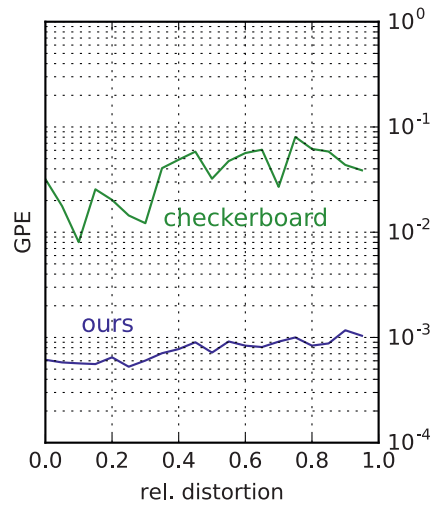

Figure 6: A plot of the GPE obtained dependent on the relative radial distortion, compare Figure 5(b). The fractal target provides both better GPE scores as well as a lower variability of the results. Both methods achieve slightly worse results with increased distortion.
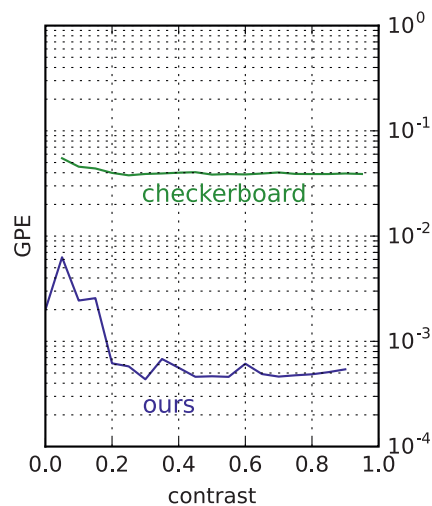

Figure 7: With decreasing contrast the GPE stays constant, but there comes a point where the quality starts to decrease. Reaction is stronger for the fractal target, but the results are always better than for the checkerboard target.

In Figure 6 the GPE is evaluated under varying radial distortion, as shown in Figure 5(b). Both methods are quite insensitive to radial distortion, with only slight increases in GPE with increased radial distortion. The fractal target has over a magnitude lower GPE for the tested range.

Figure 7 shows another variable where both methods show strong performance. Under varying contrast the results a very identical for a large range, with an increase in GPE at low contrast values. Here the results are nearly two orders of magnitude better for the fractal target.

Again similar are the results when the camera is rotated around the camera z axis, see Figure 8. In this case both methods show an interesting symmetry, although the checkerboard method has a higher variance.

One area where the fractal target delivers worse results is under strong perspective, e.g. when looking at shallow angles, see Figure 9. In this case the performance drops under the performance of the checkerboard detection from an angle of 60 degrees from the normal.

An area where the fractal target shows a particularly large advantage is with strong background gradients, as present with uneven illumination as in Figure 5(c). In Figure 10 we can see that the checkerboard target quickly deteriorates under even mildly uneven illumination, where the fractal target stays quite stable. The reason for this behavior is the localization based on the orthogonal gradient method in OpenCV which is very sensitive to gradiation changes, while our fractal target locates individual dots which are nearly unaffected by image gradients.

The effects of Gaussian blur and noise are shown in Figure 11, which shows that our fractal target is very robust with regards to noise, but looses accuracy under strong blur, while the checkerboard deteriorates under noise, but is more robust to blur.
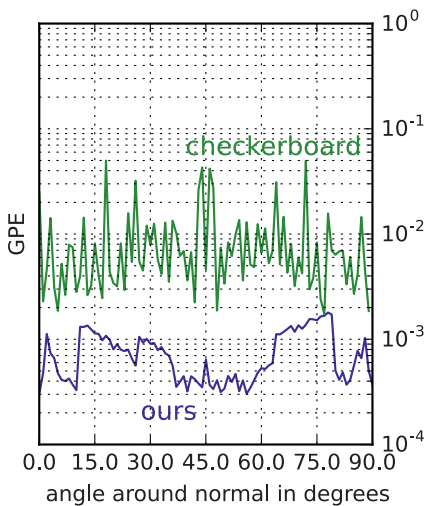

Figure 8: Plot of the dependency of the GPE from rotation around the z axis. Both methods show a distinctive symmetry which hints at some underlying systematic dependency on the z-angle. Results for the fractal target have a lower variance and are more stable.
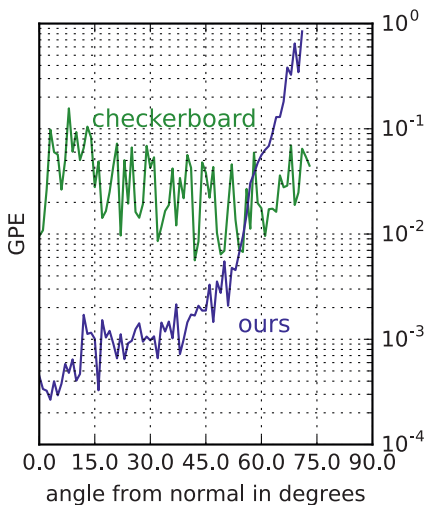

Figure 9: Plot of the dependency of the GPE on the view angle. Zero degrees means the camera points directly towards the target, while at 90 degrees it is oriented parallel to the target. The checkerboard target is insensitive against the view angle, but the fractal target shows a strong dependency and eventually is surpassed by the checkerboard target at around 60 degrees from the normal.
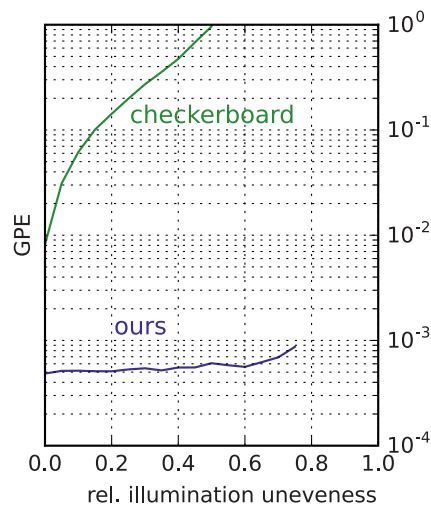

Figure 10: This plot shows the influence of an uneven illumination which increases the gradient overlayed on the calibration target. Because the fractal target derives localization information from individual calibration points there is nearly no influence on the accuracy, as each calibration point is composed of gradients in all directions. On the other hand the checkerboard target derives calibration information by observing the orthogonal gradients around a checkerboard corner and is very sensitive to strong image gradients.
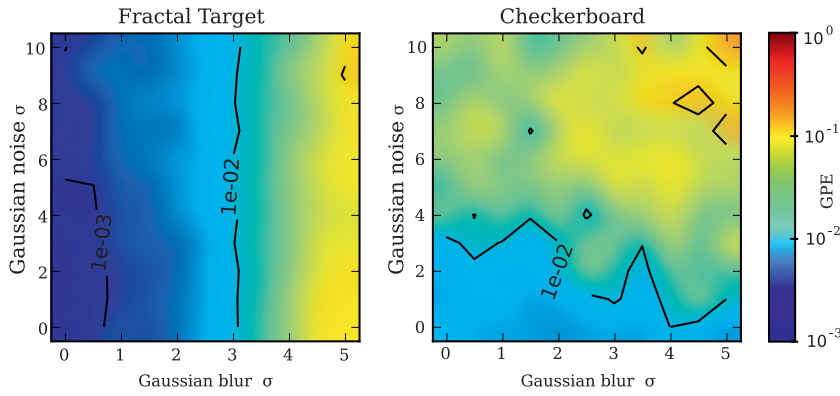
Figure 11: The plot shows the quality of the calibration dependent on both noise and blur. The two targets show quite orthogonal behavior, where the fractal target is robust to noise and the checkerboard target is more robust to blur. Note that the fractal target starts at a much lower error value, hence the checkerboard target only has an advantage in the case of strong blur but very weak noise.
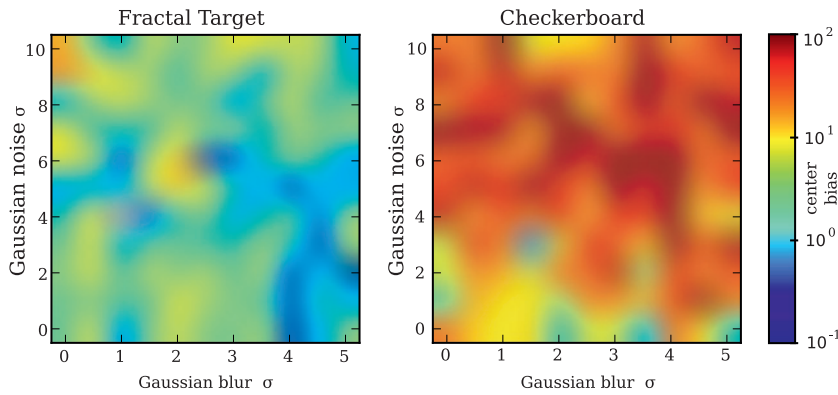

Figure 12: Evaluation of the center bias as defined in Equation (4). The plots show the reduction of center bias by our fractal target. Note that while the overall calibration quality depends on noise and blur, the center bias, which is basically the difference between center and corner results, seems to be independent of those parameters, and is dominated by noise. In addition the center bias is around an order of magnitude smaller for the fractal target. This can be explained by the ability of the fractal target to place calibration points close to the image border and hence improve calibration results in those areas.

One reason for the introduction of the new target was the goal of reducing center bias by adding calibration points close to the edges of the target. To evaluate center bias, the GPE is measured at the four corner pixels, and at the center of the images. The center bias B_c is then defined as:

$$B_c = \frac{GPE_{corner} - GPE_{center}}{GPE_{center}} \qquad (4)$$

A perfect result would thus be a B_c of zero, while a B_c of one means that the corners only reach half the quality of the center. As is visible in Figure 12 the checkerboard calibration has a very strong center bias between 10 and 100. On the other hand the fractal target achieves values for the center bias close to one, and even outliers are below 10. On average the fractal pattern again has an advantage of around one order of magnitude.

These values also explain the even larger advantage of the fractal target in Figure 13 which shows the GPE results for the corners only.

Furthermore Figure 14 shows the number of calibration points that were actually detected, again for different values of blur and noise, while Figure 15 shows the RMS error of the detected individual markers, i.e. not averaged over the calibration (the TPE as defined by Douxchamps and Chihara [10]). From those two plots we can see that, while the RMS error of the individual calibration points is higher in most cases for the fractal target, the number of markers is also much higher which offsets the advantage of the individually higher accuracy markers of the checkerboard pattern, just as argued in Section 4. Indeed when comparing Figures 14 and 15 with Figure 11 it is visible that the quality of the checkerboard calibration is directly linked to the quality of the individual markers as the number of markers is constant, while quality of the calibration
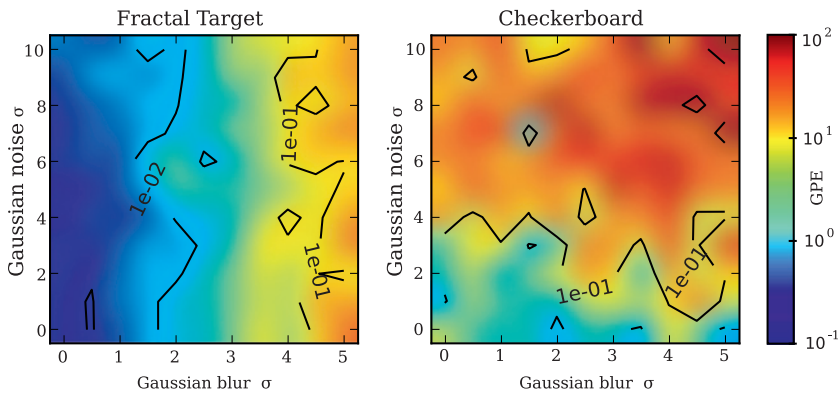
Figure 13: Evaluation of the accuracy in the image corners. Compared to Figure 11 the corners show a more pronounced advantage of the fractal pattern, which also confirms the results of Figure 12.
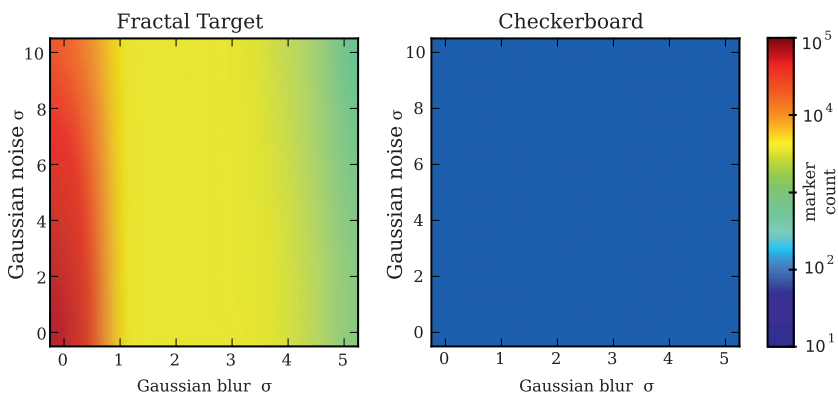

Figure 14: This plot shows the number of individual calibration points found for the targets. The checkerboard always returns the same number (else detection would not be possible), while the fractal target finds a much higher number in good conditions, and reduces this number slightly as the noise increases and much more strongly as the blur increases.
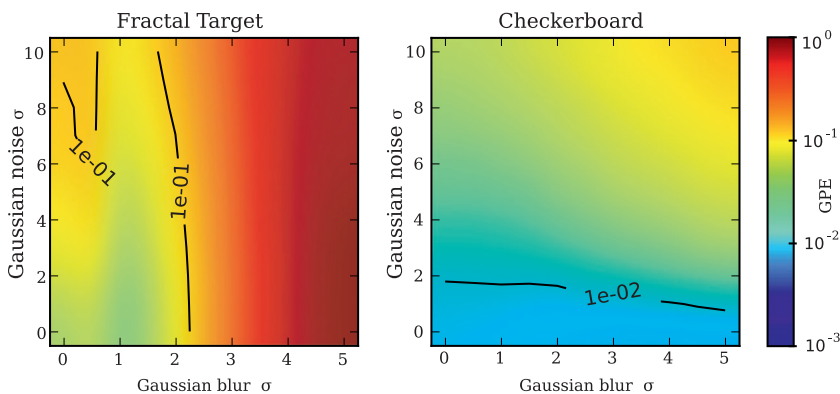

Figure 15: The marker RMS error (TPE) against ground truth shows a similar behavior to the marker count, compare Figure 14, for the fractal target, while the quality decreases only with the increase of noise for the checkerboard target. This is in line with the other results, and shows that the checkerboard target performance depends solely on the performance of the checkerboard corner detection.

with the fractal target is reduced as the number of markers decreases and the individual marker RMS increases due to blur.

7 Discussion

While the recursive target delivers most of the anticipated advantages, with highly improved general accuracy, it fails

to completely remove center bias and it performs badly under strong blur and shallow view angles.

Regarding the center bias, it seems a dense target is not enough to completely solve the problem. A possible solution could be to introduce a weighting scheme into the calibration method which weights corner samples stronger, however more work is required to estimate good weighting factors and the influence on the overall calibration results.

The second shortcoming, low robustness under shallow view angles and strong blur is a direct results of the desired characteristic of uncoupling the pattern detection from the calibration process. As Mallon and Whelan point out [7], most patterns need to incorporate the calibration within the pattern refinement to remove perspective and distortion bias from the pattern localization. While the checkerboard refinement does not suffer from this drawback, the 2D Gaussian fits used in our target do have this problem. The workaround is to use very small calibration points where this bias is very small, which works for most cases as the fractal structure ensures that the smallest calibration points can be used independently from magnification factor. This fails only if the smallest scale cannot be used for another reason than the image scale, which are blur and the perspective distortion, which appears at shallow view angles.

Reliable correction of this bias needs the calibration information, which would tie the pattern detection to the camera calibration. However this would reduce the universality of the calibration pattern which in the current form can be used completely independent of the camera calibration.

8 Conclusions

For computer vision tasks that require a high accuracy, the calibration target can be a limiting factor in the camera calibration. As we have shown in this work, the use of a fractal calibration target can dramatically improve accuracy and reduce center bias. The target was evaluated and is quite robust under a range of imaging conditions. With an accuracy between a hundredth and a thousandth of a pixel, the target provides an excellent basis for the development of new calibration models, as deficiencies in the camera model are much more apparent with more precise raw calibration data, as provided by the fractal target. Due to the small feature size the pattern detection can be completely decoupled from the camera calibration, which simplifies development of new calibration methods. To support further research our implementation is available under an open source license [5].

References

1. Itseez. Open source computer vision library. https://github.com/itseez/opencv, 2015.
2. M. Diebold, O. Blum, M. Gutsche, S. Wanner, C. Garbe, H. Baker, and B. Jähne. "Light-field camera design for high-accuracy depth estimation", in *SPIE Optical Metrology*. International Society for Optics and Photonics, 2015, pp. 952803–952803.
3. B. Atcheson, F. Heide, and W. Heidrich. "CALTag: High Precision Fiducial Markers for Camera Calibration", in *Vision, Modeling, and Visualization (2010)*, R. Koch, A. Kolb, and C. Rezk-Salama, Eds., The Eurographics Association, 2010.
4. S. Daftry, M. Maurer, A. Wendel, and H. Bischof. "Flexible and user-centric camera calibration using planar fiducial markers", in *BMVC*, 2013.
5. H. Siedelmann. "A high density fractal calibration pattern", http://hci-repo.iwr.uni-heidelberg.de/hsiedelm/hdmarker, 2016.
6. L. Lucchese and S. K. Mitra. "Using saddle points for subpixel feature detection in camera calibration targets", in *APCCAS (2)*. IEEE, 2002, pp. 191–195.
7. J. Mallon and P. F. Whelan. "Which pattern? biasing aspects of planar calibration patterns and detection methods", *Pattern recognition letters*, vol. 28, no. 8, pp. 921–930, 2007.
8. S. Placht, P. Fürsattel, E. A. Mengue, H. G. Hofmann, C. Schaller, M. Balda, and E. Angelopoulou. "Rochade: Robust checkerboard advanced detection for camera calibration", in *ECCV (4)*, ser. Lecture Notes in Computer Science, D. J. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars, Eds., vol. 8692. Springer, 2014, pp. 766–779.
9. A. Datta, J.-S. Kim, and T. Kanade. "Accurate camera calibration using iterative refinement of control points", in *Computer Vision Workshops (ICCV Workshops), 2009 IEEE 12th International Conference on*. IEEE, 2009, pp. 1201–1208.
10. D. Douxchamps and K. Chihara. "High-accuracy and robust localization of large control markers for geometric camera calibration", *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 2. pp. 376–383, 2009.

11. H. Siedelmann. "Recording, Compression and Representation of Dense Light Fields", Diplomarbeit, Universität Stuttgart, Fakultät Informatik, Elektrotechnik und Informationstechnik, Germany, Januar 2015.
12. Blender Online Community. *Blender – a 3D modelling and rendering package*. Blender Foundation, Blender Institute, Amsterdam, 2016 [Online]. Available: http://www.blender.org.

Bionotes

Hendrik Schilling
Heidelberg University, Heidelberg Collaboratory for Image Processing (HCI), Berliner Straße 43, 69120 Heidelberg, Germany
Hendrik.Siedelmann@iwr.uni-heidelberg.de

Hendrik Schilling has received his Diploma in Computer Science in 2015 at the University of Stuttgart. He is currently working on his PhD thesis at the Heidelberg Collaboratory for Image Processing (HCI), in collaboration with Sony Europe Limited. His research includes camera calibration and light eld imaging.

Maximilian Diebold
Heidelberg University, Heidelberg Collaboratory for Image Processing (HCI), Berliner Straße 43, 69120 Heidelberg, Germany
Maximilian.Diebold@iwr.uni-heidelberg.de

Maximilian Diebold is Leader of the Light-Field Imaging Group at the Heidel- berg Collaboratory for Image Processing since April 2016. He studied Electrical Engineering and Information Technology at the University of Karlsruhe (Dipl.- Ing.) and received his PhD in computer science at the combined faculty for the natural science and mathematics at the University of Heidelberg (Dr.rer.nat.) in April 2016. The thesis was collaborated with Sony Europe Limited. His current interest is to decompose image information by taking a holistic point of view about depth, material, reectance, color, polarization and illumination from a large set of optical measurements i.e. light-eld measurements and use this information to generate what he terms tunable ground truth.

Marcel Gutsche
Heidelberg University, Heidelberg Collaboratory for Image Processing (HCI), Berliner Straße 43, 69120 Heidelberg, Germany
Marcel.Gutsche@iwr.uni-heidelberg.de

Marcel Gutsche has received his Master degree in Physics from Heidelberg University in 2014. He is currently working on his PhD thesis at the Heidelberg Collaboratory for Image Processing (HCI), in the area of object and BRDF reconstruction from light eld data.

Bernd Jähne
Heidelberg University, Heidelberg Collaboratory for Image Processing (HCI), Berliner Straße 43, 69120 Heidelberg, Germany
Bernd.Jaehne@iwr.uni-heidelberg.de

Bernd Jähne received his Diploma, Doctoral degree and Habilitation degree in Physics from Heidelberg University in 1977, 1980, and 1985, respectively, and a Habilitation degree in Applied Computer Science from the University of Hamburg-Harburg in 1992. From 1988 to 2003 he hold a research professorship at the Scripps Institution of Oceanography, University of California in San Diego. Since 1994 he is professor at the Interdisciplinary Center for Scientic Computing (IWR) and Institute for Environmental Physics of Heidelberg University in 1994 and since 2008 he heads the Heidelberg Collaboratory for Image Processing (HCI). His research interests include small-scale air-sea interaction and image processing.