University of Passau
Technical Report

On the point spread function introduced by first order optics

Tobias Hanning and Oleg Smirnow *

October 16, 2007

In computer vision the blurring of observed objects in an image is modeled by a convolution with a point spread function (PSF). The extraction of basic image feature like points, lines, circles or ellipses is done generally without knowing the PSF exactly. Therefore, the point extraction must result in an erroneous position. In this article we show that even for first order optics the point spread function is not symmetrical. Furthermore, we show that it depends on the position of the observed point. Therefore, for optical measurement systems any feature point extraction should be corrected with respect to this position.

Keywords: point spread function, first order optics, feature extraction

Contents

*{hanning,smirnow}@forwiss.uni-passau.de

1 Introduction

Everyone in computer vision knows the lens makers equation: Theoretically there is a determined object plane where the observation of an object appears sharp in the image. At any other position of the object the observation must appear blurred. In general, the sensor plane, i. e. the plane of the imaging sensor in the camera, does not coincide with the image plane. This can be observed by a blurring of the image.
The common camera model in computer vision does not include this effect. It simply models a point to point relation and leaves the extraction of the observed feature point to the point extraction algorithm. To deal with the blurring effect the image is modeled as a convolution of an ideal image with a mollifying kernel. Without any information of the optical device one generally assumes that the kernel is symmetrical and identical for every pixel. But, as we show, even within the theory of first order optics this assumption is not true.

1.1 Related work

The problem of the determined object plane is not new to computer vision. In fact, there are many articles which propose to use the blurring to estimate the depth of an object. The dependency of sharpness on depth can be used to estimate the depth of an object (see e.g. [AFM98], [SG87] or [Asl03]). These methods use a very simple PSF model which does not supply an accuracy that is needed in optical measurement tasks. In most cases the PSF is simply modeled as a Gaussian.
Another way to achieve a completer sight of the optical behavior of an observed object is introduced by the 4D light field theory (see [AB91] for an introduction). The 4D light field is defined as a function from the set of all lines (light rays) to a light value. There are two principle ways to estimate this function, which is also called plenoptic function: Either by a well-defined movement of the camera ([MP04]) or by a special plenoptic camera ([AW92], [NLB+05]). The plenoptic camera is a camera with a special sensor array of micro-cameras. Such a camera can be used to approximate a point spread function, but in general the resolution is not sufficient to observe a non-symmetrical PSF.
We omit such a holistic approach to the PSF, since we are not interested in the whole light field but in the light distribution on the sensor plane. For our theoretical derivation we use the model of paraxial geometric optics with a thin lens. Furthermore, we can simplify the problem by assuming a point light source as Dirac impulse. We derive the PSF as response of the optical system on this impulse.

1.2 Overview

In the following section we introduce the first order optics with the "thin" lens assumption from which the standard camera model can be derived. Although this is a very simple optical model it will lead to an asymmetric PSF. In the third section we discuss the impact of the third order optics on the PSF. We conclude that every point extraction

based on a symmetric PSF will lead to erroneous results. In the fourth section we show by a simple experiment that there is a depth correlated effect in camera calibration.

2 Geometric optics for computer vision

2.1 The "thin lens" assumption and first order optics

First order optics simplifies the refraction law of Snell[1] $n_1 \sin(\theta_1) = n_2 \sin(\theta_2)$ to

$$n_1\theta = n_2\theta_2 \tag{1}$$

assuming that all involved angles θ are so small that $\sin(\theta) \approx \theta$ holds. (n_1, n_2 are the refraction indexes of two medias, θ_1, θ_2 the angles of the light beam to the normal of the interface of two medias). Furthermore, we assume a thin lens: The (spherical) lens is assumed to be a infinitesimal thin. From the refraction point of view the thin lens behaves like a spherical lens. But the distance which a light ray covers inside the lens is infinitesimal small. So, from a localization point of view the lens is a plane. A light ray passing the thin lens is refracted at the surface air/lens satisfying (1) and immediately after that at the surface lens/air (again satisfying (1)). An immediate consequence of the thin lens assumption is that any light ray passing the thin lens at the optical axis will not be refracted, since both surfaces of the thin lens are parallel at this point (see e. g. [FP02]).
Another consequence of the thin lens model combined with the paraxial optics simplification is that for the light ray emitted at a point p at one side of the lens passing the lens meet in a point i_p at the other side of the lens such that

$$\frac{1}{d_p} + \frac{1}{d_i} = \frac{1}{f} \tag{2}$$

holds, with $f = \frac{r}{2(n_l-1)}$ and d_p resp. d_i being the distance of the object p resp. the image point i_p to the lens plane and n_l the refraction index of the lens (see [Hec87] for more details). (2) is often called *lens maker's equation*. This means there is a determined distance behind the lens plane, in our sketch noted as d_i, at which an observation of a point from the object side appears sharp. An object plane parallel to the lens plane determines a so called *image plane* parallel behind the lens plan, where the points of the object plane appear sharp.

2.2 PSF by first order optics

For every observed object point there is one ray emitted by an object point, which passes the center of the lens. We call this ray the *center ray*. Let f_{ideal} be the image which is obtained by the center ray only. One can imagine this as the image of a pinhole camera. Of course, this image can not be seen anywhere. The real sensor input function f is

$$f = f_{\text{ideal}} * k \tag{3}$$

[1] Named after Willebrord van Roijen Snell, 1580 - 1626

for a mollifier $k \in \mathcal{L}^2(\mathbb{R}^2, \mathbb{R}) \cap \mathcal{C}^\infty(\mathbb{R}^2, \mathbb{R})$. The center ray function f_{ideal} may be not continuous, but piecewise continuous. For mathematical reasons we assume that f_{ideal} is twice integrable. Therefore, the input function for the sensor array f is continuous as a convolution of a $\mathcal{L}^2$-function with a smooth function. We call k the *point spread function* (PSF). The PSF can be seen as the pulse response of the imaging system. All rays emitted from an infinitesimal small object point which pass the lens will hit an sensor plane parallel to the lens plane in a circle, the *circle of confusion* (see [Hec87]). For any plane but the image plane this circle has a positive diameter. The circle of confusion is also the support of the PSF. A simple derivation of the fact that the PSF is not symmetric can be found in [Hor86]. There it is shown that the irradiance of a observed point depends on the angle of the object to the lens. We extend this approach to determine the PSF on the circle of confusion for an image obtained off the image plane. In our approach it is a function of the position of the object. For our considerations we represent every point in an cylindrical coordinate system defined by the optical center O: Every point in this coordinate system can be represented by its distance z to the lens plane, its distance to the optical axis r and an angle ϕ to a fixed axis in the lens plane. Let the object be a point $p = (\phi_p, r_p, z_p)$ emitting light in every direction. Further let $i_p = (\phi_{i_p}, r_{i_p}, z_{i_p})$ be the position of the ideal projection of the object point in the image plane and $c = (\phi_c, r_c, z_c)$ the center of the circle of confusion for the point object in the sensor plane. We assume that the sensor plane is located at a distance d to the image plane.

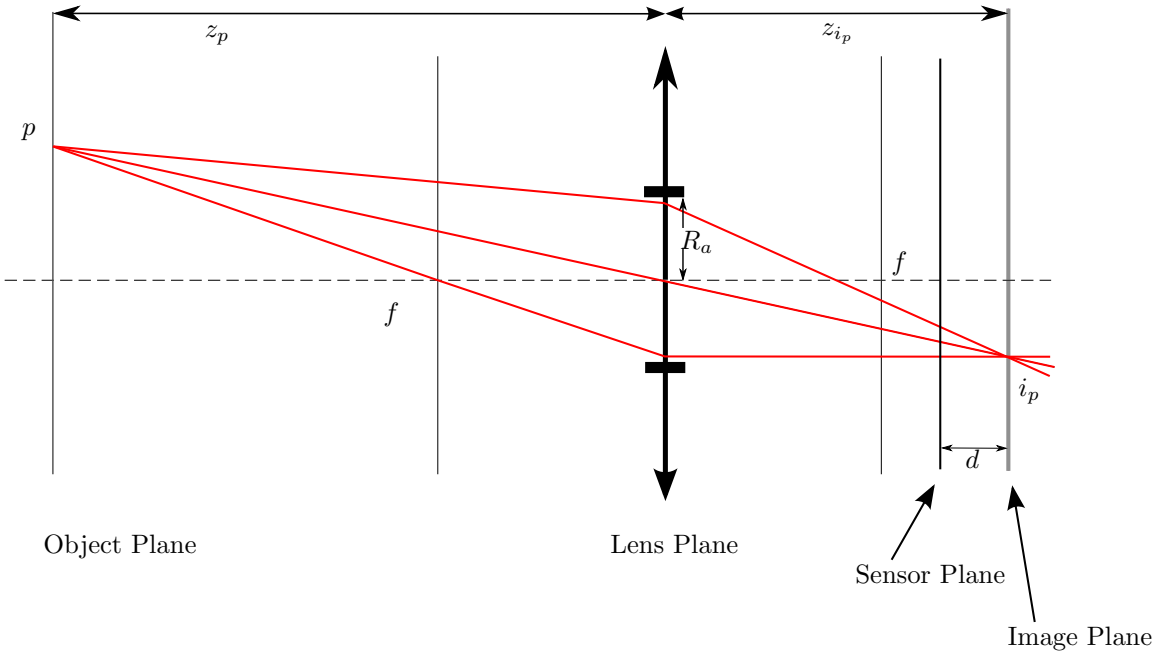

Figure 1: Identifiers for the object point p, ideal point projection i_p lens, image plane and sensor plane in cylindrical coordinates

The coordinates $(\phi_{i_p}, r_{i_p}, z_{i_p})$ of the ideal image point on the image plane can be expressed w.r.t. the object point p by

$$\phi_{i_p} = \phi_p + \pi, r_{i_p} = \frac{r_p f}{z_p - f} \text{ and } z_{i_p} = -\frac{z_p f}{z_p - f} \tag{4}$$

This result follows by simple considerations (and applying the theorem on intersecting lines) w.r.t. to first order optics and a thin lens with focal length f.
The coordinates (ϕ_c, r_c, z_c) of the center of the observed spot in the sensor plane are

$$\phi_c = \phi_{i_p}, r_c = r_{i_p} + \frac{dr_{i_p}}{z_{i_p}} \text{ (it is } z_{i_p} < 0!) \text{ and } z_c = z_{i_p} + d. \tag{5}$$

Let now R_a be the radius of the aperture, which in our model is the radius of the translucent circle of the lens. Then the radius R_s of the circle of confusion is $R_s = \left|\frac{R_a d}{z_{i_p}}\right|$.

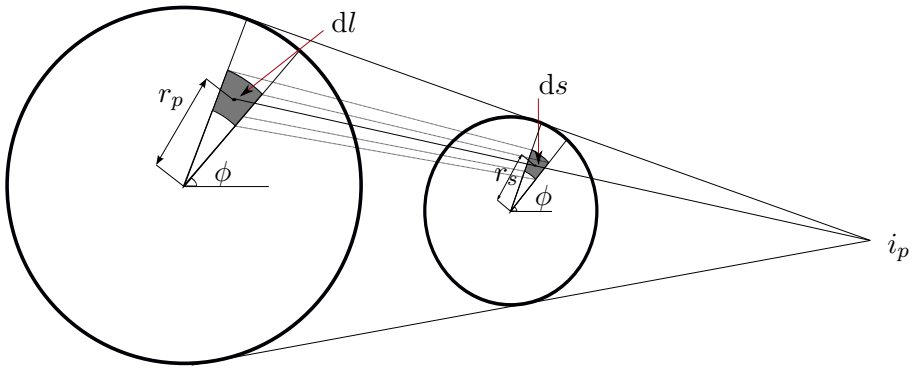

Figure 2: Identifiers for the infinitesimal areas dl on the lens and ds on the sensor plane

Let $\mathrm{d}s = r_s \mathrm{d}s\mathrm{d}\phi$ be an infinitesimal area of the spot and dl be the corresponding area on the lens plane. Then $\mathrm{d}l = \left(\frac{z_{i_p}}{d}\right)^2 \mathrm{d}s$ and $r_l = \left|\frac{z_{i_p}}{d}\right| r_s$ and $\phi_l = \phi_s$ holds.
The distance h from the infinitesimal area on the lens dl to object p is

$$h^2 = (r_p \cos(\phi_p) - r_l \cos(\phi_l))^2 + (r_p \sin(\phi_p) - r_l \sin(\phi_l))^2 + z_p^2 \tag{6}$$

The angle between the normal of the lens and the ray from the object is $\alpha = \arccos(\frac{z_p}{h})$. Let d$l'$ be the infinitesimal area perpendicular to this ray: $\mathrm{d}l' = \mathrm{d}l \cos(\alpha) = \mathrm{d}l \frac{z_p}{h}$. The (infinitesimal) solid angle dw spanning this area is given by

$$\mathrm{d}w = \frac{\mathrm{d}r_l'}{l^2} = \mathrm{d}r_l \frac{z_p}{l^3} = \frac{z_p z_{i_p} r_s}{l^3 d^2} \mathrm{d}r_s \mathrm{d}\phi \tag{7}$$

Now, we are considering an isotropic light source with intensity of emitted licht I. Let dI be the intensity emitted over dw, then it is

$$\mathrm{d}I = I \mathrm{d}w = I \frac{z_p z_{i_p}^2}{h^3 d^2} \mathrm{d}s \tag{8}$$

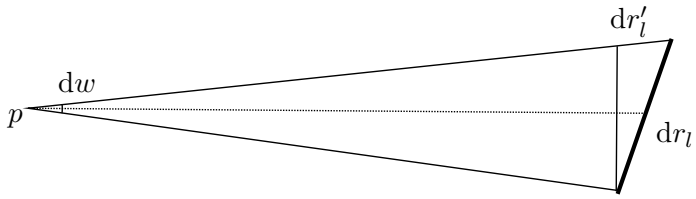

Figure 3: The object p, the area $\mathrm{d}l$ and the area $\mathrm{d}l'$ perpendicular to the light ray

and therefore the intensity N of the light in the spot s becomes

$$N = I\frac{z_p z_{i_p}^2}{h^3 d^2} \tag{9}$$

Fig. 4 and 5 show two examples of the the point spread function N for different lens parameters. Fig. 4 shows an unrealistic parametrization of a lens since the distance of the object to the lens is smaller than the twice the focal length. The purpose of this exaggeration is to point out the asymmetry of the point spread function. The resulting PSF for a more realistic setup is displayed in Fig. 5. As one in this case there is only a slight asymmetry. But, in contrast to the common assumption, the PSF is not a Gaussian at all.

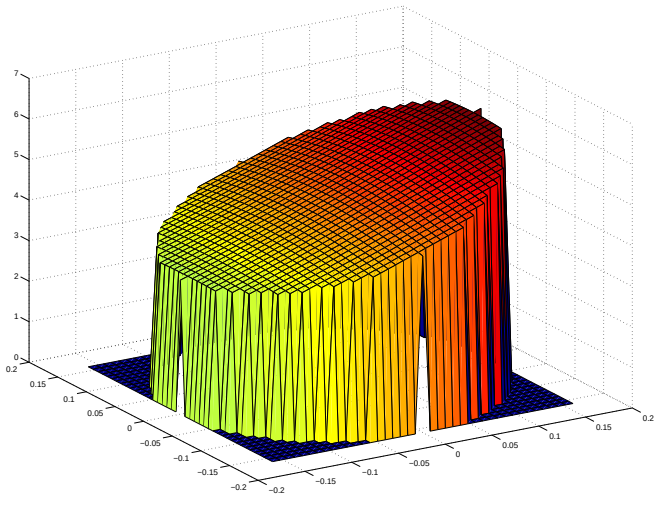

Figure 4: Intensity N of a spot in the sensor plane for an object at $(\phi_p, r_p, z_p) = (0, 4, 7)$ for a lens with focal length 4 and radius of aperture 3. The displacement d of the sensor plane is 0.5

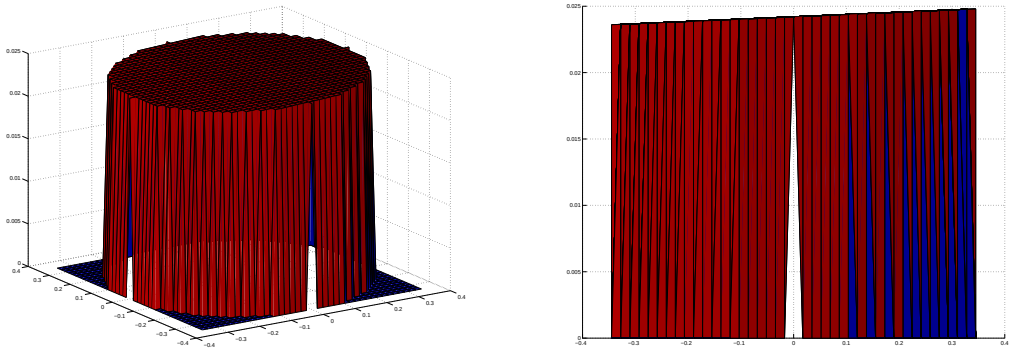

Figure 5: Intensity N of a spot in the sensor plane for an object at $(\phi_p, r_p, z_p) = (0, 20, 50)$ for a lens with focal length 4 and radius of aperture 3. The displacement d of the sensor plane is 0.5

2.3 The impact of third order optics

The simplification $\sin(\theta) \approx \theta$ is only applicable for small angles θ. For wide angle lenses the Taylor series of the sine must expanded to the third order: $\sin(\theta) \approx \theta - \frac{1}{6}\theta^3$. This leads to the so called third order optics. It is easy to see that in third order optics the refraction of a light ray depends on its distance to this axis. So, the third order optics with a thick lens yields monochromatic aberrations in every optical system. The so called *five Seidel aberrations*[2] can all be modeled and derived by this model assumption (see [Hec87]). The Seidel aberrations are (in this particular order):

i. Spherical aberrations

   Rays that hit the surface of a spherical lens at a greater distance to the optical axis will be more focused to the apex than rays at a lower distance. This leads to a circle of confusion even in the image plane.

ii. Coma

   A thick lens can be modeled as two thin lenses introducing two lens planes. But in fact, these "planes" can only be treated as planes near the optical axis. So, rays coming from an object not on the optical axis, will be focused off the optical axis, which is described by these two planes.

iii. Astigmatism

   A pencil of rays emitted from object at a great distance to the optical axis will hit the surface of the lens not symmetrically. This also causes a deformation of the circle of confusion. Modern lenses suppress astigmatism.

[2]Named after Ludwig Seidel, 1821 - 1896

iv. Curvature of field

The refraction behavior of the third order optics has also impact on the relationship of object and image points. For a given object plane parallel to the lens plane the area of focused image points is not a plane but a curved area. Only for points near the optical axis it can be approximated by a plane. So, given a planar sensor plane it is theoretically not possible to get a sharp image in every pixel.

v. Distortion

In third order optics the transversal magnification in the image plane becomes a function in the distance of the observed image to the optical axis. Instead of the aberrations above the distortion also effects the center ray. The distortion as described by Seidel is completely determined by it's observation in the sensor plane.

Only the distortion can be handled by a 3D-2D camera mapping. All other aberrations deflect the circle of confusion. So far, we have not carried out a closed form solution for the PSF introduced by third order optics. But, we assume that the shape of a PSF derived by third order optics will be more like as Gaussian.

3 Experiments

To substantiate the result of a position dependent PSF we determined the distortion parameters at different distances to the camera. For our experiments assume a well calibrated camera $\mathcal{K} : \mathbb{R}^3 \to \mathbb{R}^2$ with $\mathcal{K} = P \circ \delta \circ \mathrm{P}_z$. Where it is

$$\mathrm{P}_z : \begin{array}{c} \mathbb{R}^2 \times \mathbb{R} \setminus \{0\} \to \mathbb{R}^2 \\ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} x/z \\ y/z \end{pmatrix} \end{array} . \tag{10}$$

the central projection and

$$P : \begin{array}{c} \mathbb{R}^2 \to \mathbb{R}^2 \\ \begin{pmatrix} u \\ v \end{pmatrix} \mapsto \begin{pmatrix} \alpha & \gamma \\ 0 & \beta \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} + \begin{pmatrix} u_0 \\ v_0 \end{pmatrix} \end{array} \tag{11}$$

the change of the camera coordinate system of the image plane to the observed pixel coordinate system. See [HG07] for more details on the camera mapping.
The distortion δ is placed after the projection and before the transformation from the camera coordinate system to the pixel coordinate system. The most common distortion model is the one of a radial distortion (see e.g. [Atk96]).

$$\delta \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} u + u \sum_{i=1}^{D} k_i (u^2 + v^2)^i \\ v + v \sum_{i=1}^{D} k_i (u^2 + v^2)^i \end{pmatrix} \tag{12}$$

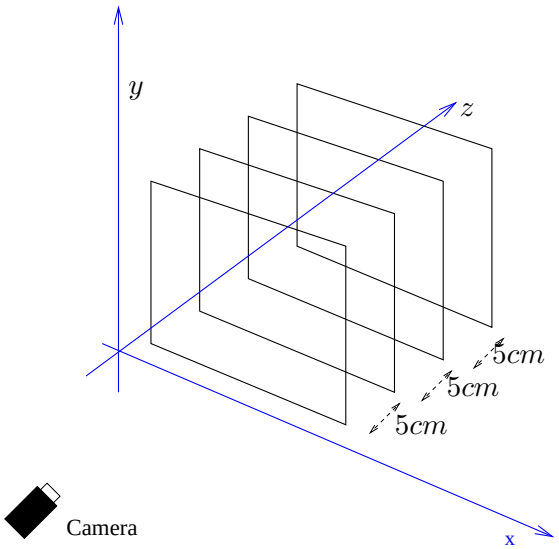

Figure 6: Setup for the distortion test

with parameters $k_1, \ldots, k_D$. Many contributions can be found in the literature to perform this task: [HS97] or [Zha98] deliver useful algorithms for this purpose. They all have in common that they minimize the error

$$\Psi_1 : \mathcal{K} \mapsto \sum_{p \in \mathbf{P}} \|\mathcal{K}(p) - i_p\|^2 \tag{13}$$

This means to find a camera mapping such that extracted observed points i_p fit best to the projected points $p \in \mathbf{P}$ for a calibration pattern $\mathbf{P} \subset \mathbb{R}^3$. In our case the camera calibration was performed following the algorithm of Zhang (see [Zha98], esp. his remark on parallel calibration targets).
In our experiment the observation of the prototype filled out the whole image. For our experiments we choose a radial distortion with four coefficients (i. e. $D = 4$ in (12)). We used a calibration pattern on a plane at different distances located equidistantly nearly parallel to the sensor planes. The planes had a distance of 5cm to each other covering distances from 60 to 135cm to the camera. The calibration target on the plane was a 51 x 49 point grid with a point distance of 20mm to each other. The diameter of the points was 2mm and we used the gray value weighted barycenter of the points as feature extractor. The camera had a common 2/3" CCD sensor and a Pentax 6mm lens.
For each plane we determine the parameters of the radial distortion mapping. In Fig. 7 - 10 we displayed each distortion parameter $k_1, \ldots, k_4$. The first plane is the plane nearest to the camera (i. e. the plane a at distance of 60cm to the camera).One can see that there is an observable systematic change of the value of all distortion parameters obtained by this setup.

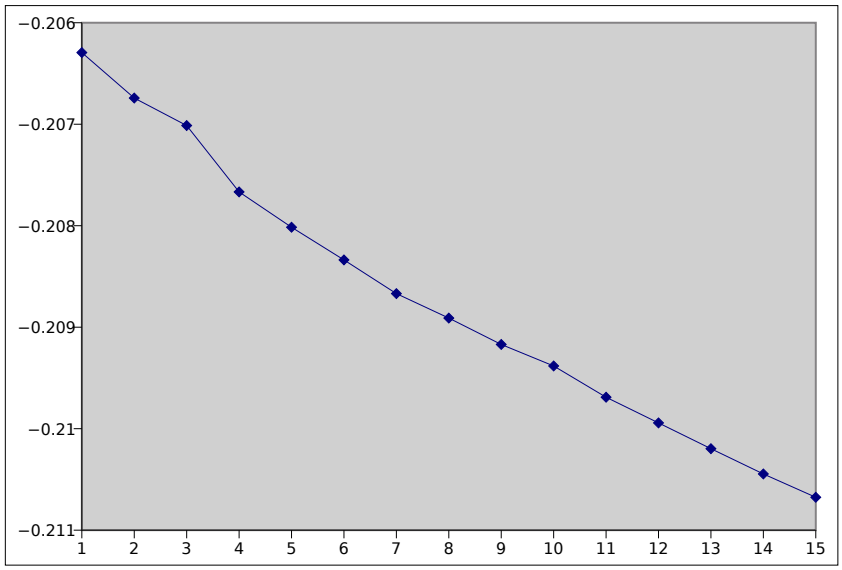

Figure 7: Value of the parameter k_1 (ordinate) w.r.t. the depth of the observed plate (abscissa)

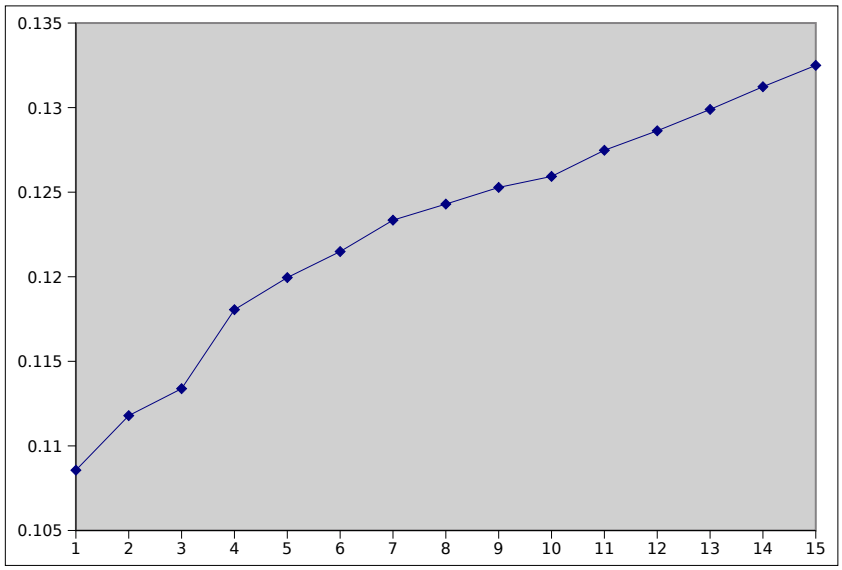

Figure 8: Value of the parameter k_2 (ordinate) w.r.t. the depth of the observed plate (abscissa)

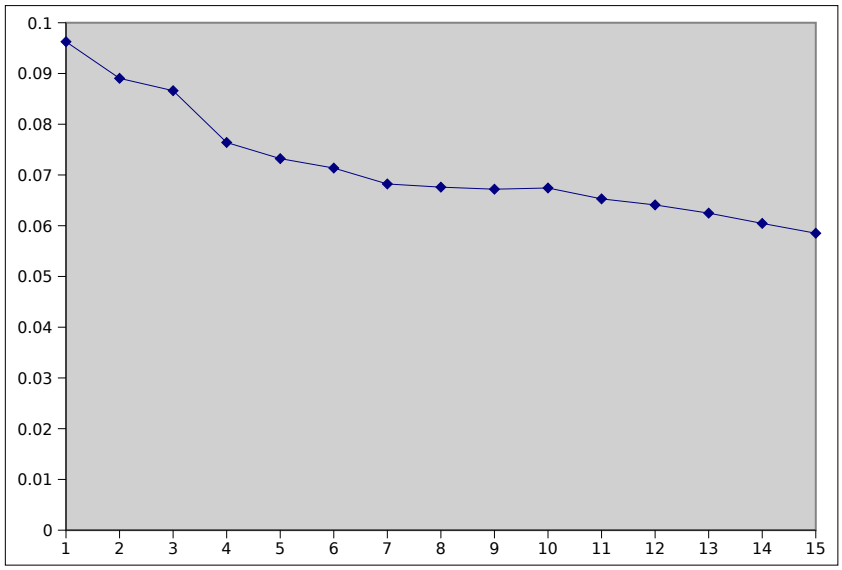

Figure 9: Value of the parameter k_3 (ordinate) w.r.t. the depth of the observed plate(abscissa)

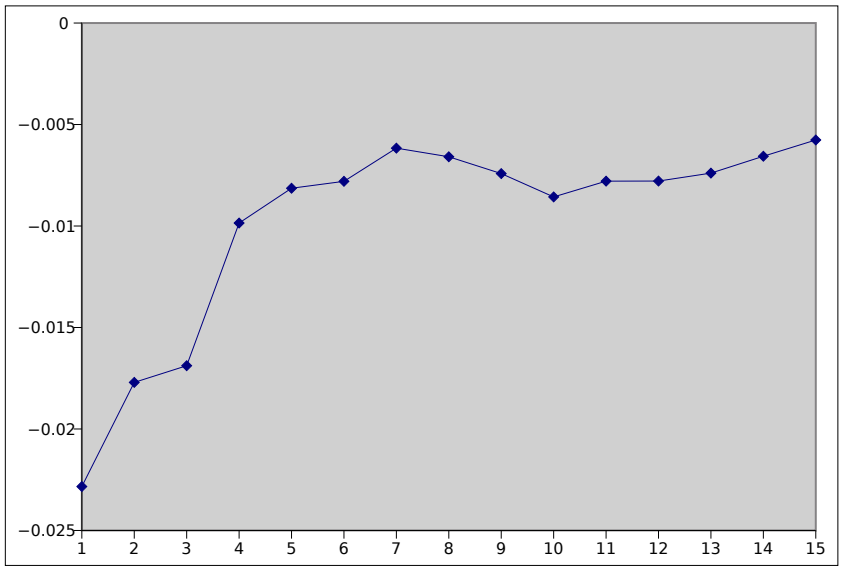

Figure 10: Value of the parameter k_4 (ordinate) w.r.t. the depth of the observed plate (abscissa)

4 Conclusion and discussion

Point extraction methods assume a symmetric and identical PSF for every pixel. We showed that this assumption is wrong even within the assumption of first order optics and a thin lens. The PSF depends on the position of the pixel in the image and on the depth of the observed object. Since the feature point extraction is based on the assumption of a symmetrical and identical PSF for every pixel position, camera calibration algorithms will result in an erroneous configuration. In most cases a camera will be calibrated w.r.t. a specific plane, where most measurements will be made. But, when the measured objects appear out of focus problems arise. A calibration algorithm for a camera setup which has to covers a range with "out of focus" areas will lead to poor reconstruction results.

References

[AB91] E. H. Adelson and J. R. Bergen. The plenoptic function and the elements of early vision. *M. Landy and J. A. Movshon, (eds) Computational Models of Visual Processing*, 1991.

[AFM98] N. Asada, H. Fujiwara, and T. Matsuyama. Edge and depth from focus. *International Journal of Computer Vision*, 26(2):498 – 503, 1998.

[Asl03] V. Aslantas. Estimation of depth from defocusing using a neural network. *Internation Journal of Computer Intelligence (IJCI)*, 1(2):305 – 309, 2003.

[Atk96] K.B. Atkinson, editor. *Close Range Photogrammetry and Machine Vision*. Whittle Publishing, 1996.

[AW92] E. H. Adelson and J. Y. A. Wang. Single lens stereo with a plenoptic camera. *IEEE Transactions on Pattern Analysis and machine Intelligence*, 14(2):99–106, 1992.

[FP02] D. A. Forsyth and J. Ponce. *Computer Vision: A Modern Approach*. Prentice Hall, New York, 2002.

[Hec87] E. Hecht. *Optics*. Addison-Wesley, Reading, 2. edition, 1987.

[HG07] T. Hanning and S. Graf. Euclidean vs. projective camera calibration: Algorithms and effects on 3d-reconstruction. Technical Report MIP-0708, Fakultät für Informatik und Mathematik, Universität Passau, 2007.

[Hor86] B. K. P. Horn. *Robot vision*. MIT Press, Cambridge, MA, USA, 1986.

[HS97] J. Heikkilä and O. Silven. A four-step camera calibration procedure with implicit image correction. In *Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1106 – 1112, San Juan, Puerto Rico, 1997. IEEE.

[MP04] W. Matusik and H. Pfister. 3d tv: a scalable system for real-time acquisition, transmission, and autostereoscopic display of dynamic scenes. *ACM Trans. Graph.*, 23(3):814 – 824, 2004.

[NLB+05] R. Ng, M. Levoy, M. Bredif, G. Duval, M. Horowitz, and P. Hanrahan. Light field photography with a hand-held plenoptic camera. Technical report, Stanford Computer Graphics Laboratory, 2005. Computer Science Tech Report CSTR 2005-02.

[SG87] M. Subbarao and N. Gurumoorthy. Depth recovery from blurred edges. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 498 – 503, Ann Arbor, Michigan, 1987.

[Zha98] Z. Zhang. A Flexible new technique for camera calibration. Technical report, Microsoft Research, 1998. Technical Report MSR-TR-98-71.