On Chromatic and Geometrical Calibration

Jørgen Folm-Hansen

Lyngby 1999
IMM-PHD-1999-61
Ph.D. Thesis

ISSN 0909-3192

Printed by IMM, Technical University of Denmark, DK-2800 Lyngby.
Bound by Hans Meyer.

Preface

This thesis fulfills in part the official requirements for obtaining the degree of Ph.D. in engineering from the Technical University of Denmark. The thesis has been prepared at the Section for Image Analysis, Department of Mathematical Modelling (IMM).

The main subject of the thesis is different methods for the geometrical and chromatic calibration of cameras in various environments. Emphasis is put on methods for calibrating monochrome, colour and multi-spectral images to get a high accuracy of measurement. The presented treatment is not intended to cover all aspects of the considered subject. It is rather meant as a presentation of topics selected from the author's knowledge and experience, for which some amount of apparently new insight has been made during this Ph.D.-study.

The reader will most probably benefit from having a basic knowledge of mathematics, statistics and digital image analysis.

Lyngby, 30 April 1999

Jørgen Folm-Hansen

Acknowledgments

It is a pleasure for me in this place to thank my colleagues in the image analysis group at IMM for having created a very pleasant and inspiring atmosphere during my 4 years work at the department.

I would like to thank my supervisors Professor Knut Conradsen, Dr. Bjarne Ersbøll and Dr. Jens Michael Carstensen for their interest in the project, valuable advices and professionally inspiring discussions. Without their help this project would never have succeeded or even started.

I would especially like to thank my two office mates Dr. Karsten Hartelius and M.Sc. Søren Falch Jørgensen for their pleasant and encouraging company. Their enthusiasm and light spirit have been an invaluable source of encouragement in times where that was needed. I also had the benefit of borrowing a private bedroom at Karsten Hartelius's apartment, which, because of my change of address to the southern part of the Danish islands, has been a supreme help in the last months of the writing process of this thesis.

In carrying out the work reported here, I have received important assistance from many people beside those mentioned here, and I wish to express my truly gratitude to everybody who has helped one way or the other.

I would like to express my sincere appreciation of the great support and especially patience of my family, who had to put up with me having much less time for them during long periods of heavy work load.

Finally I wish to express my gratitude to my lovely wife Anne-Mette and especially for the love and understanding she was able to give during the writing process, despite my absence both mentally and physically during much of her pregnancy.

Summary

The main subject of the present thesis is different methods for the geometrical and chromatic calibration of cameras in various environments.

For the monochromatic issues of the calibration we present the acquisition of monochrome images, the classic monochrome aberrations and the various sources of non-uniformity of the illumination of the image plane. Only the image deforming aberrations and the non-uniformity of illumination are included in the calibration models. The topics of the pinhole camera model and the extension to the Direct Linear Transform (DLT) are described. It is shown how the DLT can be extended with non-linear models of the common lens aberrations/errors some of them caused by manufacturing defects like decentering and thin prism distortion. The relation between a warping and the non-linear defects are shown. The issue of making a good resampling of an image by using the correct interpolation method is described.

For the chromatic issues of calibration we present the acquisition of colour and multi-spectral images, the chromatic aberrations and the various lens/camera based non-uniformities of the illumination of the image plane. It is described how the monochromatic calibrations are extended to multi channel images. Since accurate colour images require equal (uniform) intensity levels in all channels, the various sources, in both open and closed scenes, for the non-uniform intensities and some corresponding calibration methods are described.

The various possibilities to design calibration targets for both geometrical and chromatic calibration are described. We present some possible systematical errors on the detection of the objects in the calibration targets, if viewed in a non orthogonal angle, if the intensities are uneven or if the image blurring is uneven.

Finally we present the implementation of a complete calibration method for an accurate colour texture measurement device called VMX2000, the calibration for uneven laser sheet illumination in a flow measuring system and the use of automatic detection of calibration targets for a DLT/warping in a 3D PIV system.

Resumé

Hovedemnet i denne afhandling er forskellige metoder til geometrisk og kromatisk kalibrering af kameraer i forskelligartede miljøer.

Med hensyn til de monokromatiske emner indenfor kalibrering er præsenteret optagelsen af monokrome billeder, de monokrome aberrationer og de forskellige kilder for ujævn belysning af billedplanet. Kun de billeddeformerende aberrationer og den ujævne belysning er inkluderet i kalibreringerne. Emnerne omkring pinhole kameraer og udvidelserne til den direkte lineære transformering (DLT) er beskrevet. Det er vist, hvordan en DLT kan udvides med ikke lineære modeller af de almindelige linse aberrationer/fejl, hvoraf nogle af dem skyldes produktionsfejl af typen decentrering og tynd prisme fortegning. Relationerne mellem en warpning og de ikke lineære aberrationer/fejl er blevet vist. Det er blevet beskrevet, hvorledes valget af den rigtige interpolation har betydning for en god resampling af et billede.

Med hensyn til de kromatiske emner indenfor kalibrering er præsenteret optagelse af farve og flerkanals billeder, de kromatiske aberrationer og de forskellige linse/kamera baserede kilder for ikke jævn belysning af billedeplanet. Det er beskrevet, hvorledes de monokromatiske kalibreringer er udvidet til flerkanalsbilleder. Da nøjagtige farvebilleder kræver ens (jævne) intensitetsniveauer i alle kanaler, er det beskrevet, hvorledes de forskellige grunde til ujævnheden er opstået, og hvorledes der kan kalibreres for dem.

Der er beskrevet en masse forskellige muligheder for kalibreringsmål til brug for både de geometriske og de kromatiske kalibreringer. Der er blevet præsenteret en række forskellige systematiske fejl på detekteringen af kalibreringsobjekter, når de ses i en ikke orthogonal vinkel, ujævne intensiteter og ujævn uskarphed i billedet.

Der er blevet præsenteret implementeringen af en komplet kalibreringsmetode til et præcist farve tekstur måleapparat kaldet VMX2000, en kalibrering for ujævne intensiteter i et laser plan placeret i et luftstrøm målesystem og en automatisk detektering af kalibreringsobjekter til DLT/warping i et 3D PIV system.

Contents

Preface iii

Acknowledgments v

Summary vii

Resumé ix

Contents xi

Introduction 1
1.1 Main Issue of the Thesis 1
1.2 Outline of the Thesis 3

Geometrical Calibration 5
2.1 Image Acquisition 5
2.2 Lenses and Aberrations 11
2.2.1 Point Blurring Aberrations
2.2.2 Image Deforming Aberrations
2.2.3 Other Defects
2.3 Pinhole Model 19
2.4 Direct Linear Transform DLT 21

2.5 Nonlinear Camera Calibration 28
2.5.1 Radial-Symmetrical Distortion
2.5.2 Decentering Distortion
2.5.3 Thin Prism Distortion
2.5.4 Nonlinear Minimization
2.6 Warp Based Camera Calibration 40
2.7 Resampling Methods 48
2.7.1 Interpolation Methods
2.8 Summary 61

Chromatic Calibration 62
3.1 Colours 63
3.2 Acquisition 68
3.2.1 Single chip colour cameras
3.2.2 Beam-splitter colour cameras
3.2.3 Multiple shot colour cameras
3.2.4 Scanners
3.2.5 Colour Alignments
3.3 Chromatic Aberrations 77
3.3.1 Transverse chromatic aberration
3.3.2 Longitudinal chromatic aberration
3.4 Full camera calibration 80
3.5 Warp based colour calibration 81
3.6 Uneven intensities 84
3.6.1 Camera and lens based
3.6.2 Scene and light source based
3.6.3 Correction and Calibration
3.7 Summary 109

Calibration Targets 110
4.1 Different Devices 110
4.1.1 Common for Almost all Devices
4.1.2 Monochrome Cameras
4.1.3 Colour Cameras
4.1.4 Scanners and Line-scan Cameras

4.2 Miscellaneous Targets 117
4.3 Detection 126
4.3.1 Manual
4.3.2 Automatic and Semi-Automatic
4.4 Uneven Condition Displacements 151
4.4.1 Uneven Distributed Light
4.4.2 Reflection
4.4.3 Uneven Blurring in the Image Plane
4.5 Perspective Projection Displacement 160
4.5.1 Illustrative Example
4.5.2 A Model for PPD
4.5.3 Parallel Image and Object Plane
4.6 Summary 170

Implementations 171
5.1 Advanced Colour Texture Measurements 171
5.1.1 The Surfix Texturemeter
5.1.2 The Integrated Sphere/VMX2000
5.1.3 The SurfView
5.1.4 Measurements
5.2 Laser Sheet Measurements 204
5.2.1 Measuring Concentration Distribution
5.2.2 Particle Image Velocimetry in 3D
5.3 Summary 213

Conclusion 214

Appendix A 218
Developed Software

Appendix B 222
Line-Jitter

Appendix C 228
Warping Experiments

Appendix D 234
Inverse Nonlinear Distortion Experiment

References 237

Chapter 1

Introduction

The reader will already be aware from the title of this thesis that its overall topics concerns geometrical and chromatic calibrations. By the geometrical calibration we mean the calibration of cameras or other image capturing devices for the image deforming distortions. This includes the spatial misalignment of the channels in multi-spectral images. By chromatic calibration we mean the calibration used to get absolute, repeatable and uniform intensity values in the whole image plane. This, of course, also includes the colour and multi-spectral images.

1.1 Main Issue of the Thesis

The main practical implementation in this project has been the design of suitable calibration methods for obtaining highly accurate colour images from a colour camera combined with a frame-grabber and a computer. This has, of course, also influenced the theoretical contents of this thesis.

There has been made a lot of work on the study of the classical camera calibration methods from photogrammetry. These methods all deal with the geometrical calibration, since they traditionally have been used for measuring e.g. building geometry and not for grey tone/colour measurements. In order to get a better understanding of these calibration models, optics and aberrations of lenses have been studied and compared to the non-linear calibrations. Although it is possible to calibrate a camera with a bad lens to obtain a very high accuracy of measurement, this does not mean that the images are useful for image analysis. The monochromatic and chromatic aberrations also includes point blurring aberrations that blur the image significantly, the use of high quality lenses can therefore be needed despite the use of a very good calibration.

To make an accurate colour image, where the colours in the image can be measured in some standard colour coordinate system, it is needed not only to have a good alignments of the channels but also to have a uniform intensity distribution on the image plane. This has shown to be quite a difficult task since the lens, camera, illumination and scene setup all will influence on the intensity distribution in the image. There has been made a lot of work on solving these problems and some apparently new insight on this subject has been developed.

The studied and developed methods have been implemented in three cases.

- An accurate colour texture measurement device called VMX2000.
- Measuring concentration distribution of injected gas in a swirling bulk flow using a laser-sheet.
- The initial measurement for a prototype of a 3D (stereo) Particle Image Velocimetry (PIV) system.

1.2 Outline of the Thesis

It has not always been easy to divide the material into chapters and sections, since much of it is closely related. So some of the dividing into chapters should be considered a compromise to having one big chapter.

The first of the subsequent chapters (**chapter 2**) gives a description of the aspects of calibration that mainly are related to monochrome images (ignoring the chromatic issues), this includes the lens aberrations, uneven radiance on the image plane, classical camera calibration, warp based geometrical calibration and resampling/interpolation issues.

Chapter 3 is dedicated to the more chromatic (colour) related calibrations, this is among others the chromatic aberrations. It is described how various types of colour cameras work and what the advantages / disadvantages are. The relations to the classical camera calibration and warping are described. Even though it can be interesting to calibrate for uneven intensities with monochrome cameras this subject has been placed in this chapter. It is described how the lens/camera errors, combined with the scene based uneven illumination, can influence the intensities, and how to make various calibrations for this effect.

Chapter 4 gives a description of various calibration targets that can be used for both the geometrical and chromatic calibrations. We also present methods for automatic detection of simple calibration targets. It is shown how unevenness in the intensities or the blurring can make systematical subpixel displacements of the calibration objects. Finally we present a simple model for calculating the perspective projection displacement (PPD) that can be used for evaluation of a practical setup.

Chapter 5 describes three practical implementations, made within this project, of some of the calibration methods described in the previous chapters.

Appendix A contains a brief description of the developed software during this project.

Appendix B describes a practical experiment with line-jittering.

Appendix C describes a practical experiment with warping used to illustrate the relation between a perspective transformation and a polynomial warping.

Appendix D shows a practical experiment when using the non-linear distortions as estimates of the inverse distortions.

Chapter 2

Geometrical Calibration

This chapter will describe the aspects of monochrome images and the related calibrations. This includes the acquisition, interlacing, monochromatic aberrations, the pinhole model and the related extensions for modelling the distortions, the classical photogrammetry models and warping. Aspects about obtaining correct/uniform intensities across the image plane and chromatic aberrations, are discussed in the next chapter on chromatic calibration.

2.1 Image Acquisition

Digital images can come from a variety of devices. In this thesis only scanners and especially camera based acquisition will be discussed. A monochrome camera is made of a ray collecting device (like a lens) and some sort of normally planar 2D-array of sensors (chip) to digitalize the image. A variant is the line-scan cameras. They also use a ray collecting device, but they only have a 1D-array of sensors, and either the camera or the object must be moved in order to make a 2D image. A scanner normally uses a 1D-array of sensors that are moved at a very close range over (or under) the object to be scanned.

Digital images are as a minimum specified by their size and dynamics, the size is the number of spatial element (pixels) and the dynamics is the maximum number of grey tones that can be represented. The dynamics is usually given in bits, where e.g. an 8 bit image maximally can hold 256 grey tones.

Micro-Technology

The most common (1998) chip type for digital image capturing is the CCD (charge coupled device) chip where each sensor element (pixel) is made of an electron well. The CCD technology gives a high fill factor, up to 95%, meaning that the active photon counting area is 95% of the on-chip pixel area. The alternative to the CCD chips are the CMOS based chips which are rapidly coming into the market, instead of an electron well they use a small photo diode for each sensor. They are cheaper to produce and the layout of the chip makes it possible to add simple image processing hardware into the chip. The active sensor (the photo diode) needs some surrounding circuitry to function which, unfortunately, causes the fill factor to be very small, only about 30%. The small fill factor of the CMOS chips is the largest disadvantage of this technology, since the number of photons to hit the chip must be larger than with a CCD chip to get the same intensity and thereby the same signal/noise ratio. A CMOS camera therefore normally requires stronger light than the CCD chip and as described later in this chapter a larger aperture on the lens gives more aberrations. There are actual plans of placing micro-lenses onto each of the pixels in the CMOS chip to focus the light directly on the photo diode. For further details see e.g. the Sound Vision Incorporated Internet homepage at *http://www.soundvisioninc.com/.* In this project there has only been used CCD cameras, but if using CMOS based cameras the low fill factor must be taken into consideration, if the focus of the lens is better than the size of a pixel the low fill factor could e.g. cause cut-lines effect since the “picture” on a single pixel only is captured in a small region, this can be compared to making a heavy undersampling of a digital image without pre-blurring.

Scanners

Scanners differ significantly from cameras in the sense that they usually are used for planar objects, like photos, negatives and textiles, and must have the objects at a very close range. A scanner uses an array of sensors that moves across the object where the array not necessarily can cover the entire image. The most common is to use a full width single line array that is moved perpendicular to the line in order to capture a complete 2D-image. Some variants use a small 1D-array that is moved vertically over (or under) the objects line by line, this can cause some reflection problems with glossy objects that gives an uneven light across the sensor. In this project there has been observed visually large vertical stripes on a very expensive scanner (Eskofot 2450) when grabbing negatives/positives caused by the illumination at the top and bottom of the small 1D sensor to be different.

With most scanners there is no lens in traditional camera sense and the models for classic lens aberration cannot be used. The line-array is typically very straight and since it is factory mounted it is also very perpendicular to the moving direction. Nowadays (1998) traditional scanners use a single line CCD chip which captures image data through a lens-and-mirror optical system. In such a system the optics can become misaligned giving manufacture specific displacements errors. These displacements can be in both directions. The new type of scanners that are coming rapidly into the scanner market is CIS scanners. A CIS (contact image sensor) scanner replaces the optical technique with a single row of sensors, which is positioned about 1-2 mm below the object to be scanned. Illumination is provided by a closely packed row of red, blue, and green LEDs to create white light. Because the CIS uses far less power, are simpler, smaller and cheaper to produce than the CCD scanners they (or some similar product) will most likely control most of the market in the future.

Noise

The sensor elements work by counting the number of arrived photons inside the active sensor area, the number of arrived photons at a specific sensor is a stochastic variable but since the number of arrived photons normally are very high this should be a neglect able noise source in most systems. But some of the counted photons are thermally generated within the sensor (black current) and this is normally a significant noise source for digital cameras. This noise can be reduced by cooling the sensor array and many non-consumer cameras come with built-in cooling for that purpose.

Since the largest noise source is thermally generated the signal/noise ratio is very dependent on the exposure time and the amount of illumination on the scene/object. As exposure time increases the accumulation of thermal noise is also increased. But to get the best signal (without over exposure) the exposure time must be increased until the full dynamics almost is reached. The best is always the highest possible illumination, a low lens aperture and a short exposure time. The aperture should only be enlarged if the possible lens aberrations can be accepted. Many digital cameras have an adjustable gain that can move the range of dynamics electronically, this can be used for very weak illuminations or exposure times. A low gain should only be used if it is strictly necessary since this also increases the noise, the thermal noise can easily be observed if capturing images with a very low exposure time (e.g. 1/10000) and using the gain to adjust the intensities of the image.

In some applications sequences of images of moving objects are captured. In such applications there can be a requirement for very low exposure times to decrease blurring from the movement. In order to improve the signal/noise ratio in such sequences cooled cameras or a very strong light must be used. Very strong light can come from a strobe synchronized with the camera in order to reduce the power consumption and the heat generation.

The thermal noise can, if the scene/object position and illumination is constant and time allows it, be used in a constructive way. From [Conradsen84] can be seen that if averaging n observations of equally $N(\mu,\Sigma)$ distributed stochastic

variables, the average of the observations will follow a $N(\mu, \Sigma/n)$ distribution. If the noise is so large that each pixel can jump between some integer values the average of e.g. 25 images will have a numerically 80 % lower standard deviation. If the images are kept in e.g 8 bits this will give significantly better images. If the images are averaged in floating point precision the dynamic of an image can actually be increased, since the mean value will be closer to the true pixel value (non integer). This would in most cases not have been possible if the camera had no thermal noise, since the pixels then would stay at their integer value. This way of enlarging the dynamics has been used very much in this project, especially when working with textures where the dynamics only was about 10 - 20 grey tones in the region of interest in the image (this is described in more detail in chapter 5).

Blooming

When a sensor element (pixel) receives a certain number of photons it will be filled and if more photons are received they must be absorbed by the surroundings. With some sensor arrays these extra photons will affect the neighbouring sensors causing a highlighted area to spread out, this phenomena is called *blooming*. Over exposure must therefore in general be avoided, unless the technical specifications specifically state that the chip is blooming free. This is a very common error and has been observed with all cameras that have been over exposed in this project.

Interlacing

When the first TV signals were designed, the bandwidth was not high enough for the wanted resolution and frame rate, and a technique called interlacing was used to solve the problem. The idea is to double the frame rate by sending only the even lines and the odd lines in turns (half-frames or fields). In the European PAL/CCIR systems this means that there is sent 50 fields/sec which can be collected using a frame grabber to full 25 frames/sec. This looks very good on e.g. a television but when using a video camera and a frame grabber the small

difference of time between the fields (odd and even lines) gives very poor results when using it for moving scenes/objects. Unfortunately, the interlacing has been applied to most cameras for video signals, but the non-interlaced *progressive scan* cameras are becoming more and more common.

The effect of interlacing can be seen in figure 2.1 (a) which is a small extract from a surveillance video camera (3 fr/s), the bicycle moving horizontally is shown in odd and even fields (b) and the linear interpolated odd and even fields in (c).

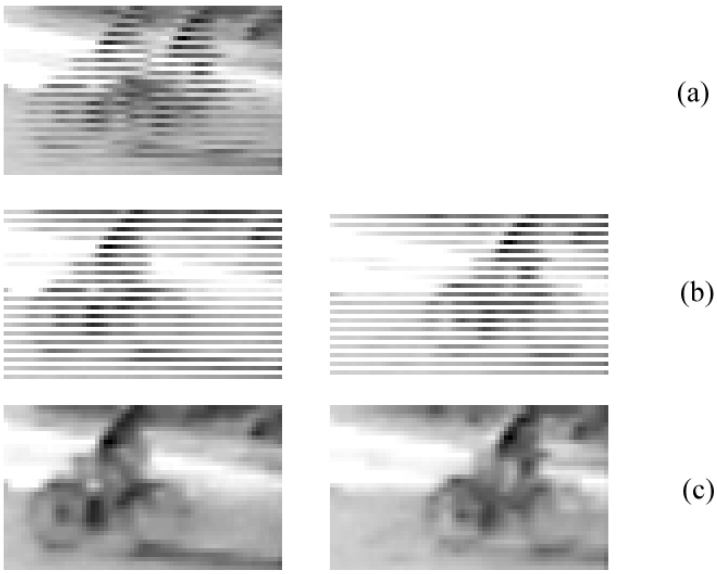


Figure 2.1 Bicycle moving horizontal from surveillance camera (a) interlaced (b) odd and even fields (c) interpolated odd and even fields.

The problem regarding handling interlaced sequences is in general not simple. It can be solved for objects moving strictly horizontally by shifting e.g. the odd lines by some subpixel interpolation method. When objects are moving in any direction the problem is very complex, if e.g. a bicycle is moving away from the camera the small difference of time will cause the bicycle to have different size in the two fields.

There has been made considerable work with the problems about interlacing in this project. The two types that are described later in this thesis are three cases for the Danish Federal Police about interlaced surveillance cameras and a case

with image based measurement on a corner-fired boiler made in collaboration with AFM at DTU. In the collaboration with AFM the work with interlacing has been used for getting 50 frames/sec. in order to get a better representation of the high speed smoke (6 m/s) that moves about half the width of the image in 1/25 sec.

2.2 Lenses and Aberrations

In this section some general aspects about lenses that are useful for the next sections and chapters are described, this includes a description of the five monochromatic aberrations divided into blurring, distortion aberrations.

A lens for camera use is normally built up of several elements to reduce the aberrations. An element that collects a beam of parallel rays is called positive and an element that spreads is called negative. A single element or elements cemented together are called a group. Each new element gives the designer of the lens a few more degrees of freedom to reduce the aberrations but since each element can reflect a little light, too many elements could cause flare (see further down). In figure 2.2 is an illustration of a 3 group lens with 4 elements. Normally, the aperture is made of a diaphragm placed between the elements inside the lens (not illustrated).

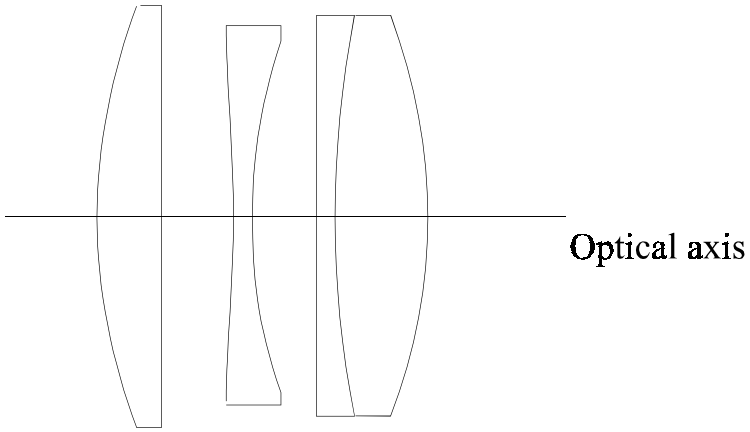


Figure 2.2 Illustration of the Tessar lens. From [Dalsgaard89].

2.2.1 Point Blurring Aberrations

Spherical Aberration (SA)

The first of the classic aberrations is called Spherical Aberration and is caused by the light not being focussed at the same point, dependent on if it hits the lens at the top or the center, see figure 2.3. This is the only of the 5 classic aberrations that can occur on the optical axis.

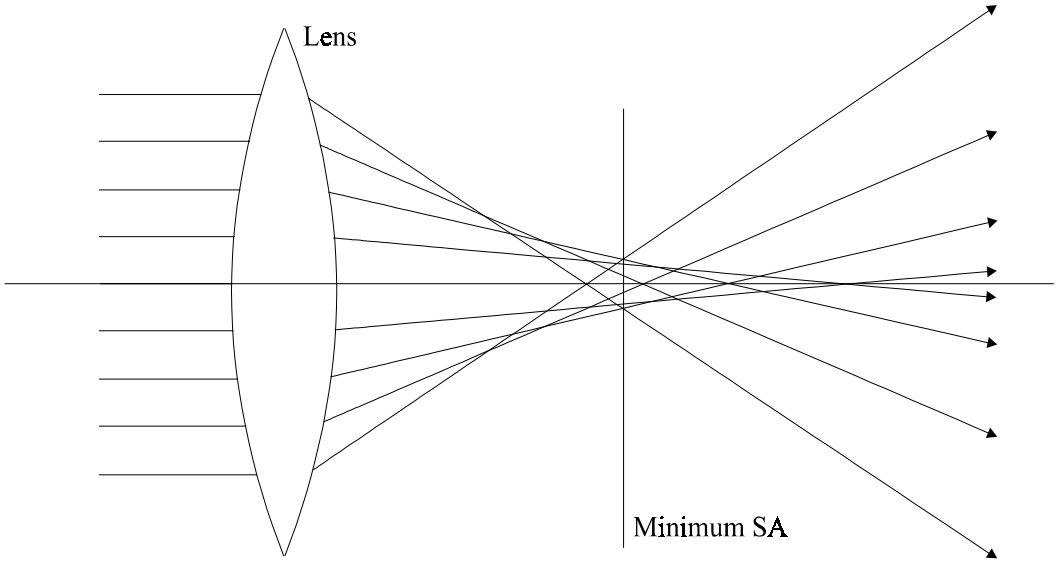


Figure 2.3 Spherical Aberration (SA).

The SA can be reduced significantly by combining a positive lens element with a negative with same numerical SA value, this can be done as in figure 2.4 where the SA can be reduced to less than 1% [Dalsgaard89]. The SA can also be reduced by using a smaller aperture.

Figure 2.4 (a) normal positive lens element (b) positive element cemented with suitable negative to reduce SA.

Coma Aberration

The second classic aberration is called Coma and is caused by different magnification of a point if the rays are passing through the lens near the optical center or near the edge. This is illustrated in figure 2.5. As in the figure the effect of the Coma is that a point is depicted as a series of small circles, resulting in the comet-like appearance that has given the aberration the name Coma.

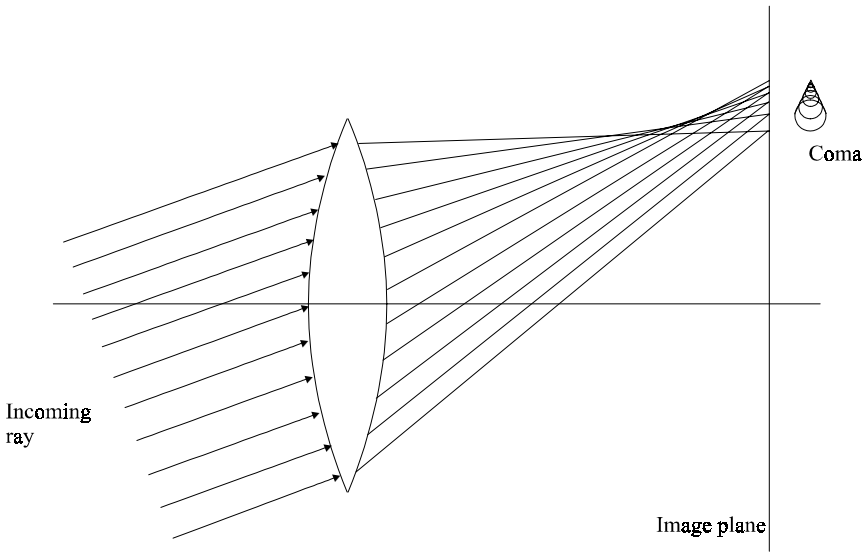


Figure 2.5 Coma Aberration.

In order to design a lens without Coma aberration it must satisfy the Abbe Sine Condition (see e.g. [Dalsgaard89] for a good description of the Sine Condition) and be free of Spherical aberration e.g. by using a combination of a positive and negative lens like in figure 2.4 [Dalsgaard89]. A lens fulfilling the Abbe Sine Condition is called an aplanatic lens.

Note: Ernst Abbe (1840-1905) was a German professor from the University of Jena who in cooperation with Carl Zeiss invented the apochromatic lens system and discovered the *Abbe Sine Condition* that gives the conditions for a lens to form an image without Coma and Spherical aberration.

Astigmatism

The last of the aberrations that causes point blurring is Astigmatism that as Coma only affects off-axis points. The effect of astigmatism is that off-axis points are blurred in their radial or tangential direction. Focussing can only reduce one at the expense of the other but cannot bring both radial and tangential into focus at the same time. Astigmatism can be a considerable error with a two lens system that is corrected for SA and coma. It can be reduced by using an smaller aperture or adding one more suitable lens element to the lens system [Dalsgaard89].

Note on deblurring

In classical camera calibration it is only the distortions and not blurring that is taken into consideration but it is possible to estimate the Point Spread Function (PSF) by e.g. grabbing an image of a white spot on a dark background and then use a method like a Wiener Filter to deblur the image [Sonka93]. Unfortunately, deblurring has some artifacts like ringing and since the blurring is not normally constant across the image plane (unless the image has been blurred from e.g. a moving camera) the quality can be questioned. Deblurring is not considered further in this project and no actual experiments have been performed.

2.2.2 Image Deforming Aberrations

Field Curvature

When a lens system is designed for reduced astigmatism the plane of focus will follow a Petzvals surface (that is curved) giving a Field Curvature aberration [Dalsgaard89]. The effect of this error is that if an object is in focus at the center it will be out of focus near the edge and visa versa, consequently, the whole image plane cannot be in focus at the same time. The effect of the Field Curvature aberration can be reduced by using a smaller aperture or by designing a complex lens system where a high degree of freedom gives the possibility of controlling the astigmatism and thereby reducing the curvature of the Petzvals surface [Dalsgaard89]. If it is technically possible in the application a suitable curved image plane could also eliminate the effect of this error.

Distortion

This last of the five monochromatic aberrations called distortion does not cause any blurring, a point will be depicted as a point but gives a displacement of an off-axis point. It is caused by the magnification being dependent on the angle to the optical axis. It can be reduced by a suitable design of the lens system. If the distortion causes increased magnification when moving away from the optical center it is called pincushion and if the magnification decreases it is called barrel distortion, this is illustrated in figure 2.12. This is the only of the five monochromatic aberration where it is possible to make a classic camera calibration to simulate a perfect lens. If an image is free of distortion it is called rectilinear, all straight lines are depicted as straight lines.

2.2.3 Other Defects

Diffraction

The limit of how sharp an image a lens can present, is given by the diffraction caused by the light beam spread out after passing the lens (aperture). Figure 2.6 is an illustration of the diffraction phenomena where a light beam hits a surface with a hole (like a pinhole camera or a diaphragm).

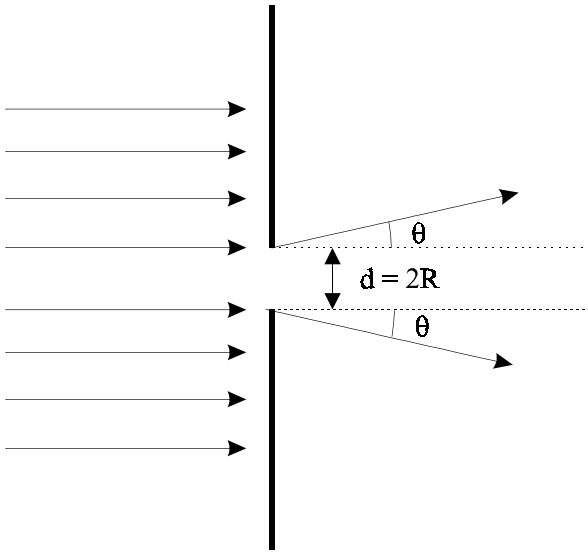


Figure 2.6 Diffraction of a planar light beam through hole with diameter d.

The diversion angle θ is approximately given in (2.1) where λ is the wavelength of the light (visible light about 400 to 750 nm).

$$\theta \approx 1.22 \cdot \frac{\lambda}{d} \tag{2.1}$$

The limit of the size of a focus point is then given by d_A in (2.2) which is the size of the Airy disc from the diffraction pattern, see [Dalsgaard89] for a

description of the Airy disc. *f* is the focal length of the lens. This should be compared to the CCD-chip pixel size.

$$d_A = 2 \cdot \theta \cdot f = 2.44 \cdot \lambda \cdot \frac{f}{d} \quad (2.2)$$

Using a 1/3" CCD-chip (pixel size about 6 µm) with a 25 mm lens, 4 mm aperture and 555nm light gives an Airy disc of 8,5 µm which is a little larger than a pixel. This illustrates that diffraction can be a considerable blurring error in normal vision systems.

As described earlier in this chapter most aberration can be reduced by using a smaller aperture but from (2.2) it can be seen that the smaller the aperture the larger is the diffraction. When a manufacturer denotes a lens “perfect” or error free he normally means diffraction limited, that is, the aberrations are smaller than the diffraction error.

The diffraction is also the reason why pinhole cameras (see figure 2.8) give smeared images and therefore are very rare in vision systems. The smaller the hole, the larger the blurring from diffraction and the larger the hole the more blurring from geometric views. It can be mentioned that when optimized for 555nm light the diameter (in millimetre) of the hole in a pinhole camera should be 0.036*sqrt(f) where f is the distance between hole and image plane in millimetres [Jacobson97].

Flare

Since each element in a lens can reflect a little light multiple elements could cause light being reflected forward and backward inside the lens, these reflections can cause a defect called flare. The more elements in a lens the higher the risk of flare. The internal material used in the housing of the lens can also cause troubles if it is too reflective. Flare is typically seen as bright spots

(often shaped as the diaphragm inside the lens) when a bright light is in the picture. It can be reduced by a good design of the lens, the use of non-reflective material in the housing and the use of good glass. Flare is likely to happen if the light source in a scene can light directly (or indirectly e.g. through a mirror) at the front glass of the lens. This has been observed when working with a camera looking at a hole in an illuminated integrated sphere causing a slight design change of the system, see chapter 5. If flare appears in a machine vision system another lens (probable more expensive) must be used or if possible the direct light must be shielded. Some times a lens shade at the front can keep the unwanted light from striking the lens and causing flare, but it should be sized to the particular lens to avoid vignetting.

Chromatic effects on monochrome images

Since white light consists of all visible wavelengths, a lens with chromatic aberrations (see chapter 3) will cause blurring when used with a monochrome camera. This can be a significant blur and the use of an achromatic lens or spectrally narrow light must be considered, if possible.

Decentering and Thin Prism Distortion

When assembling a lens it is of course possible to misalign the elements. This can either be an alignment error of the optical axis or a tilt of the element as illustrated in figure 2.7. A decentering of a lens element causes a distortion with both a radial and a tangential component and is a commonly modelled distortion in a camera calibration. For more details see page 32. A tilt of a lens element (or CCD chip) causes a Thin Prism distortion which also have both a radial and a tangential component, the name “Thin Prism Distortion” comes from the fact that the tilt can be modelled by an adjunction of a thin prism in the optical system. The Thin Prism distortion is not commonly used in camera calibration, see page 35 for more details.

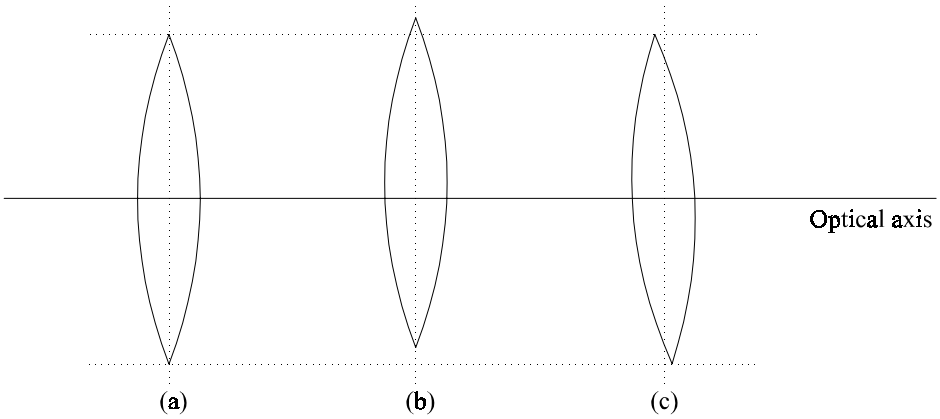


Figure 2.7 Lens elements with various misalignments
(a) perfectly mounted (b) decentered (c) tilted.

2.3 Pinhole Model

The pinhole camera illustrated in figure 2.8 represents the ideal camera, free from any errors like distortion. It can be made in practice using a box with a white sheet as image plane and a small hole in the middle of the opposite side of the “image plane”. Such a small pinhole can be used as a camera and the main advantage is that it has unlimited depth of field, the disadvantage is that the light beams that go through the small hole will cause diffraction and the image will always be quite blurred, pinhole cameras are therefore only used for very special applications.

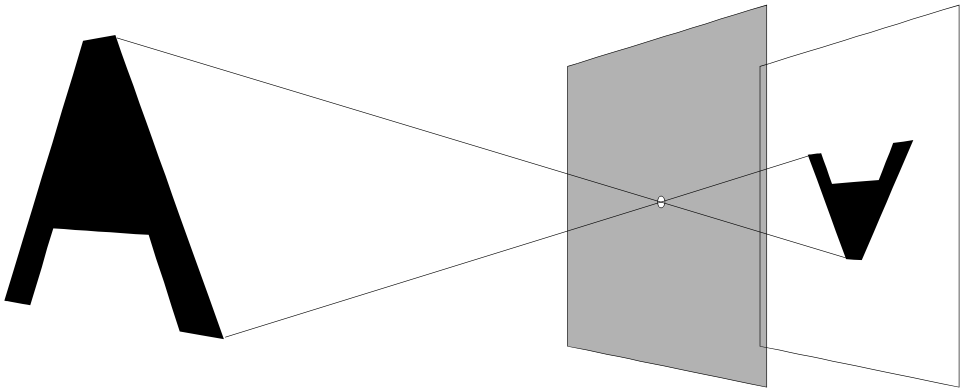


Figure 2.8 The pinhole camera. Only the direct rays from object to image plane come through the small hole.

The pinhole camera is, because of its simplicity and perspective similarity to a lens based camera, used as a mathematical model for the perfect camera. The parameters used for the pinhole model which is the foundation for the further camera models are illustrated in figure 2.9.

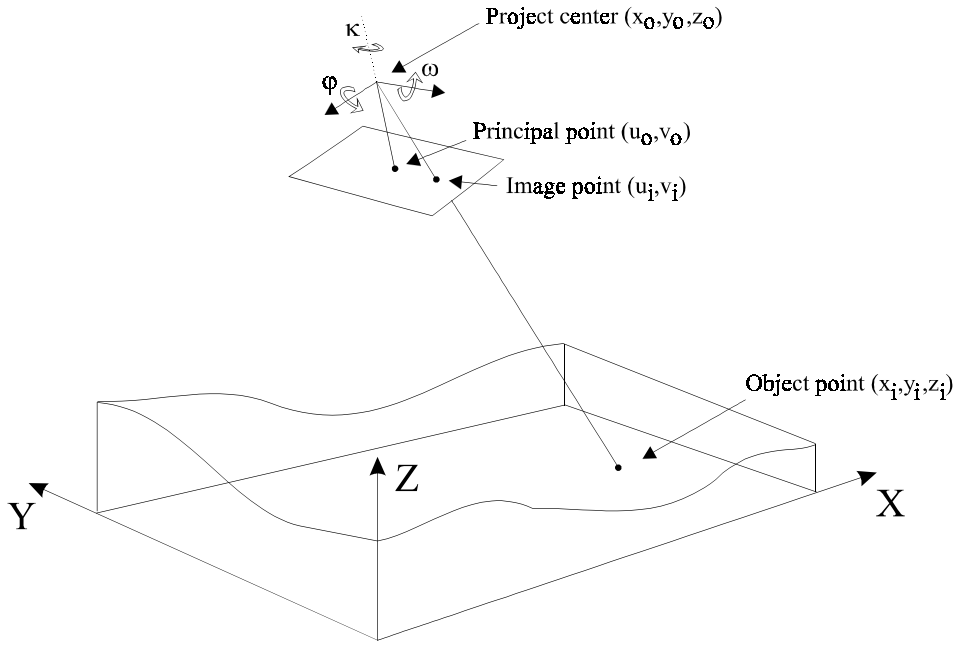


Figure 2.9 The pinhole camera model with parameters.

In (2.3) are the formulas for the relation between world coordinates p_i (x_i,y_i,z_i) and image coordinates q_i (u_i,v_i). for a pinhole model in homogeneous coordinates, it is made of a translation, a rotation and a perspective transformation.

$$\begin{bmatrix} w_i u_i \\ w_i v_i \\ w_i \end{bmatrix} = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & -x_0 \\ 0 & 1 & 0 & -y_0 \\ 0 & 0 & 1 & -z_0 \end{bmatrix} \cdot \begin{bmatrix} x_i \\ y_i \\ z_i \\ 1 \end{bmatrix} \qquad (2.3)$$

$$or$$

$$q_i = F \cdot M \cdot T \cdot p_i$$

The rotation matrix M actually only consists of three parameters which traditionally have been the Euler angles: first ω around the X axes then φ around the Y axes and finally κ around the Z axes (see figure 2.9), where the order of the rotations is important. (2.4) shows the rotation matrix M with the three angles.

$$M = \begin{bmatrix} \cos(\varphi)\cos(\kappa) & \sin(\omega)\sin(\varphi)\cos(\kappa) - \cos(\omega)\sin(\kappa) & \cos(\omega)\sin(\varphi)\cos(\kappa) + \sin(\omega)\sin(\kappa) \\ \cos(\varphi)\sin(\kappa) & \sin(\omega)\sin(\varphi)\sin(\kappa) + \cos(\omega)\cos(\kappa) & \cos(\omega)\sin(\varphi)\sin(\kappa) - \sin(\omega)\cos(\kappa) \\ -\sin(\varphi) & \sin(\omega)\cos(\varphi) & \cos(\omega)\cos(\varphi) \end{bmatrix} \quad (2.4)$$

The rotation matrix M is orthogonal, the inverse matrix equals the transposed matrix. For details on the rotation matrix see [Jacobi84].

2.4 Direct Linear Transform DLT

The direct linear transform (DLT) can be regarded as an extension of the pinhole model that includes a shift of the image origin and a linear distortion in the image plane. It is normally divided into the 3x4 DLT for a full 3D transformation and the 3x3 DLT for planar word coordinates (z=0), both transforms will be discussed in this section.

(2.5) shows the 3x4 DLT (homogeneous coordinates) between world coordinates p_i (x_i,y_i,z_i) and image coordinates q_i (u_i,v_i).

$$\begin{bmatrix} w_i u_i \\ w_i v_i \\ w_i \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix} \cdot \begin{bmatrix} x_i \\ y_i \\ z_i \\ 1 \end{bmatrix} \quad (2.5)$$

$$or$$

$$q_i = A \cdot p_i$$

The linear distortion in the DLT consists of a scale difference and lack of orthogonality between the image axes [Melen94]. The linear distortion from image coordinates (u_i,v_i) to corrected image coordinates (u_i´, v_i´) can be represented by the equation in (2.6).

$$\begin{bmatrix} u_i' \\ v_i' \\ 1 \end{bmatrix} = \begin{bmatrix} 1+b_1 & b_2 & 0 \\ b_2 & 1-b_1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & -u_0 \\ 0 & 1 & -v_0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} \qquad (2.6)$$

$$or$$

$$q_i' = B \cdot V \cdot q_i$$

Where b_1 and b_2 are the linear distortion coefficients and (u_0,v_0) is the principal point.

When combined with the pinhole model the **A** matrix from (2.5) can be represented by equation (2.7).

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix} = \lambda \cdot \begin{bmatrix} 1 & 0 & -u_0 \\ 0 & 1 & -v_0 \\ 0 & 0 & 1 \end{bmatrix}^{-1} \cdot \begin{bmatrix} 1+b_1 & b_2 & 0 \\ b_2 & 1-b_1 & 0 \\ 0 & 0 & 1 \end{bmatrix}^{-1} \cdot \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & -x_0 \\ 0 & 1 & 0 & -y_0 \\ 0 & 0 & 1 & -z_0 \end{bmatrix}$$

$$or$$

$$A = \lambda \cdot V^{-1} \cdot B^{-1} \cdot F \cdot M \cdot T \qquad (2.7)$$

Where $\lambda \neq 0$ is an arbitrary overall scaling factor, its value will only effect w_i which is divided out to find u_i and v_i, anyway.

The DLT matrix A can of course easily be calculated from the camera parameters. The decomposition of an estimated DLT matrix (**A**) to the camera parameters is quite complicated and was first solved completely in 1987. Notice that 12 parameters need to be estimated since the rotation matrix only has 3 parameters. Non-iterative and fast methods for this decomposition can be found

in many references and it will not be discussed further in this thesis, good references are [Melen94] and [Knudsen98]. The decomposition of the 3x4 DLT matrix described in [Melen94] has been software implemented within this Ph.D.-project and has been used in the work done in this project as well as other projects at the department IMM with non planar calibration targets.

The parameters are traditionally divided into the exterior (translation and rotation) and the intrinsic parameters which with the DLT are f, u_0, v_0, b_1 and b_2. The point about exterior and intrinsic parameters is that if a camera is moved from one image to another the intrinsic parameters are constant and the exterior gives the position of the camera for each position.

When calibrating a camera using a planar calibration target from various positions, the planar target is assumed to be the Z=0 plane and the DLT matrix for each camera position is reduced to a 3x3 matrix as seen in equation (2.8).

$$\begin{bmatrix} w_i u_i \\ w_i v_i \\ w_i \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \cdot \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} \qquad (2.8)$$

$$or$$

$$q_i = A_{3x3} \cdot p_i$$

The representation of the A_{3x3} as an extended pinhole is given in equation (2.9), the only difference is actually the new translation matrix T_{3x3}.

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \lambda \cdot \begin{bmatrix} 1 & 0 & -u_0 \\ 0 & 1 & -v_0 \\ 0 & 0 & 1 \end{bmatrix}^{-1} \cdot \begin{bmatrix} 1+b_1 & b_2 & 0 \\ b_2 & 1-b_1 & 0 \\ 0 & 0 & 1 \end{bmatrix}^{-1} \cdot \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & -x_0 \\ 0 & 1 & -y_0 \\ 0 & 0 & -z_0 \end{bmatrix}$$

$$or$$

$$A_{3x3} = \lambda \cdot V^{-1} \cdot B^{-1} \cdot F \cdot M \cdot T_{3x3} \qquad (2.9)$$

The decomposition of the 3x3 DLT matrix cannot be solved completely without the prior knowledge of at least 3 parameters, because the A_{3x3} matrix only gives 9 degrees of freedom and 12 (including λ) are needed. [Melen94] gives three different approaches with assumptions about a subset of parameters (b_1, b_2, f , u_0 or v_0) and since the DLT mostly is used as a starting guess for more advanced method this is normally enough. If the image plane and the planar object are parallel, then none of the methods can decompose the focal length and the principal point. One of the methods for decomposition of the 3x3 DLT matrix in [Melen94] has been software implemented within this Ph.D.-project and has successfully been used in work done in this project as well as other projects on IMM where the use of planar calibration targets has been used.

When making a camera calibration the DLT matrix and its decomposition(s) are very useful tools for estimating the camera parameters, which again can be used as initial guesses for more advanced methods. After obtaining a set of image coordinates with corresponding world coordinates (using calibration targets) the next step is to estimate the DLT matrix from the coordinate set(s). This is a straight forward method that can be divided into estimating the full 3x4 DLT matrix or the 3x3 DLT matrix.

In the case of a coordinate set where all the world coordinates are not in one plane, the 3x4 DLT matrix must be used. Multiplying the DLT matrix in (2.5) with the world coordinates $\mathbf{p_i}$ and dividing the two upper rows by w_i the two equations in (2.10) (one for u_i and one for v_i) are obtained.

$$\begin{aligned} a_{11}x_i + a_{12}y_i + a_{13}z_i + a_{14} - a_{31}x_iu_i - a_{32}y_iu_i - a_{33}z_iu_i - a_{34}u_i = 0 \\ a_{21}x_i + a_{22}y_i + a_{23}z_i + a_{24} - a_{31}x_iv_i - a_{32}y_iv_i - a_{33}z_iv_i - a_{34}v_i = 0 \end{aligned} \quad (2.10)$$

This can be written in matrix format.

$$\begin{bmatrix} x_i & y_i & z_i & 1 & 0 & 0 & 0 & 0 & -x_iu_i & -y_iu_i & -z_iu_i & -u_i \\ 0 & 0 & 0 & 0 & x_i & y_i & z_i & 1 & -x_iv_i & -y_iv_i & -z_iv_i & -v_i \end{bmatrix} \cdot \begin{bmatrix} a_{11} \\ a_{12} \\ a_{13} \\ a_{14} \\ a_{21} \\ a_{22} \\ a_{23} \\ a_{24} \\ a_{31} \\ a_{32} \\ a_{33} \\ a_{34} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

or

$$L' \cdot a' = 0 \quad (2.11)$$

The equation in (2.11) can be used for a least square fit when estimating the DLT matrix *A* where each point gives two rows in the L' matrix. Unfortunately the matrix is singular and an additional constraint must be applied, the most common is $a_{34}=1$ but since it will fail if the correct a_{34} is close to zero another (singularity free) constraint $a_{31}^2 + a_{32}^2 + a_{33}^2 = 1$ is also often used [Melen94]. In the work done in this project the constraint $a_{34}=1$ has been used overall and has not given any problems, the singularity with the $a_{34}=1$ constraint will only occur when the equation in (2.12) is true and the center of perspective is located in one of the three planes defined by the axis of the world coordinate system. This can easily be seen from the rotation matrix in (2.4) and the extended pinhole model in (2.7) when considering that λ=0 cannot be true since it would

imply that the DLT matrix **A** is zero.

$$-\sin(\varphi)x_0 + \sin(\omega)\cos(\varphi)y_0 + \cos(\omega)\cos(\varphi)z_0 = 0 \tag{2.12}$$

If using the a_{34}=1 constraint the equation from (2.10) is turned into (2.13).

$$\begin{aligned} a_{11}x_i + a_{12}y_i + a_{13}z_i + a_{14} - a_{31}x_iu_i - a_{32}y_iu_i - a_{33}z_iu_i &= u_i \\ a_{21}x_i + a_{22}y_i + a_{23}z_i + a_{24} - a_{31}x_iv_i - a_{32}y_iv_i - a_{33}z_iv_i &= v_i \end{aligned} \tag{2.13}$$

In (2.14) is shown the equation containing the N observed image coordinates and the corresponding “known” world coordinates in a 11x2*N matrix **L** and a 2*N vector **l**.

$$\begin{bmatrix} & & & & & & & & & & \\ & & & & & & & & & & \\ & & & & & & & & & & \\ x_i & y_i & z_i & 1 & 0 & 0 & 0 & 0 & -x_iu_i & -y_iu_i & -z_iu_i \\ 0 & 0 & 0 & 0 & x_i & y_i & z_i & 1 & -x_iv_i & -y_iv_i & -z_iv_i \\ & & & & & & & & & & \\ & & & & & & & & & & \\ & & & & & & & & & & \end{bmatrix} \cdot \begin{bmatrix} a_{11} \\ a_{12} \\ a_{13} \\ a_{14} \\ a_{21} \\ a_{22} \\ a_{23} \\ a_{24} \\ a_{31} \\ a_{32} \\ a_{33} \end{bmatrix} = \begin{bmatrix} \cdot \\ \cdot \\ \cdot \\ u_i \\ v_i \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} \tag{2.14}$$

or

$$L \cdot a = l$$

The least squares fit is made by calculating the expression in (2.15).

$$a = (L^T L)^{-1} L^T l \tag{2.15}$$

To solve (2.15) at least six points are needed to have full rank eleven, but no geometric distribution of six points will guarantee full rank for all camera positions therefore at least seven points should be used [Melen94]. In all cases the points must not be in one plane (in world coordinates) since it will make the system singular.

The system to estimate the 3x3 DLT matrix is similar to the above description of the estimation of the 3x4 DLT matrix and the estimation is shown in (2.16) with the constraint that $a_{33}=1$.

$$\begin{bmatrix} & & & & & & & \\ & & & & \vdots & & & \\ & & & & & & & \\ x_i & y_i & 1 & 0 & 0 & 0 & -x_i u_i & -y_i u_i \\ 0 & 0 & 0 & x_i & y_i & 1 & -x_i v_i & -y_i v_i \\ & & & & \vdots & & & \\ & & & & & & & \end{bmatrix} \cdot \begin{bmatrix} a_{11} \\ a_{12} \\ a_{13} \\ a_{21} \\ a_{22} \\ a_{23} \\ a_{31} \\ a_{32} \end{bmatrix} = \begin{bmatrix} \vdots \\ u_i \\ v_i \\ \vdots \end{bmatrix} \tag{2.16}$$

or

$$L_{3x3} \cdot a_{3x3} = l_{3x3}$$

Where the least squares fit is given in (2.17).

$$a_{3x3} = (L_{3x3}{}^{T} L_{3x3})^{-1} L_{3x3}{}^{T} l_{3x3} \tag{2.17}$$

2.5 Nonlinear Camera Calibration

In classical photogrammetry the distortion has normally been modelled by two nonlinear distortion components called the radial and tangential distortion instead of the linear distortion from the DLT. These two distortions are depicted in figure 2.11 to give a visual explanation of the names. The radial and tangential distortion can also be used as an addition to the linear distortion [Melen94] but this will cause an increased amount of intrinsic parameters to be estimated and should be unnecessary in most applications since the additional linear distortion only can model non-orthogonal image axes which should be a very small and insignificant error [Heikkilä97b]. There are various different models of the nonlinear distortion components and the model described in this section is the one used and software implemented in this project.

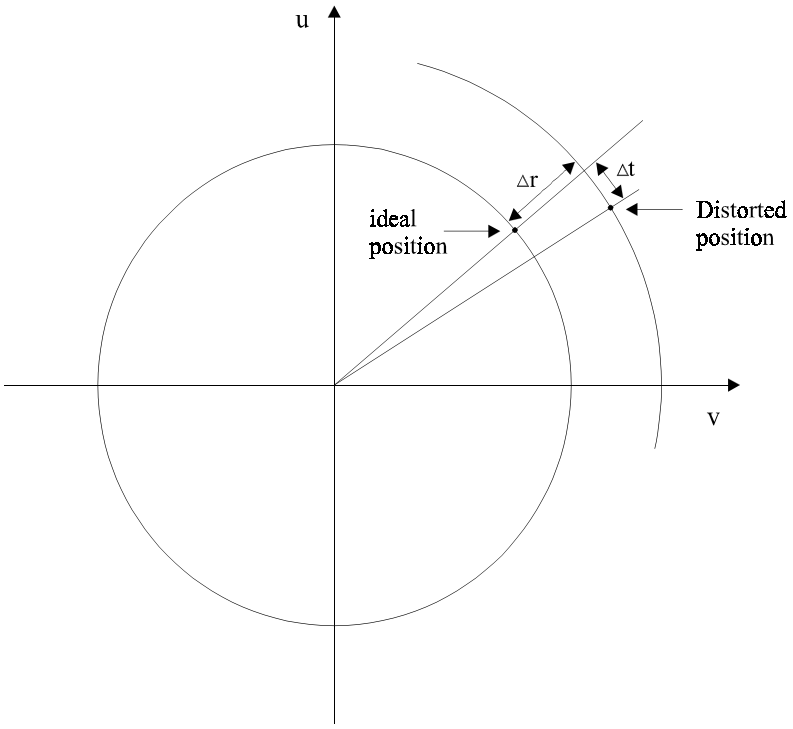


Figure 2.11 Nonlinear distortions, Δr is the radial distortion and Δt is the tangential.

If using the image coordinates (u_i^*, v_i^*) from the raw pinhole model which have their origin in the optical center, the transformation to real image coordinates can be expressed as (2.18).

$$\begin{bmatrix} u_i' \\ v_i' \end{bmatrix} = \begin{bmatrix} s_u(u_i^* + \delta u_i^{(r)} + \delta u_i^{(t)} + \delta u_i^{(p)}) \\ v_i^* + \delta v_i^{(r)} + \delta v_i^{(t)} + \delta v_i^{(p)} \end{bmatrix} + \begin{bmatrix} u_0 \\ v_0 \end{bmatrix} \tag{2.18}$$

$\delta u_i^{(r)}$ and $\delta v_i^{(r)}$ are the radial distortion, $\delta u_i^{(t)}$ and $\delta v_i^{(t)}$ are the tangential (decentering) distortion, $\delta u_i^{(p)}$ and $\delta v_i^{(p)}$ are the Thin Prism distortion and s_u is a scale factor to compensate for a difference in scale between the horizontal and vertical image coordinates (e.g. caused by line-jittering). If pixel coordinates are used instead of metric units (e.g. mm) two conversion factors D_u and D_v must be added to (2.18) giving the expression in (2.19). The exact values of D_u and D_v are not necessary because they are linearly dependent on the focal length f and the scale factor s_u. If the correct focal length is wanted they must of course be correct, the array size and number of pixels in a CCD chip can normally be found in the data sheet for the camera.

$$\begin{bmatrix} u_i' \\ v_i' \end{bmatrix} = \begin{bmatrix} D_u s_u(u_i^* + \delta u_i^{(r)} + \delta u_i^{(t)} + \delta u_i^{(p)}) \\ D_v(v_i^* + \delta v_i^{(r)} + \delta v_i^{(t)} + \delta v_i^{(p)}) \end{bmatrix} + \begin{bmatrix} u_0 \\ v_0 \end{bmatrix} \tag{2.19}$$

2.5.1 Radial-Symmetrical Distortion

The part of the radial distortion that is symmetrical around the optical center is called the radial-symmetrical distortion and is caused by imperfect manufacturing of the lens elements in the composed lens. In most calibration routines this is the only modelled distortion since it is far the largest component. The size of the radial-symmetrical distortion is therefore only dependent on the distance r_i to the principal point. In figure 2.12 is depicted the radial-symmetrical which with positive values gives a pincushion distortion and with negative values gives a barrel distortion. The radial distortion causes straight lines to be curved especially near the borders. The radial-symmetrical

distortion is zero in the principal point (optical center) and straight lines going through the principal point will still be straight lines (but the length will not be “correct”).

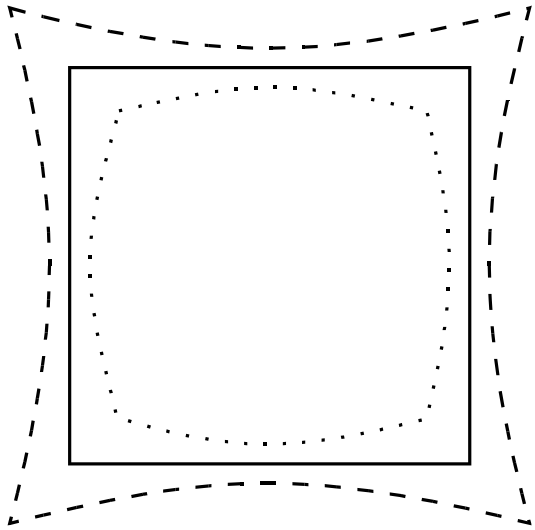


Figure 2.12 Radial-symmetrical distortion, solid line no distortion, dashed line positive distortion, dotted line negative distortion.

The radial-symmetrical distortion is traditionally described by a correction $\triangle r$ at the distance r_i from the principal point (see figure 2.13 (a)) modelled by a polynomial with uneven exponents normally up til the 7^{th} power, see equation (2.20).

$$\Delta r = a_1 r_i + a_2 r_i^3 + a_3 r_i^5 + a_4 r_i^7 \qquad \text{where: } r_i = \sqrt{(u_i^*)^2 + (v_i^*)^2} \tag{2.20}$$

Since the coefficient a_1 and the focal length f are interdependant a_1 is normally set to zero and the tangent of $\triangle r$ will therefore be zero for r_i=0. This is illustrated in figure 2.13 (b) which is a typical appearance of a radial distortion $\triangle r$ for an ordinary lens.

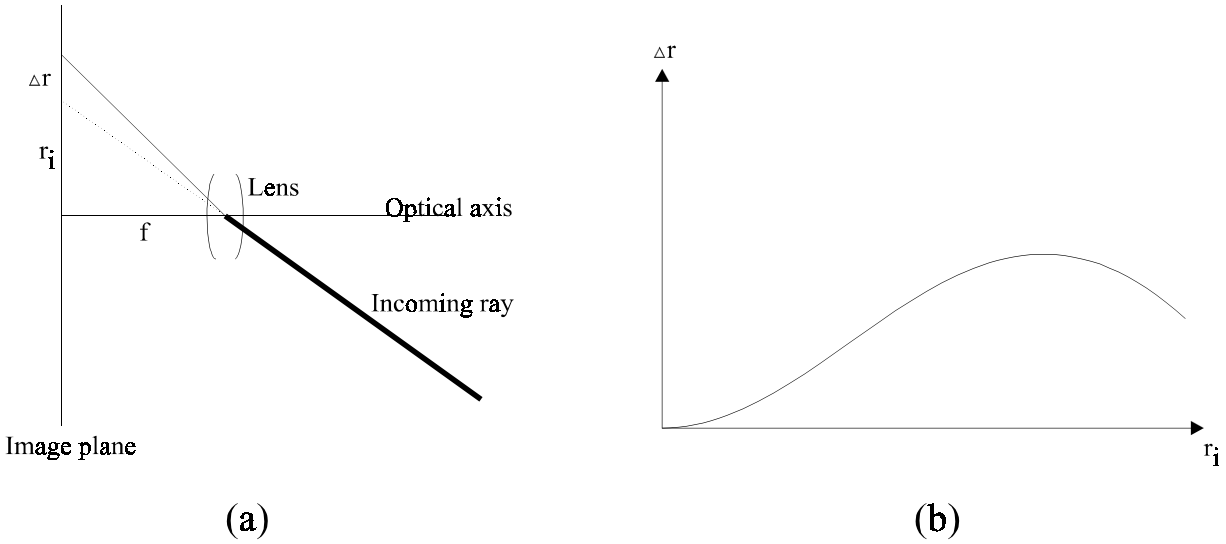


Figure 2.13 Radial-symmetrical distortion. (a) the out-coming ray of the lens does not follow the pinhole model (dotted line). (b) typical appearance of radial-symmetrical distortion, notice the zero tangent at r_i=0.

In order to use the equation in (2.20) in equations (2.18) or (2.19) the correction $\triangle r$ is projected into the image coordinate system giving the expression in (2.21).

$$\begin{bmatrix} \delta u_i^{(r)} \\ \delta v_i^{(r)} \end{bmatrix} = \begin{bmatrix} u_i^* (k_1 r_i^2 + k_2 r_i^4 + k_3 r_i^6) \\ v_i^* (k_1 r_i^2 + k_2 r_i^4 + k_3 r_i^6) \end{bmatrix} \tag{2.21}$$

In many cases (lens dependant) two coefficients are enough for an usable calibration [Heikkilä96b] and the two coefficient expression that was used and software implemented in this project are given in (2.22). The number of coefficients for the radial distortions must be chosen for the individual task dependent on the accuracy of the calibration target and its detection. In the references the number of coefficients varies from one [Weng92] ([Tsai86] two coefficient in article but only one in downloadable software) to three coefficients [Melen94].

$$\begin{bmatrix} \delta u_i^{(r)} \\ \delta v_i^{(r)} \end{bmatrix} = \begin{bmatrix} u_i^* (k_1 r_i^2 + k_2 r_i^4) \\ v_i^* (k_1 r_i^2 + k_2 r_i^4) \end{bmatrix} \tag{2.22}$$

2.5.2 Decentering Distortion

Decentering distortion occurs when the alignment between the lens elements in the composed lens is imperfect in such a way that the optical centres are not strictly collinear. If using polar image coordinates this can be described analytically by the expression in (2.23) from [Weng92] (notice that the *i* index is omitted in some of the next formulas due to software notation problems).

$$\begin{bmatrix} \delta_{\rho d} \\ \delta_{td} \end{bmatrix} = \begin{bmatrix} 3\cdot(j_1\rho^2 + j_2\rho^4 + j_3\rho^6 + ...)\cdot\sin(\varphi - \varphi_0) \\ (j_1\rho^2 + j_2\rho^4 + j_3\rho^6 + ...)\cdot\cos(\varphi - \varphi_0) \end{bmatrix} \tag{2.23}$$

Where φ_0 is the angle between the positive u axis and the line of maximum tangential distortion, see further down. The transformation to the image coordinate system (u,v) is given in (2.24).

$$\begin{bmatrix} \delta u_i^{(t)} \\ \delta v_i^{(t)} \end{bmatrix} = \begin{bmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{bmatrix} \cdot \begin{bmatrix} \delta_{\rho d} \\ \delta_{td} \end{bmatrix} \tag{2.24}$$

Combining (2.23), (2.24), $u = \rho\cos(\varphi)$ and $v = \rho\sin(\varphi)$ using the addition formulas for trigonometric functions and introducing the coefficients (2.25)

$$p_2 = -\, j_1\sin(\varphi_0) \quad \text{and} \quad p_1 = j_1\cos(\varphi_0) \tag{2.25}$$

the decentering can be expressed as the right side of (2.26) where $r^2=u^2+v^2$

$$\begin{bmatrix} \frac{u}{\rho} & \frac{-v}{\rho} \\ \frac{v}{\rho} & \frac{u}{\rho} \end{bmatrix} \cdot \begin{bmatrix} 3\cdot(j_1\rho^2 + j_2\rho^4 + j_3\rho^6 + ...)\cdot\left(\frac{v}{\rho}\cdot\frac{p_1}{j_1} - \frac{u}{\rho}\cdot\frac{-p_2}{j_1}\right) \\ (j_1\rho^2 + j_2\rho^4 + j_3\rho^6 + ...)\cdot\left(\frac{u}{\rho}\cdot\frac{p_1}{j_1} + \frac{v}{\rho}\cdot\frac{-p_2}{j_1}\right) \end{bmatrix} = \begin{bmatrix} \frac{(j_1 + j_2\rho^2 + j_3\rho^4 + ...)}{j_1}\cdot(2\cdot p_1\cdot v\cdot u + p_2\cdot(r^2 + 2\cdot u^2)) \\ \frac{(j_1 + j_2\rho^2 + j_3\rho^4 + ...)}{j_1}\cdot(2\cdot p_2\cdot v\cdot u + p_1\cdot(r^2 + 2\cdot v^2)) \end{bmatrix} \tag{2.26}$$

Assuming that

$$j_2 \rho^2 + j_3 \rho^4 + \ldots \approx 0$$

the commonly used expression for decentering (see [Weng92], [Melen94] and [Heikkilä97b]) is obtained in (2.27).

$$\begin{bmatrix} \delta u_i^{(t)} \\ \delta v_i^{(t)} \end{bmatrix} = \begin{bmatrix} 2p_1 u_i^* v_i^* + p_2 (r_i^2 + 2(u_i^*)^2) \\ 2p_2 u_i^* v_i^* + p_1 (r_i^2 + 2(v_i^*)^2) \end{bmatrix} \tag{2.27}$$

The decentering distortion has both a radial (non-symmetrical) and a tangential component (like in figure 2.11) that are modelled together. There is a line through the principal point along which the radial component is zero and along the line perpendicular to this line the tangential component is zero and the radial component is zero where the tangential component has it maximum and vice versa [Melen94].

To illustrate the shape of a decentering distortion the model from (2.27) has been software implemented and performed on the 900x900 pixel image of a circle with center in the principal point in figure 2.14 (a) with the coefficients p_1=0.0001 and p_2=0.0003 and setting the principal point to (450,450). The result of this simulation is shown in figure 2.14 (b) where both the radial and the tangential distortion are very visible. In figure 2.15 the original image is overlapped by the result, notice that the center is unchanged and that the line with maximum tangential distortion is perpendicular to the line with maximum radial distortion. From (2.25) it is obvious that $p_2 = 0$ makes the line of maximum tangential distortion horizontal (φ=0) and $p_1 = 0$ makes the line vertical ($\varphi=\pi/2$).

The modelling of the decentering distortion is often called the tangential distortion, but this is not completely correct since it contains both a non-symmetrical radial and a tangential component.

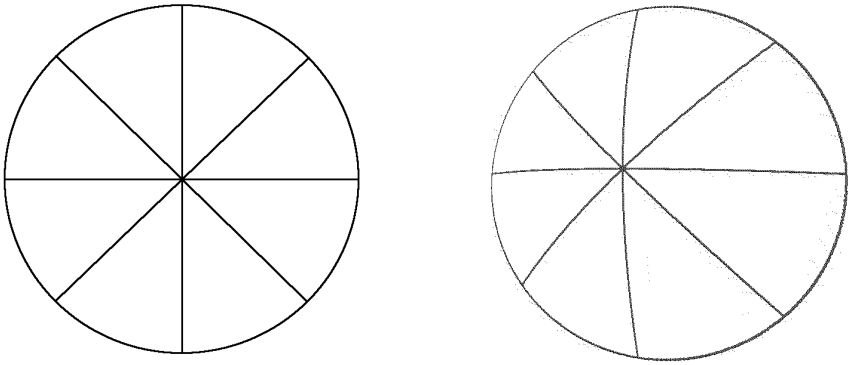


Figure 2.14 (a) circle without distortion (b) circle with tangential distortion p_1=0.0001 and p_2=0.0003 (u and v are between -450 and 450 pixels).

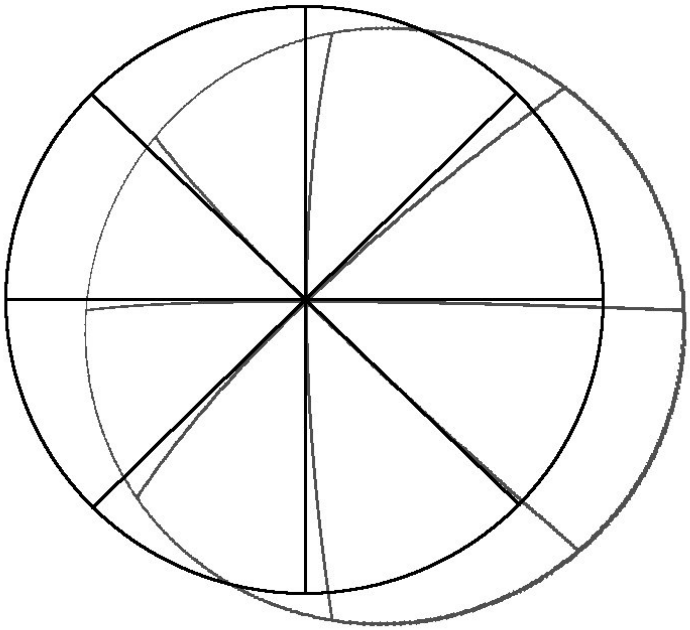


Figure 2.15 Circle with four lines through center without (dark) and with tangential distortion with coefficients p_1=0.0001 and p_2=0.0003.

2.5.3 Thin Prism Distortion

The thin prism distortion arises from imperfect manufacturing of lens elements and camera assembly like a small tilt of lens elements or the CCD-array, it can be adequately modelled by the adjunction of a thin prism to the optical system causing additional amounts of radial and tangential distortions [Weng92]. The thin prism distortion can be described analytically by the expression in (2.28) which is very similar to the decentering in (2.23) the only difference is that the constant “3" is removed from the δ_{pd}.

$$\begin{bmatrix} \delta_{\rho p} \\ \delta_{tp} \end{bmatrix} = \begin{bmatrix} (i_1\rho^2 + i_2\rho^4 + i_3\rho^6 + ...) \cdot \sin(\varphi - \varphi_1) \\ (i_1\rho^2 + i_2\rho^4 + i_3\rho^6 + ...) \cdot \cos(\varphi - \varphi_1) \end{bmatrix} \tag{2.28}$$

The angle φ_1 is like the φ_0 in decentering the angle between positive u axis and the axis of maximum tangential distortion, but φ_1 is not necessarily the same as φ_0, as the coefficients i_1 ... i_n are not the same as j_1 ... j_n. Similar to the coefficients in (2.25) the two coefficient s_1 and s_2 in (2.29) are introduced.

$$s_1 = - i_1\sin(\varphi_1) \quad \text{and} \quad s_2 = i_1\cos(\varphi_1) \tag{2.29}$$

After calculations and assumptions similar to (2.26) the common expression for thin prism distortion in (2.30) is achieved.

$$\begin{bmatrix} \delta u_i^{(p)} \\ \delta v_i^{(p)} \end{bmatrix} = \begin{bmatrix} s_1 r_i^2 \\ s_2 r_i^2 \end{bmatrix} \quad , \text{ where } r_i^2 = u_i^2 + v_i^2 \tag{2.30}$$

The thin prism distortion in (2.30) can be included in (2.18) in the same way as the radial-symmetrical and tangential distortion. In figure 2.16 is a simulation of the thin prism distortion from equation (2.30) similar to the simulation made in figure 2.15 with tangential distortion. The coefficients were chosen so that the radial distortions were almost the same size as in figure 2.15 for better comparison. As seen from the figures the shape of the two distortions are very similar but the tangential component is larger (almost double) with the same radial component in the thin prism distortion compared to the tangential (this is

not coefficient dependant).

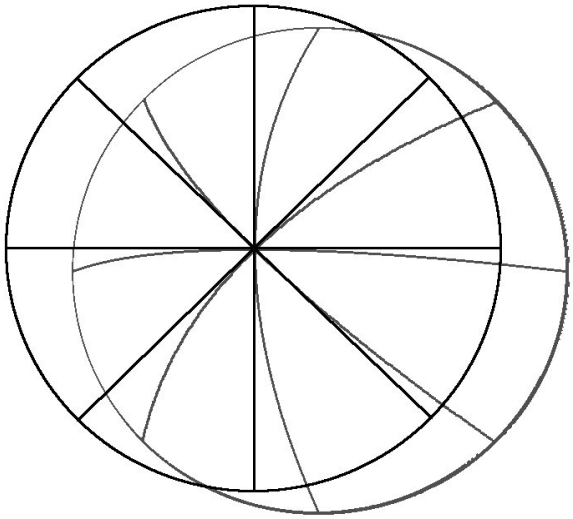


Figure 2.16 Circle with four lines through center without (dark) and with thin prism distortion with coefficients s_1=0.0009 and s_2=0.0003.

The thin prism distortion is not used much in calibrations and in most of the references it is mentioned but not used and in others it is not even mentioned. Although there has not been found references describing the size of thin prism distortion for actual CCD-cameras it is very likely (my opinion) that the error arising from tilt of CCD-array and lens components when assembling cameras is small compared to the decentering and radial distortion from lens manufacturing. Even if the thin prism distortion has a significant size the radial-symmetrical and tangential distortion modelling will compensate for the missing modelling. In [Heikkilä96b] experiments have shown that two coefficients for radial-symmetrical and two for decentering distortion are enough for most machine vision applications. It should also be considered that the radial-symmetrical distortion is the largest and that many calibration routines only calibrate for this distortion with good results, see e.g. [Tsai86]. The thin prism distortion has not been tested, or implemented in calibration software in this project. The thin prism distortion is though indirectly modelled using a 3rd or higher order polynomial warping model, see chapter 2.6 for more information.

2.5.4 Nonlinear Minimization

In order to estimate the parameters in (2.18) an iterative algorithm that minimizes the sum of squared residuals between the model and the observations is generally applied.

If N observations with image coordinates (U_i,V_i) and world coordinates (X_i, Y_i, Z_i) where i=1,..,N are obtained, the objective function to minimize can be expressed as J in (2.31), where the covariance matrix R_i can be a common matrix for all points using the common covariance matrix R in (2.32) [Melen94].

$$J = \sum_{i=1}^{N} \hat{e}_i^T R_i^{-1} \hat{e}_i$$

$$or \qquad (2.31)$$

$$J = \sum_{i=1}^{N} \begin{bmatrix} U_i - u_i \\ V_i - v_i \end{bmatrix}^T \begin{bmatrix} \mathrm{var}(U_i) & \mathrm{cov}(U_i,V_i) \\ \mathrm{cov}(U_i,V_i) & \mathrm{var}(V_i) \end{bmatrix} \begin{bmatrix} U_i - u_i \\ V_i - v_i \end{bmatrix}$$

$$R = \frac{1}{N} \sum_{i=1}^{N} R_i \qquad (2.32)$$

The covariance matrix R must of course either be known, estimated or set to a fixed value. In [Melen94] there is a method for estimating it, using multiple images with even conditions of the same calibration target. The most common is to use the unit matrix I since the image noise can reasonably be assumed to have a zero mean and be uncorrelated [Weng92] and [Heikkilä97b].

In order to start a non-linear minimization for the parameters in (2.18), a quite good initial guess of the parameters must be applied since bad initial parameters can cause the minimization to diverge [Heikkilä97b]. The number of iterations is also dependent on the quality of the initial guess. The common method for obtaining a good initial guess is to estimate the DLT matrix and obtain the

parameters from the decomposition, this method has been used in all the relevant references. The DLT decomposition only gives information about the linear distortion which cannot be used for the non-linear distortion and it is therefore common to set the nonlinear distortion parameters to zero in the initial guess, this method has also been used in the software implementation made in this project.

In [Melen94] is a method that can be used for obtaining initial parameters for the nonlinear distortion after estimation and decomposition of the DLT matrix. The idea is to compute the rest error $\tilde{e}_i = [\tilde{e}_{ui} \quad \tilde{e}_{vi}]^T$ given in (2.33) using the transformation matrix T, rotation matrix M and perspective matrix F from the decomposition of the DLT matrix.

$$\tilde{e}_i = q_i - (v + q_i^{'}) = q_i - (v + FMTp_i) \tag{2.33}$$

where **v** is the principal point, $\mathbf{q}_i$ are the measured image coordinates, $\mathbf{p}_i$ are the world coordinates of the objects in the calibration target.

Assuming the rest error equals the nonlinear distortion and since the nonlinear distortion is linear in the coefficients, the coefficients can be estimated solving an equation like the one in (2.34) where each point contributes with two equations. The equation in (2.34) uses the number of coefficients that are used for the nonlinear distortion in this project, but it can of course be modified to other coefficients.

$$\begin{bmatrix} \cdot \\ \cdot \\ \cdot \\ \tilde{e}_{ui} \\ \tilde{e}_{vi} \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} = \begin{bmatrix} & & \cdot & \\ & & \cdot & \\ & & \cdot & \\ u_i^* r_i^2 & u_i^* r_i^4 & 2u_i^* v_i^* & r_i^2 + 2u_i^{*2} \\ v_i^* r_i^2 & v_i^* r_i^4 & r_i^2 + 2v_i^{*2} & 2u_i^* v_i^* \\ & & \cdot & \\ & & \cdot & \\ & & \cdot & \end{bmatrix} \cdot \begin{bmatrix} k_1 \\ k_2 \\ p_1 \\ p_2 \end{bmatrix} \tag{2.34}$$

$$or$$

$$\tilde{e} = Q \cdot l$$

An equation like (2.34) can also be used for detection of overparametrization of a non-linear camera calibration, if e.g. the Q^TQ matrix is singular the number of coefficients for nonlinear distortion should be reduced. If using too many parameters the measurement errors will be absorbed in the model, and too few will give some systematic errors that are not compensated for in subsequent measurements.

After obtaining a good initial guess of the parameters an iterative nonlinear minimization routine must be applied, the method used in this project is the Levenberg-Marquardt method which has become the standard of nonlinear least-squares routines [Press86] and is used in almost all calibration routines [Heikkilä97b], [Weng92] and [Tsai86]. In the experiments made in this project the minimization has been acceptable after about 6-10 iterations with an initial guess based on the DLT and the nonlinear distortions set to zero. Other minimization routines like the quasi-Newton method can be used, see e.g. [Melen94]. In [Heikkilä97b] some results from experiments using the above model are shown.

Back-projection

A problem with the non-linear camera calibration described above is that it “only” estimates the camera parameters, these can easily be used for calculating the transformation from world coordinates to image coordinates, using the raw pinhole model in (2.3) and the distortion model in (2.18), but not the other way around (back-projection). A back-projection can be divided into two steps: first from raw image coordinates to distorted image coordinates, second from distorted image coordinates to world coordinates. The second step can easily be inverted from the pinhole model in (2.3) but the first step to convert raw image coordinates to distorted image coordinates is problematic. In (2.35) is the expression with only radial-symmetrical and decentering distortion, there is no (found) analytical solution to the inverse of this expression and a nonlinear search must be used [Heikkilä97b].

$$\begin{bmatrix}\delta u_i \\ \delta v_i\end{bmatrix} = \begin{bmatrix}u_i^*(k_1 r_i^2 + k_2 r_i^4) + 2p_1 u_i^* v_i^* + p_2(r_i^2 + 2(u_i^*)^2) \\ v_i^*(k_1 r_i^2 + k_2 r_i^4) + 2p_2 u_i^* v_i^* + p_1(r_i^2 + 2(v_i^*)^2)\end{bmatrix}, \text{ where } r_i^2 = (u_i^*)^2 + (v_i^*)^2 \quad (2.35)$$

It is also possible to use an approximation to the inverse model. The easiest is to use the distortion model itself as the inverse model [Weng92], in appendix D is a description of experiments made with this approximation on expression (2.35) which shows a good back-projection when the distortion is small. Another more advanced approximation is to model the inverse from a series of synthetic tie-points (>1000 points for good results) using a geometric transformation [Heikkilä97b].

2.6 Warp Based Camera Calibration

When the scene or setup makes it impossible to gain 3D information from a calibration target, the classic nonlinear calibration cannot be used since it requires the 3D information from either a 3D calibration target or multiple images from different places of a coplanar calibration target. This is often the case in close-range images where the depth of field is very low or when using devices like microscopes and scanners. In such cases the use of warping (geometric transformation) can be used instead. The use of warping for camera calibrations is restricted to planar world coordinates where the z-coordinate is omitted, since it is a transformation from 2D to 2D.

A warp which is a transformation from one coordinate system (x,y) to another (x´,y´) can be defined through the use of 2D-polynomials. In expression (2.36) is a general transformation of order N where the number of coefficients (a_{rk}´s and b_{rk}´s) are (N+1)*(N+2).

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \sum_{r=0}^{N}\sum_{k=0}^{N-r} a_{rk} x^r y^k \\ \sum_{r=0}^{N}\sum_{k=0}^{N-r} b_{rk} x^r y^k \end{bmatrix} \tag{2.36}$$

Using a planar calibration target it is possible to obtain a series of image coordinates (x,y) with corresponding calibration target 2D coordinates (x´,y´), also called tie-points, see figure 2.17. These tie-points can then be used to estimate the coefficients using a least square fit. In (2.37) is a linear system for a 2nd order warp, where the coefficient vector **l** can be least squares estimated from expression (2.38) assuming the covariance matrix Σ=I [Conradsen84]. A very practical thing about using warping as a camera calibration is that the direct transformation between image coordinates in pixels to calibration target coordinates in e.g. millimetre and visa versa can be estimated, this also means that information like the size of the CCD chip and focal length are not necessary.

$$\begin{bmatrix} \cdot \\ \cdot \\ \cdot \\ x_i' \\ y_i' \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} = \begin{bmatrix} & & & & & & \cdot & & & & & \\ & & & & & & \cdot & & & & & \\ & & & & & & \cdot & & & & & \\ 1 & y_i & y_i^2 & x_i & x_i y_i & x_i^2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & y_i & y_i^2 & x_i & x_i y_i & x_i^2 \\ & & & & & & \cdot & & & & & \\ & & & & & & \cdot & & & & & \\ & & & & & & \cdot & & & & & \end{bmatrix} \cdot \begin{bmatrix} a_{00} \\ a_{01} \\ a_{02} \\ a_{10} \\ a_{11} \\ a_{20} \\ b_{00} \\ b_{01} \\ b_{02} \\ b_{10} \\ b_{11} \\ b_{20} \end{bmatrix} \tag{2.37}$$

or

$$p' = W \cdot l$$

$$l = (W^T W)^{-1} W^T p' \tag{2.38}$$

A problem with a transformation like (2.36) is the very high number of coefficients to be estimated, if using e.g. a 3rd order warp it would require (3+1)(3+2) = 20 coefficients. Such a high number of coefficients requires many tie-points. There are various methods for automatic examination of the number or composition of coefficients, but a maximum order must be considered first (5th order is in most practical cases an acceptable maximum). In [Conradsen84] is described some methods "backwards elimination", "forward selection", "stepwise regression" and complete examination that are suitable for this purpose.

Complete examination ("all possible subsets regression") of all combinations is of course the best method, but this requires many computer resources. The number of combinations is given by 2^N where N is the number of coefficients to examine, if trying up to the 4th power, it requires the search of $2^{(4+1)(4+2)} = 1.07 \cdot 10^9$ combinations, if one combination requires $^1/_{100}$ sec. of computer time the complete search would take 124 days (examining up to the 3rd power would take three hours and up to the 5th power 1394 years).

In backwards elimination equation (2.38) is first solved for all coefficients and then a test value (F-value) is calculated for each coefficient and the one with the smallest value is removed. This is repeated until the smallest value is larger than a specific fractile in the relevant F-distribution [Conradsen84]. Another variant of this could be to compute all the W_j matrices where coefficient l_j is removed and then remove the coefficient that produced the largest D_j vector in expression (2.39), this is repeated until the lowest wanted accuracy is reached.

$$D_j = \left(p' - W_j \cdot l\right)^T \left(p' - W_j \cdot l\right) \tag{2.39}$$

Forward selection starts with the 0 and 1st order coefficients and all W_j matrices are estimated by adding one extra coefficient from 2nd to e.g. 5th power, the

coefficient that gives the lowest D_j vector in (2.39) is then added to the coefficient vector. This is repeated until the wanted accuracy is reached. The forward selection method is probably the best when very high order coefficients are wanted for the warping, since the W^TW matrix could be singular from the start with the backwards elimination method. High order coefficients are necessary if the warping should be able to model radial-symmetrical distortions with two or more coefficients since this gives at least 5^{th} order coefficients, see further down for more details.

Stepwise regression as it is described in [Conradsen84] is a forward selection method with the extension that each new coefficient is examined together with all the already selected ones, allowing already selected coefficients to be excluded. This method should be better than both backward elimination and forward selection but requires more computations.

When using high order transformations the number and composition of coefficients **must** be optimized using e.g. one of the above methods since overparametrization could cause fitting to the measurement errors and even worse give fluctuations near the image borders (or outside the calibration target area), which can give errors in subsequent measurements.

Relations to classic camera calibration

When using warping as a camera calibration the relations to the classic camera calibration models must be considered. When working with planar objects the 3x3 DLT from (2.8) is used for the linear camera calibration. If calculating (u,v) from the homogeneous coordinates in (2.8) and dividing out a_{33}, the transformation from planar world coordinates (x,y) to image coordinates can be expressed as in the so-called perspective transformation (2.40) where 8 new parameters a_1, a_2, a_3, b_1, b_2, b_3, c_1 and c_2 have been introduced.

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \dfrac{a_1 x + a_2 y + a_3}{c_1 x + c_2 y + 1} \\ \dfrac{b_1 x + b_2 y + b_3}{c_1 x + c_2 y + 1} \end{bmatrix} \tag{2.40}$$

The 8 parameters a_1, a_2, a_3, b_1, b_2, b_3, c_1 and c_2 can then be estimated from four point sets where three of them are not on a straight line, this could typically be done from the corners of a rectangle like in figure 2.17 that illustrates a use of the perspective transformation in (2.40). The 8 parameters are simply the elements in the A_{3x3} matrix from the planar DLT with the a_{33}=1 constraint and can be least squares estimated from (2.17).

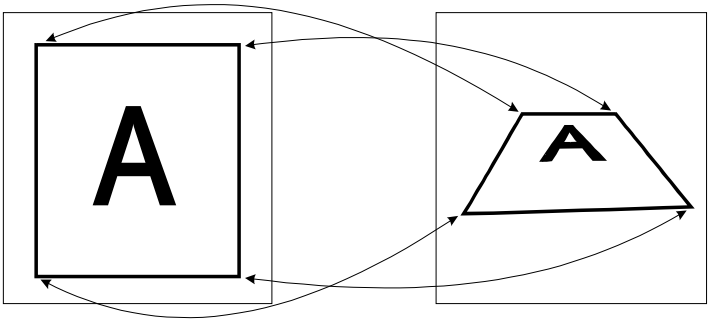


Figure 2.17 Four tie-points from (a) wanted image coordinates to (b) perspective image coordinates.

The expression in (2.40) is not linear with respect to x and y and if the expression should be implemented in the warping model in (2.36) a Taylor series expansion must be used. As a result of this the use of warping with images that contain perspective transformation (like in figure 2.17 (b)) can be modelled together with the nonlinear distortion using higher order polynomials. However, it will only be an approximation of the perspective transformation, and borders are likely to fluctuate. In appendix C is an experiment showing the relation between 1st to 6th order 2D polynomial warping and the perspective

transformation. In the work done in the corporation with Dantec Measurement A/S, see chapter 5, a full 3rd order 2D polynomial has been used with high quality lenses (low distortion) in a 45° perspective transformation with rather good result (about 0.1 pixel accuracy).

Polynomial warping can also be used in conjunction with a planar DLT, in a way so that the warping is only used for modelling the non-linear distortion, in this case very accurate results can be obtained, see next section on parallel image plane and object plane.

Parallel image plane and object plane

When working with a parallel image plane and planar object the transformation from the planar object to the image plane is quite simple. This simplification comes from the rotation matrix *M* in (2.4) since the ω and φ rotations are zero with parallel image plane and object plane. In expression (2.41) is the rotation matrix M_{2D} for parallel image and object plan.

$$M_{2D} = \begin{bmatrix} \cos(\kappa) & -\sin(\kappa) & 0 \\ \sin(\kappa) & \cos(\kappa) & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2.41}$$

In (2.42) is the result of inserting the 2D rotation matrix in the 3x3 DLT matrix expression (2.9) and calculating the corresponding DLT matrix, the linear distortion in form of the *B* matrix has been left out because of space problems (very large equation) but this does not change the meaning of the following.

$$\lambda \cdot \begin{bmatrix} 1 & 0 & -u_0 \\ 0 & 1 & -v_0 \\ 0 & 0 & 1 \end{bmatrix}^{-1} \cdot \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \cos(\kappa) & -\sin(\kappa) & 0 \\ \sin(\kappa) & \cos(\kappa) & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & -x_0 \\ 0 & 1 & -y_0 \\ 0 & 0 & -z_0 \end{bmatrix} \rightarrow \begin{bmatrix} \lambda \cdot f \cdot \cos(\kappa) & -\lambda \cdot f \cdot \sin(\kappa) & -x_0 \cdot \lambda \cdot f \cdot \cos(\kappa) + y_0 \cdot \lambda \cdot f \cdot \sin(\kappa) - u_0 \cdot \lambda \cdot z_0 \\ \lambda \cdot f \cdot \sin(\kappa) & \lambda \cdot f \cdot \cos(\kappa) & -x_0 \cdot \lambda \cdot f \cdot \sin(\kappa) - y_0 \cdot \lambda \cdot f \cdot \cos(\kappa) - v_0 \cdot \lambda \cdot z_0 \\ 0 & 0 & -\lambda \cdot z_0 \end{bmatrix} \tag{2.42}$$

As it is seen from (2.42) the conversion from homogeneous coordinates is now only a division with the constant $-\lambda z_o$ and the c_1 and c_2 constants from expression (2.40) are now zero giving the simpler expression in (2.43) which expresses the transformation between planar world coordinates and image coordinates. This expression equals the 1st order warp in expression (2.36).

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} a'_1 x + a'_2 y + a'_3 \\ b'_1 x + b'_2 y + b'_3 \end{bmatrix} \tag{2.43}$$

As a note can be mentioned that the expression in (2.42) can be used for confirming the statement that *"the focal length f and principal point (u_o,v_o) cannot be decomposed from a 3x3 DLT matrix if the image plane and object plane are parallel"*. In (2.42) we already know that $b_1=0$ and $b_2=0$ but there are still 8 parameters and only 5 equations (since $a_{11}=a_{22}$ and $a_{21}=-a_{12}$), we must therefore apply three of the four remaining intrinsic parameters f, λ, u_o and v_o. If applying the linear distortion (b_1 and b_2) to (2.42) it would only give two extra equations, which give the same result.

This means that a 1st order warp completely can emulate a linear calibration of an image with parallel image plane and object plane with translation, rotation, linear distortion and the principal point, but the physical parameters cannot be extracted.

When combining the three nonlinear distortions Radial-symmetrical, Decentering and Thin Prism from the previous section the expression in (2.44) is achieved (u,v with origin in the principal point), which obviously can be substituted with a warp with some coefficients set to zero.

$$\begin{bmatrix}\delta_u\\ \delta_v\end{bmatrix} = \begin{bmatrix} u\cdot\left[k_1\cdot\left(u^2+v^2\right)\right]+p_1\cdot 2\cdot u\cdot v+p_2\cdot\left[\left(u^2+v^2\right)+2\cdot u^2\right]+s_1\cdot\left(u^2+v^2\right)\\ v\cdot\left[k_1\cdot\left(u^2+v^2\right)\right]+p_1\cdot\left[\left(u^2+v^2\right)+2\cdot v^2\right]+p_2\cdot 2\cdot u\cdot v+s_2\cdot\left(u^2+v^2\right)\end{bmatrix} = \tag{2.44}$$

$$\begin{bmatrix} u^3\cdot k_1+u\cdot k_1\cdot v^2+2\cdot p_1\cdot u\cdot v+\left(3\cdot p_2+s_1\right)\cdot u^2+\left(p_2+s_1\right)\cdot v^2\\ v\cdot k_1\cdot u^2+v^3\cdot k_1+\left(p_1+s_2\right)\cdot u^2+\left(3\cdot p_1+s_2\right)\cdot v^2+2\cdot p_2\cdot u\cdot v\end{bmatrix}$$

As seen the nonlinear distortion gives no 1st order warping coefficients and the complete warping model for Radial-symmetrical, Decentering and Thin Prism distortion can be expressed as in (2.44), which correspond to a warp with 10 coefficients up to 3rd order. From (2.44) it is also seen that it is not possible to separate the Thin Prism and Decentering distortion when using warping as calibration.

There is however the problem with (2.44) that the distortion is defined with (0,0) at the principal point, and in order to use it with measured image coordinates a coordinate shift must be applied. This can be done by inserting $u=(a_1x+a_2y)$ and $v=(b_1x+b_2y)$ in (2.44) giving the expression in (2.45).

$$\begin{bmatrix} k_1\cdot\left(a_1\cdot x+a_2\cdot y\right)^3+\left(a_1\cdot x+a_2\cdot y\right)\cdot k_1\cdot\left(b_1\cdot x+b_2\cdot y\right)^2+2\cdot p_1\cdot\left(a_1\cdot x+a_2\cdot y\right)\cdot\left(b_1\cdot x+b_2\cdot y\right)+\left(3\cdot p_2+s_1\right)\cdot\left(a_1\cdot x+a_2\cdot y\right)^2+p_2\cdot\left(b_1\cdot x+b_2\cdot y\right)^2\\ \left(b_1\cdot x+b_2\cdot y\right)\cdot k_1\cdot\left(a_1\cdot x+a_2\cdot y\right)^2+k_1\cdot\left(b_1\cdot x+b_2\cdot y\right)^3+p_1\cdot\left(a_1\cdot x+a_2\cdot y\right)^2+\left(3\cdot p_1+s_2\right)\cdot\left(b_1\cdot x+b_2\cdot y\right)^2+2\cdot p_2\cdot\left(a_1\cdot x+a_2\cdot y\right)\cdot\left(b_1\cdot x+b_2\cdot y\right)\end{bmatrix} \tag{2.45}$$

If expanding (2.45) and introducing new constants (d_i and e_i), it can be applied to the linear model and the expression in (2.46) is achieved.

$$\begin{bmatrix}u\\ v\end{bmatrix} = \begin{bmatrix} d_1x^3+d_2x^2+d_3x^2y+d_4xy^2+d_5xy+d_6y^2+d_7y^3\\ e_1x^3+e_2x^2+e_3x^2y+e_4xy^2+e_5xy+e_6y^2+e_7y^3\end{bmatrix}+\begin{bmatrix}a'_1x+a'_2y\\ b'_1x+b'_2y\end{bmatrix}+\begin{bmatrix}a'_3\\ b'_3\end{bmatrix} \tag{2.46}$$

From (2.46) it can be seen that with parallel image plane and object plane a nonlinear camera calibration is fully modelled using a full (20 coefficients) 3rd order warp, if assuming only one coefficient radial-symmetrical distortion. If higher order radial-symmetrical distortion is needed for a satisfactory calibration, it is recommended to use a step-wise regression with e.g. up to 5th order coefficients.

A good benefit from using warping in an image with parallel image plane and object plane instead of classic nonlinear calibration is that the inverse model can be estimated without approximations.

Note on non-camera devices

When calibrating devices like scanners and line-scan cameras, the classic nonlinear models for cameras cannot be used and warp-based calibration with step-wise or forward regression is probably the best solution with such devices.

2.7 Resampling Methods

In many image processing tasks it is very desirable to obtain distortion and/or perspective transformation free images. This could e.g. be for aligning images from different cameras or camera views, for making texture measurements or simply for making visually good images. Such images can be made after a calibration where the expression to convert from world coordinates to image coordinates or distorted image coordinates to non-distorted has been obtained.

When making an output image from an input image, it is important to notice that the transformation is made from integer to integer coordinates, since it is digital images. It is not necessarily given that all integer coordinates in the input image will produce an output at all integer coordinates in the output image and there is also the possibility that several integer coordinates in the input image

are transformed into the same integer coordinate in the output image. It is therefore important to use the transformation from the output image to the input image (if it is accessible) [Niblack85]. In this way the floating point coordinates in the input image can be calculated from the integer coordinate output image. The floating point coordinates in the input image can then be used in an interpolation to get a suitable input image value.

When creating an output image the size and resolution must be decided. If the resolution is reduced it is called subsampling and if it is increased it is called oversampling. When a resampling is done in order to create an image without distortion and/or perspective transformation the most common is to decide the resolution so that the size of the pixel in the principal point is the same in the input and the output image. When using that method there will normally occur both subsampling and oversampling.

When subsampling an image to a much smaller resolution (< 0.5X) it is very common first to use a Gaussian mean filter on the input image, when using that method a nearest neighbour interpolation should be sufficient. The experiments made in this project have not shown any visual improvement of using a gaussian mean filter instead of a uniform mean filter. The uniform mean filter has the advantage that it can be software implemented to be very fast.

2.7.1 Interpolation Methods

Interpolation is a group of methods to estimate a continuous function from a set of discrete data. Meaning, find the pixel value that probably would have been in some floating point coordinate between the original pixels. In this section the three traditional interpolation methods and one more unusual will be described. There are various methods besides the four mentioned here (some based on B-splines) but the four described here are the most common in image processing.

A reconstruction of a piecewise continuous function from unit-spaced samples can be modelled as a linear combination of the input data and a reconstruction kernel like in (2.47) where h(s) is the reconstruction kernel and f_i are sample

values.

$$f(x) = \sum_{i=-\infty}^{\infty} f_i \cdot h(x - i) \tag{2.47}$$

There is of course the possibility of many different reconstruction kernels. In this thesis only four based on piecewise local polynomials will be mentioned, but a more detailed description of various kernels can be found in [Dodgson97].

When a 1D reconstruction is expanded to 2D (linear to bilinear etc.), it is first reconstructed in the horizontal direction and each of these reconstructions is then used for a reconstruction in the vertical direction (or opposite) [Niblack85].

Nearest Neighbour

When using Nearest Neighbour interpolation the value of the nearest pixel is chosen. The main advantage is that it is the fastest method and it is normally only used when high speed is required. Another advantage is that it does not introduce new pixel values which can be good in some statistical and image processing circumstances. It can also be used with heavy subsampling combined with a local (gaussian) mean filter. The disadvantage is the very low quality interpolation with various errors like thin lines being cut and curves to be jagged. In figure 2.18 is the 1D reconstruction kernel for the nearest neighbour interpolation.

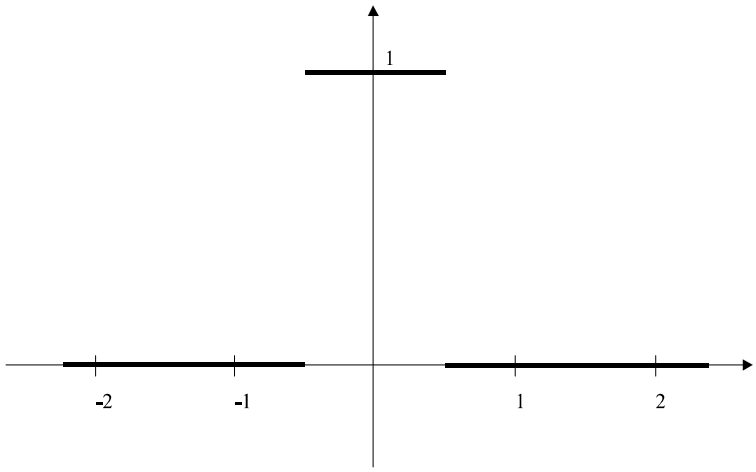


Figure 2.18 Reconstruction kernel for Nearest Neighbor Interpolation.

Bilinear

The linear interpolation is the 1st order interpolation, its 1D reconstruction kernel is in figure 2.19, when expanded to 2D it is called bilinear.

The Bilinear interpolation is slower than the nearest neighbour interpolation but is also much better. The main disadvantage of this methods is that it blurs the image a little. This can especially be seen if applying many resamplings to the same image and it should not be used in applications that requires multiple resamples. Another disadvantage is that the reconstructed signal is not differential, when oversampling this gives small ridges where the original pixels was. The advantages of this method is that it is the fastest of the more complex methods, it gives nice images and, highly importantly it does not introduce pixel values higher or smaller than any of the four neighbours. It can easily be hardware implemented for extremely fast interpolations, which is very common on to days front-end video cards. This method is the one mainly used in the work done in this project.

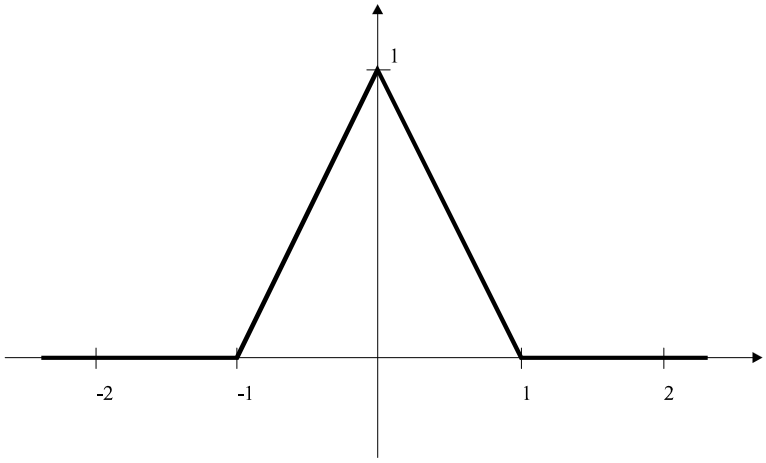


Figure 2.19 Reconstruction kernel for linear interpolation.

The calculation of the interpolated pixel value p(x',y') is given by the expression in (2.48) where an explanation of the notation can be seen from figure 2.20.

$$\begin{aligned} p(x',y') &= p(x,y')\cdot(1-dx) + p(x+1,y')\cdot dx \\ p(x,y') &= p(x,y)\cdot(1-dy) + p(x,y+1)\cdot dy \\ p(x+1,y') &= p(x+1,y)\cdot(1-dy) + p(x+1,y+1)\cdot dy \end{aligned} \tag{2.48}$$

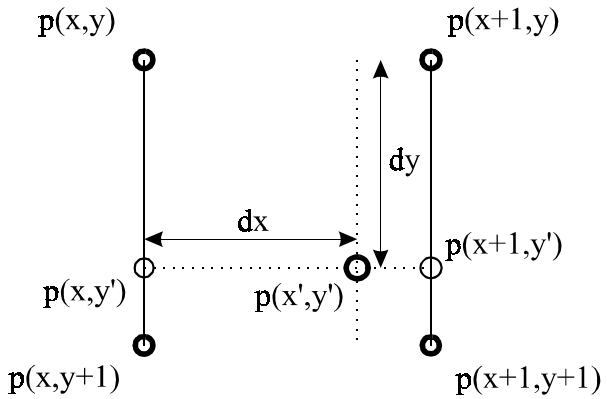


Figure 2.20 Bilinear interpolation

This approach is equivalent to calculating the 4 coefficients in expression (2.49).

$$I(dx,dy) = a \cdot dx + b \cdot dy + c \cdot dx \cdot dy + d \tag{2.49}$$

This can simply be done by solving the equation in (2.50), where p(x,y) is the pixel value at the (x,y) coordinate.

$$\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}^{-1} \cdot \begin{bmatrix} p(x,y) \\ p(x+1,y) \\ p(x,y+1) \\ p(x+1,y+1) \end{bmatrix} \tag{2.50}$$

In figure 2.21 is an example of a bilinear interpolation (10X oversampling) with the values *45, 25, 3* and *21* in the four original pixels (at the corners). Notice from the figure and the equation in (2.49) that the bilinear interpolation is not linear, since four points not necessarily are in one plane (three points are).

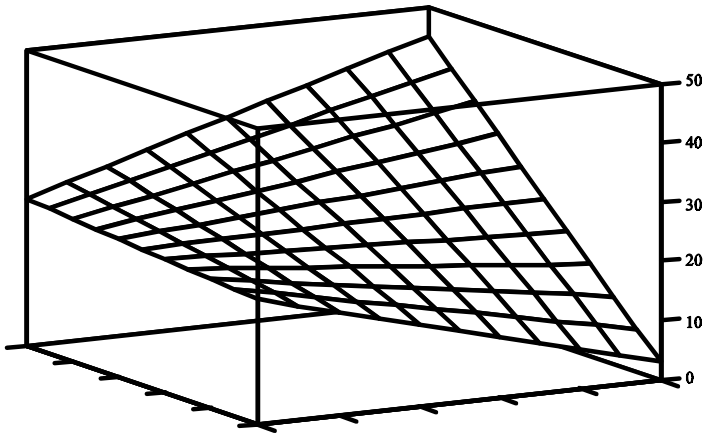


Figure 2.21 Bilinear interpolation. 10X oversampling.

BiQuadratic

The quadratic interpolation uses three samples (nine in biquadratic) with the kernel given in (2.51), the kernel is illustrated in figure 2.22. Notice the two non differentiable points at the small arrows. Biquadratic interpolation is very rare in image processing, but [Dodgson97] gives a very positive description of this kernel, and claims that it visually is very close to the higher order bicubic interpolations and that it is 55-63% faster than a bicubic interpolation. The BiQuadratic interpolation has traditionally not been used in image processing since it has been assumed to introduce phase distortion, but [Dodgson97] has proved that this is not the case. The main disadvantage of the quadratic/biquadratic interpolation is that the reconstructed signal is non-differentiable and as bilinear interpolation it gives small ridges where the original pixels were.

$$h(x) = \begin{cases} -2 \cdot |s|^2 + 1 & , \quad |s| \le 0.5 \\ |s|^2 - 2.5 \cdot |s| + 1.5 & , \quad 0.5 < |s| \le 1.5 \\ 0 & , \quad otherwise \end{cases} \tag{2.51}$$

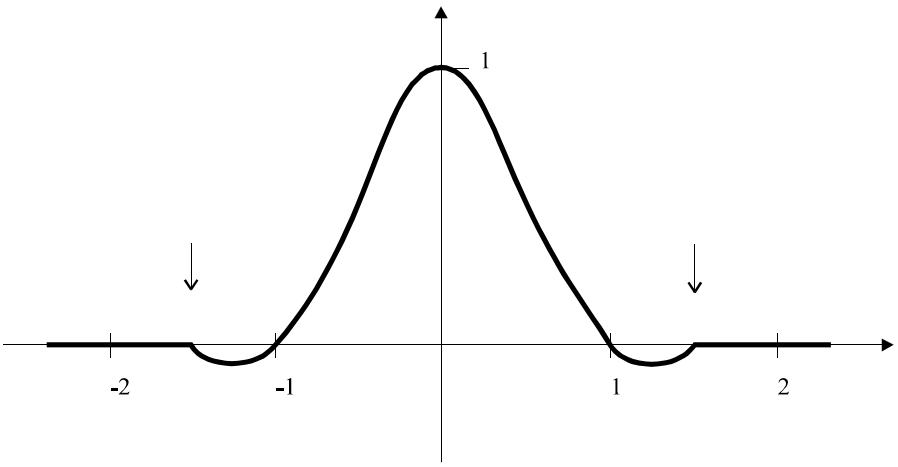


Figure 2.22 Reconstruction kernel for Quadratic interpolation.

BiCubic

The cubic interpolation uses four samples and the reconstruction kernel is based on two cubic polynomials. The most used kernel type is the “two-parameter family” where one of the parameters makes the kernel an approximating spline [Dodgson97]. Since splines are quite blurry for interpolation and therefore not considered in this thesis, the kernel is presented in (2.52) with one parameter, *C*.

$$h_{cc}(s) = \begin{cases} (2-C)\cdot|s|^3 + (C-3)\cdot|s|^2 + 1 & , \quad |s| \le 1 \\ -C\cdot|s|^3 + 5\cdot C\cdot|s|^2 - 8\cdot C\cdot|s| + 4\cdot C & , \quad 1 < |s| \le 2 \\ 0 & , \quad otherwise \end{cases} \tag{2.52}$$

The value of the parameter *C* can be chosen arbitrarily in the range *0<C<3*. But the two commonly used that can be found in the references are *C=0.5* and *C=1.0*. The *C=1* kernel can be found in [Niblack85] and is very common in software implementations. It is, however, not the best for oversampling. The *C=0.5* is the Catmull-Rom cubic [Dodgson97] that in this thesis is considered the best kernel for general purposes. In figure 2.23 is the reconstruction kernel for various values of *C*. Notice that they are continuous and differentiable, which also makes the reconstructed signal continuous and differentiable.

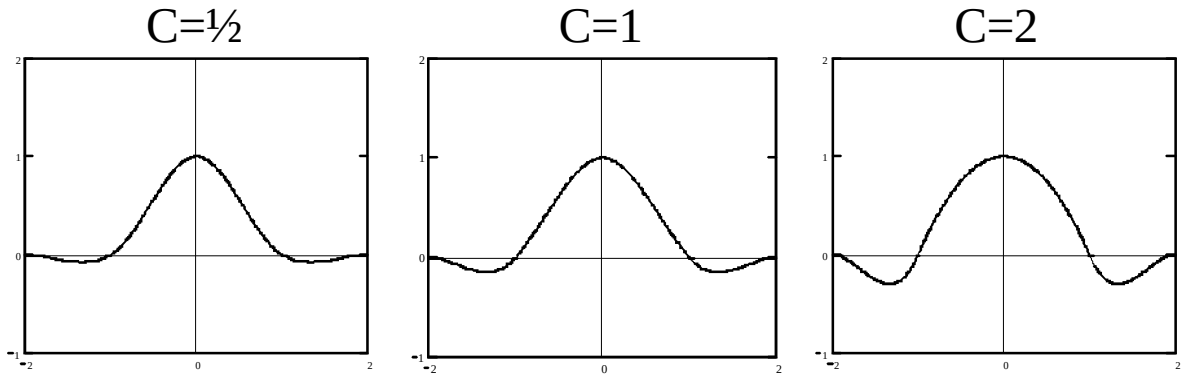


Figure 2.23 Reconstruction kernels for Cubic interpolation for various C values.

The Catmull-Rom cubic is a special case where the reconstructed signal has the

characteristic of the tangent in a point P_i to be parallel with the line going through P_{i-1} and P_{i+1}, this is illustrated in figure 2.24 for 6 data points, for information on splines and the Catmull-Rom family see [Foley92].

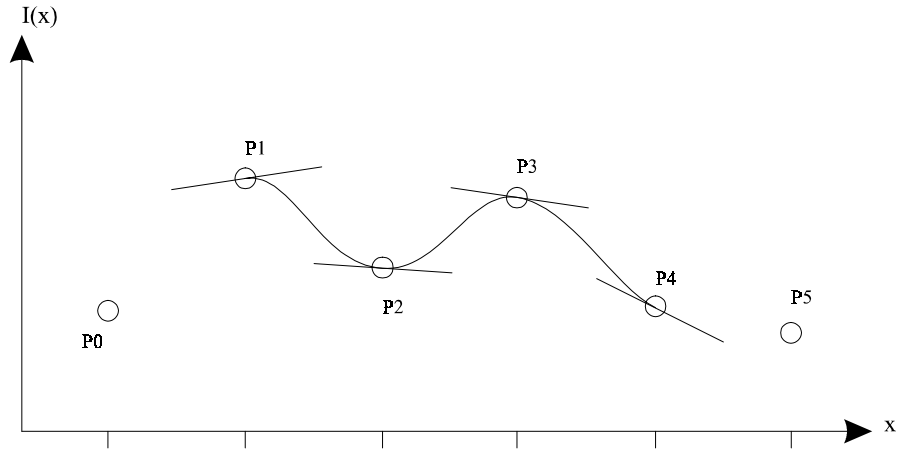


Figure 2.24 The Catmull-Rom spline family, the tangent in P_i is parallel with the line from P_{i-1} to P_{i+1}.

Actually, it is the *C* parameters that chooses the value of the tangent at a point P_i with respect to the line between P_{i-1} and P_{i+1}, to prove this the general formula *cc(x)* for the cubic interpolation is extracted below.

Calling the four original pixel values *y1*, *y2*, *y3* and *y4* at positions -1, 0, 1 and 2. The kernel *h(s)* from (2.52) can be inserted into the summation in (2.47) and when arranging the signs with respect to $0 \le x < 1$ the following expression is reached:

$$\begin{aligned} cc(x) = {} & y1\cdot\left[-C\cdot(1+x)^3+5\cdot C\cdot(1+x)^2-8\cdot C\cdot(1+x)+4\cdot C\right]+y2\cdot\left[(2-C)\cdot(x)^3+(C-3)\cdot(x)^2+1\right]+\ldots \\ & y3\cdot\left[(2-C)\cdot(1-x)^3+(C-3)\cdot(1-x)^2+1\right]+y4\cdot\left[-C\cdot(2-x)^3+5\cdot C\cdot(2-x)^2-8\cdot C\cdot(2-x)+4\cdot C\right] \end{aligned} \tag{2.53}$$

this can be simplified to:

$$\begin{aligned} cc(x) = {} & (y2\cdot(2-C)+y3\cdot(-2+C)-y1\cdot C+y4\cdot C)\cdot x^3+(2\cdot y1\cdot C+y2\cdot(C-3)-y4\cdot C+y3\cdot(3-2\cdot C))\cdot x^2+ \\ & (-y1\cdot C+y3\cdot C)\cdot x+y2 \end{aligned} \tag{2.54}$$

To find the tangent this is differentiated:

$$\frac{d}{dx}cc(x) = 3\cdot(y2\cdot(2-C)+y3\cdot(-2+C)-y1\cdot C+y4\cdot C)\cdot x^2+2\cdot(2\cdot y1\cdot C+y2\cdot(C-3)-y4\cdot C+y3\cdot(3-2\cdot C))\cdot x-y1\cdot C+y3\cdot C \tag{2.55}$$

And the tangent in *x=0* and *x=1* can be found as:

$$\frac{d}{dx}cc(0) = C\cdot(y3-y1) \qquad \frac{d}{dx}cc(1) = C\cdot(y4-y2) \tag{2.56}$$

From the tangents in (2.56) it is directly seen that with the Catmull-Rom cubic *C=½* the tangent at a point *y2* is parallel with the line from *y1* to *y3* and at point *y3* it is parallel to the line *y2* to *y4*. The expressions for the tangents can also be used to see how the differentiable reconstructed signal is generated, the tangents are preserved when moving one sample backwards or forwards.

Since it is difficult to find in the references the Catmull-Rom cubic is shown in a programming friendly expression in (2.57), this is directly comparable to a similar expression in [Niblack85] for the *C=1* kernel. When using it for bicubic interpolation this must be calculated five times for each point (first four in the horizontal or vertical direction and then one more time using the result of the first four).

$$\begin{aligned} cc(x) = {} & y2 + x\cdot((-0.5\cdot y1+0.5\cdot y3)+x\cdot((y1-2.5\cdot y2+2.0\cdot y3-0.5\cdot y4)+ \\ & x\cdot(-0.5\cdot y1+1.5\cdot y2-1.5\cdot y3+0.5\cdot y4))) \end{aligned} \tag{2.57}$$

The kernel with *C=1* that e.g. is used in [Niblack85] doubles the tangents, so with images (or image regions) that are smooth there will be very little difference between the *C=1* and the Catmull-Rom. But when the images contains edges, noise etc. the *C=1* kernel will be much more fluctuating, causing significant over- and undershooting. This can be used when re- or

undersampling an image and very little blurring is wanted for e.g. visual reasons. But when oversampling an image this can cause very unfortunate results. In the work done in this project the bilinear interpolation has given better results that the *C=1* bicubic interpolation, but when using the Catmull-Rom kernel the cubic interpolation has been far the best method. To illustrate this an example is given below.

The example uses an image of a circular black/white calibration object. A small 20x20 pixels subregion is extracted on the border of the object, this is then oversampled to 400x400 pixel using nearest neighbour, bilinear, bicubic (*C=1*) and bicubic Catmull-Rom interpolation. The result of these oversamplings is shown in figure 2.25 (a) to (d) where a small 6x6 region has been selected for better print visualization. It can immediately be observed that the nearest neighbour interpolation is very bad for oversampling. Visually the bilinear seems to be very good but the contour lines and the use of pseudo-colouring (not shown) reveals that it is not perfect. The cubic *C=1* interpolation seems to be quite bad, the original pixels can easily be seen, the contour lines are very curly and a pseudo-colouring also reveals heavy curling following the size of one original pixel. The Catmull-Rom cubic seems to very good and no errors can be revealed using contour lines or pseudo-colouring.

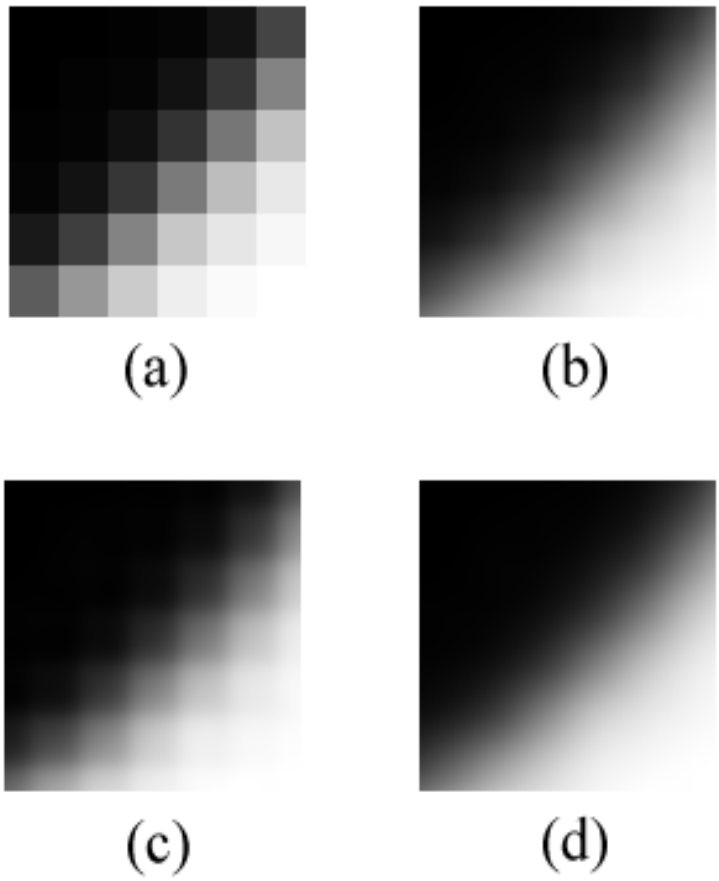


Figure 2.25 Oversampling of 6x6 subregion to a 120x120 image using (a) nearest neighbor (b) bilinear (c) bicubic (C=1) and (d) bicubic Catmull-Rom.

The contour lines of the two cubic interpolations can be seen in figure 2.26, notice the heavy curling in the *C=1* cubic.

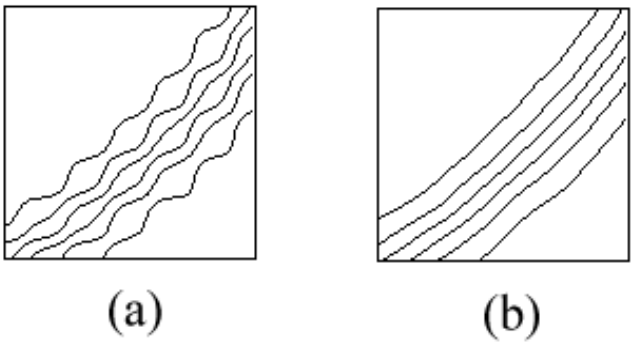


Figure 2.26 Contour lines for cubic interpolation (a) Niblacks C=1 (b) Catmull-Rom.

When looking at the line profile in figure 2.27 it can be seen that the bilinear interpolation in this case gives a better profile than the cubic *C=1* interpolation. It can also be seen that the Catmull-Rom cubic gives a much better profile than the three other methods.

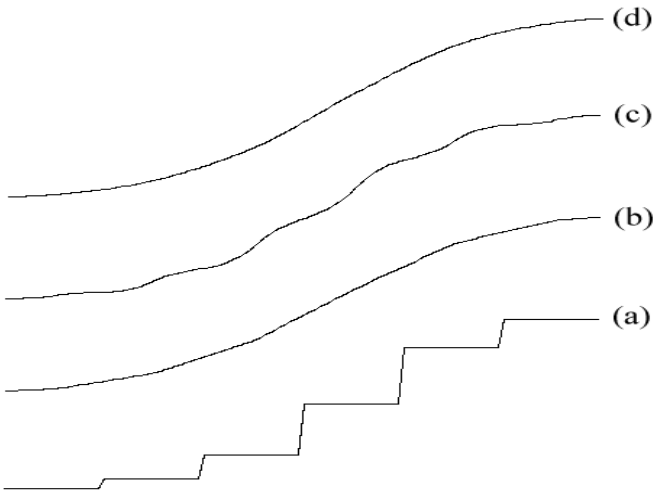


Figure 2.27 Vertical intensity profiles of a 6x6 pixel subregion interpolated to 120x120 pixels using: (a) nearest neighbor (b) bilinear (c) bicubic (C=1) and (d) bicubic Catmull-Rom.

The conclusion on the experiments made with interpolation in this project is that the Catmull-Rom cubic is far the best method for most purposes, especially for oversampling. There has been contact with other people doing image analysis that have chosen not to use cubic interpolation, since they could not

see the quality of the results compared to e.g. bilinear, this can in most cases be explained by the traditional use of the *C=1* kernel in software implementations instead of the Catmull-Rom kernel.

In general the bicubic interpolation is the slowest of the common interpolation methods in image processing but it is also the best in most cases. As described above the choice of the *C* parameter in the kernel is VERY important, since a bad selection for a specific task can make it visually worse than e.g. the bilinear interpolation. If an image has a very bad signal/noise ratio and no other preprocessing has been done the use of e.g. an approximating B-spline interpolation can be used, since they blur images by reducing peaks, see e.g. [Dodgson97] for a description of the spline kernels.

2.8 Summary

This chapter has presented useful aspects regarding the acquisition of monochrome images and a very detailed description of the aberrations and other lens defects that can influence a monochrome camera system and (hopefully) given series of good reasons to use high quality (expensive) lenses. There has been given a presentation of the classic camera calibration models (linear and non-linear) and descriptions of how to solve them. Most of these models have been software implemented in this project and have been used within this project and at external sites. In the special case where it is impossible to gain 3D information for a classic camera calibration the use of warping (geometric transformation) has been presented and how warping relates to the classic camera models. It has been shown that a 3rd order warping can make a high quality calibration that models three non-linear distortions when working with images with parallel image plane and object plane. For the applications that require rectified images, the use of resampling and various interpolation models have been presented, where the use of the Catmull-Rom cubic interpolation has been described in detail.

Chapter 3

Chromatic Calibration

The use of colour images in image analysis is common today (1999) but it is still quite new, especially with high precision measurements. Many references working with camera calibration either completely ignore the colours (chromatic aberration etc.) or simply make an assumption about chromatic errors being very small and therefore negligible.

As an example can be mentioned that John C. Russ on page 18 in [Russ90] states that "*...the software to analyze the three images in combination is difficult to design because there is almost too much information and it is hard to design algorithms that always use it correctly*." which obviously is outdated since the use of colour images has become quite common and there is no problems working with colours, unless the used routines/software packages are old. There have not been any computational problems with the use of colours in this project.

This chapter does not contain many references, mostly because of colour issues being new in image processing and especially calibration. Many of the statements in this chapter are therefore based on the experience gained from this project.

3.1 Colours

When working with colour images it is important to keep in mind that the colours we work with only is a small subset of the electric-magnetic spectrum as shown in figure 3.1. The human eye can only see wavelengths from about 395 nm to 750 nm (the visible light), but cameras and other artificial receptors can be designed to detect light in both the ultraviolet and the infrared wavelength. However it is possible for humans to detect the infrared light through their skin as heat (like when standing near a closed stove) and the effect of ultraviolet light can also be felt as burning skin after a day in strong sunlight.

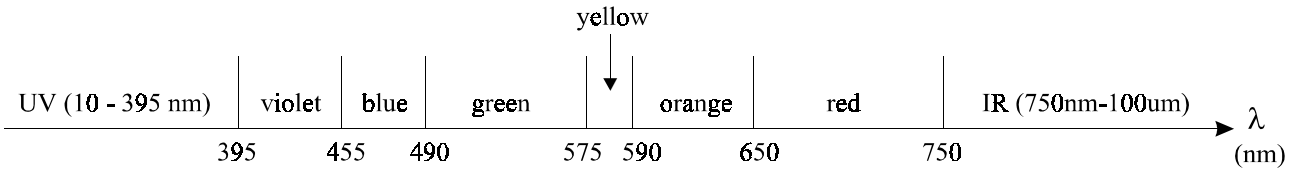


Figure 3.1 The electro-magnetic waves of interest in optics.

When designing lenses and CCD-chips it is the wavelengths to be used that is the interesting parameter, this is also the case when calibrating e.g. a camera for various wavelengths. But when working with e.g. colour textures or colour measurements in a system that relates to materials for humans, the colour system for the human eye must be taken into consideration. Colours like white, grey and purple are mixtures of several wavelengths and can therefore not be expressed as a single wavelength. On the other hand single wavelength colours like yellow and darker variant brown are only known as mixtures in the human eye and computer screens.

The special case white light (or grey) is the mixture of all the wavelengths in a composition as found in the sunlight. This can easily be seen in a rainbow or white light spread out in a prism. When working with monochrome cameras in normal daylight or using e.g. a light bulb it is therefore important to notice that the lens system should be able to work correctly with all visible/detectable

wavelengths. The opposite case single wavelength (monochromatic) light is normally only achieved with a laser.

The human eye uses three different receptors called cones in daylight and another receptor type called rods for the night vision. The rods are very sensitive compared to the cones and they can detect all the visible wavelengths, but they cannot distinguish wavelengths, this is also the reason why humans cannot see colours at night. The three cones called α, β and γ are used for colour vision, in figure 3.2 are some typical absorption curves for the three cones. Notice that the α cone which primarily is in the blue wavelength has a very low sensitivity and that there is considerable overlap between the curves. The existence of these three cones provides the physiological basis for trichromatic theory of colour vision [Pratt91].

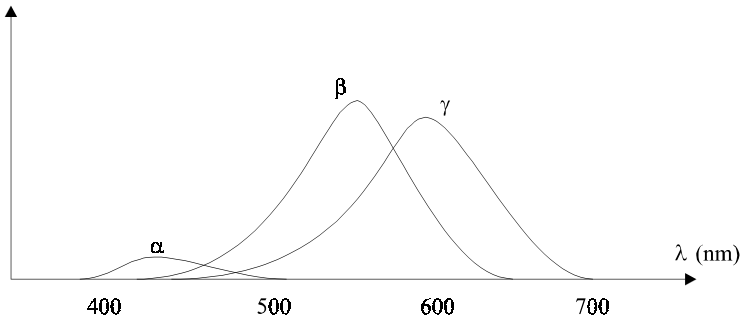


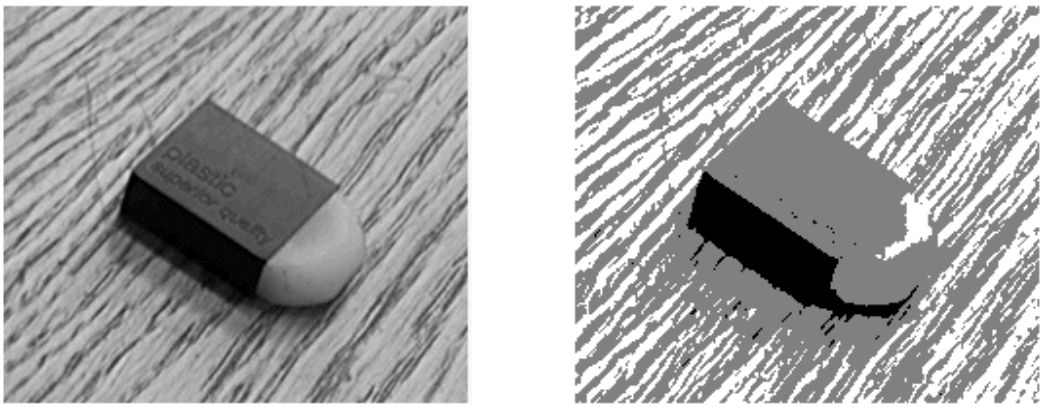
Figure 3.2 Typical spectral absorption curves of the three cones on the human retina, from [Pratt91].

The colour vision systems like television, photos, video cameras and almost all other commercialized colour devices have traditionally been using the trichromatic models from the idea that “all” colours can be mixed from three primary colours [Pratt91]. This is obviously wrong in the sense of a non-human observer, like an advanced colour measurement device, as an example can be mentioned that even very simple colour devices can distinguish between a monochromatic yellow and a yellow mixed from red and green wavelength. It is therefore very important to consider that the trichromatic colour vision only is adapted to the human eye and can produce very poor results if applied without considerations to a true multi spectral system. If e.g. some sort of system has three lights red, yellow and green (like a traffic lights) that should be kept under surveillance, a wrong designed trichromatic system could easily

mistake a single yellow light from the red and green light turned on simultaneous.

It can be mentioned that most airborne cameras (and satellites) for remote sensing use multi-spectral cameras with e.g. 64 channels from the UV to the IR wavelengths. Also very advanced scanners like the Eskofot 2540 use more channels than three to produce high quality images.

Since the human eye cannot see multi-spectrally it is difficult to illustrate how the trichromatic colour vision reduces the details, but to illustrate the problem in one dimension less, figure 3.3 contains a picture of an eraser on a wooden surface in 32 grey levels (channels) and 3 grey levels (channels). As seen in the figure the 32 grey levels gives a nice representation with many details, the 3 levels also gives information about the position of the eraser, its shadow and the grain in the surface, but all details in the grain and the text on the eraser is missing. This can give an idea about how the use of three colour components instead of e.g. 32 can reduce important information about an object surface. Like when a newspaper uses 2 levels (black and white) to create all grey level through raster the trichromatic system can also cheat the human eye.



(a) (b)

Figure 3.3 A picture of an eraser in (a) 32 grey level and (b) 3 grey level. This illustrates the loss of details when using fewer values to represent data.

When mixing colours to match other colours there are two basically different methods: additive and subtractive. The additive matching works by blending light in the primary colours, it is used with e.g. monitors, the common primary colours are red, green and blue. The subtractive matching uses white light passing through filters in the three primary colours, this is e.g. used with colour printers, that start with white paper, the most common primary colours for subtractive matching are cyan, magenta and yellow (to compensate for bad colours a black ink is often applied for a four colour print quality). A perfect yellow will absorb the blue light, the cyan absorb the red light and the magenta absorb the green light.

When working with colour cameras and scanners the most commonly used primary colours are red, green and blue, where the spectral distribution of each component completely relates to the capturing device. For scientific devices the spectral response for each channel can normally be found in the technical specifications of the device.

Unfortunately there has been no possibilities to work with other devices than trichromatic red, green and blue cameras and scanners in this project, a multi-spectral light source based on a filter wheel has been ordered for the colour texture project but has not been ready for this thesis. The rest of this thesis will therefore be concentrated on the use of trichromatic devices.

The three primary colours red, green and blue (RGB) can be treated as a 3-dimensional colour space as illustrated in figure 3.4, where the cube can contain all the possible colours to mix from the RGB, in this cube all grey tones will be on the diagonal from (0,0,0) to (1,1,1).

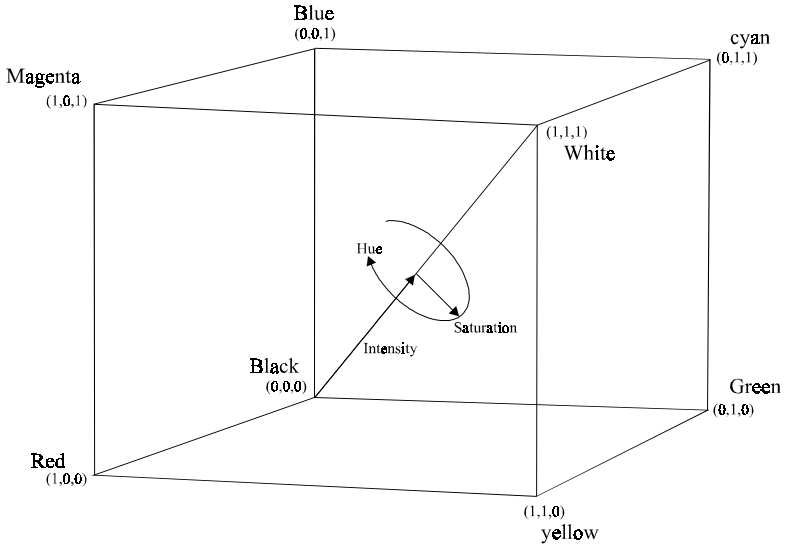


Figure 3.4 The RGB cube and the IHS system.

A very useful nonlinear transformation of the RGB coordinates is the Intensity/Hue/Saturation (IHS) system, where I represents the intensity level, S is the saturation of the colour (distance from the intensity diagonal to RGB point) and H is the colour expressed as an angle (around the intensity diagonal). The IHS system is also illustrated in figure 3.4. The transformation from RGB to IHS is in (3.1).

$$\begin{bmatrix} I \\ U \\ V \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{-1}{2} & \frac{-1}{2} & 1 \\ \frac{\sqrt{3}}{2} & \frac{-\sqrt{3}}{2} & 0 \end{bmatrix} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix}, \quad H = \tan^{-1}\left(\frac{V}{U}\right), S = \sqrt{U^2 + V^2} \tag{3.1}$$

When using the expression in (3.1) hue will be zero at the colour blue, this is also a problematic issue about the hue coordinate, it is not continuous in the colour where hue passes $0^0/360^0$, this is e.g. a problem when using IHS in

colour texture measurement. There are a number of other features similar to Hue that can be calculated from the RGB coordinates which do not suffer from discontinuities, some of these are described in chapter 5 in the section about colour texture measurements.

Another colour measurement that is based only on one value is the *temperature* of the light, it is measured in Kelvin (K) and it is defined as the radiation from a black body with the given temperature [Pratt91]. When e.g. observing a white object a low temperature light gives a reddish/greenish appearance and a high temperature a greenish/blueish.

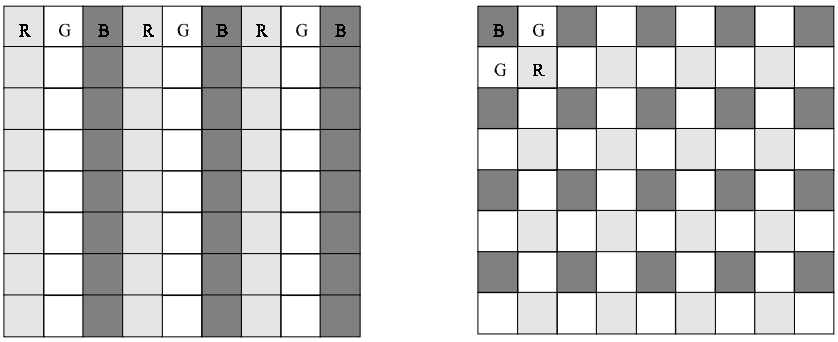
3.2 Acquisition

Colour images can come from a variety of sources but in this project there has only been used colour CCD cameras (single chip and 3xCCD), monochrome cameras in combination with coloured light and scanners. The analysis of colour images will only be focussed on images from these devices and especially 3xCCD cameras since they have been a significant part of the primary industrial targets of this project.

3.2.1 Single chip colour cameras

The most common digital colour cameras are based on a single chip since that is the most simple to produce, they are common in the consumer market and they are also the least expensive colour cameras to produce. The principle of the single chip colour cameras is to apply a colour filter individually to each cell, this can be done in several ways and the composition/distribution of the different filters must be considered, the two most common are illustrated in figure 3.5. The striped (a) has the advantage of an even distribution of the different filters and that each line have an equal number of the different filters, a disadvantage is that the distance between two equal filters is three pixels in

horizontal direction. The 2x2 Bayer mosaic (b) which is the most common has the advantage of the maximum distance between to different filters to be two and the disadvantages of an uneven filter distribution and odd and even lines to have different distribution of the different filters. The 1-2-1 distribution of the three filters in the 2x2 mosaic is often used as one for red, two for green and one for blue, since the central wavelengths (green) are considered to be the most important for the human eye.



(a) (b)

Figure 3.5 Single chip color CCD chip with R, G & B cells, (a) striped (b) 2x2 Bayer mosaic.

When comparing the digital chip with the retina in the human eye, the monochrome chip is similar to the rods and the RGB mosaic in the colour camera is similar to the cones.

The use of single chip cameras for capturing RGB images has some major disadvantages. First the resolution for each channel is at most ½ of the number of pixels in the chip (dependent on the channel and mosaic design see figure 3.5). Second the colour pixels are not located at the same places making it impossible to make a perfect fit between the red, green or blue pixels, this is also known as Colour Aliasing. This can be described simplest with a black/white step-edge in the image which will contain unwanted colours from the misalignments of the RGB pixels. This depends on the interpolation method and cannot be removed by calibration and in many cases an intentional blurring of the image is the only way around it.

In most single chip colour cameras there is a significant spectral overlap between the R, G and B channels caused by the on-chip filters. This has been seen in all the technical specifications for single-chip colour cameras that have been used in this project.

In some cameras four different filters are used instead of three to generate more correct colours in the images, in the CMOS cameras from Sound Vision Inc. the four filters are red, green, teal (green/blue) and blue, they are then converted to RGB using a 3x4 matrix like expression (3.2). In this case an even distribution of the four filters can be applied to a mosaic similar to figure 3.5 (b). Some women are born with four types of cones, called tetrachromaticity, which actually enables them to discriminate certain colours better than other humans.

3.2.2 Beam-splitter colour cameras

The best multipurpose colour cameras are the beam-splitter based cameras, these are most common as 3xCCD cameras. They can be used for almost all conditions like a moving camera, a scene with movements, oscillating lighting etc.). In figure 3.6 is an illustration of an RGB beam-splitter from a JAI 3xCCD camera (the drawing is a simplification). The prisms work as semi-transparent mirrors and are normally coated with metals and dielectrics to give them the correct reflection characteristics.

The main advantages of the beam-splitter cameras are that they give full resolution in all channels, they do not suffer from Colour Aliasing and they make a very good spectral separation compared to e.g. the single-chip colour cameras. The disadvantages are that they are difficult to manufacture because of the beam-splitter and therefore often very expensive compared to single-chip colour cameras, they suffer from a spatial intensity level difference that is different for each channel and the spatial alignment between the channels is normally not perfect.

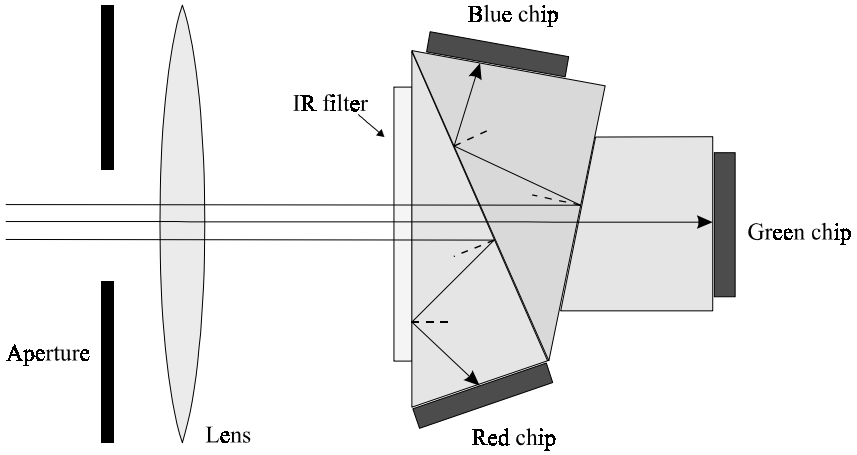


Figure 3.6 Illustration of 3xCCD camera with beam-splitter.

As described in the previous section there is a significant spectral overlap between the R, G and B channels in most single-chip (filter based) colour cameras. This is normally not the case when using a beam-splitter, where the optical prism block can make very good spectral separation. In figure 3.7 is shown a coarse plot of a spectral separation from a 3xCCD camera. The low and high frequencies are cut by the filter in front of the prism block, see figure 3.6. This can make the low cut of red and the high cut of blue different from the prism block separations. The spectral separation of the beam-splitter in e.g. a 3xCCD camera can normally be found in the technical specifications, this must be combined with the spectral response for the used chips (CCD, CMOS etc.), if not already specified in the specs.

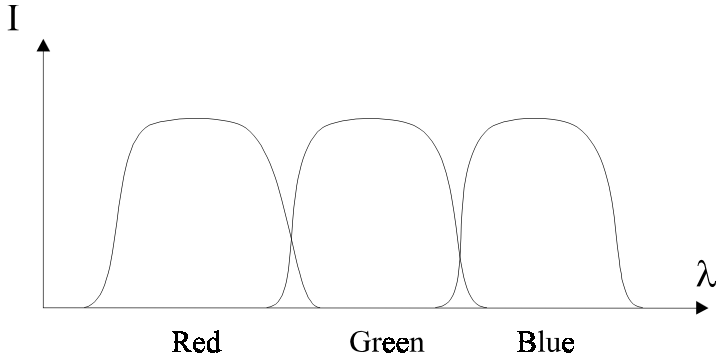


Figure 3.7 RGB response curves for 3xCCD camera.

If the Beam-splitter (optical prism block) is not manufactured well enough several error sources are possible. First spatial misalignment of the channel which has been observed at a significant level in this project. Second different focus for each channel which also has been observed in this project. Third secondary images (shadow images) this has not been observed in this project. Fourth Colour Shading (described below).

Colour Shading

The last of the major error sources is called Colour Shading (some call it Dynamic Shading) and it is caused by the fact that the optical length for the rays through the prism block are not equal, depending on where in the image plane they hit. This phenomena is typical for the beam-splitter cameras. It can easiest be observed by grabbing an image of a uniform white surface where it will cause the image to have a slightly coloured (non white) top and bottom. Most 3xCCD cameras are adjusted linearly for this phenomena and have additional adjustments for typically the red and green channels. These adjustments are simply a small electric circuit that make a linear adjustment of the level of each line. The additional adjustment is necessary since the Colour Shading is dependent on the lens. With a JVC KY-55 camera used in this project it has been the most significant error, there has been observed up to 60 grey levels (with 256 levels) difference between the red and green channel. The work with these cameras in this project has also shown that the Colour Shading is not a linear error with respect to the line number and that all channels and not just red and green have this error, see figure 3.15 and 3.16.

3.2.3 Multiple shot colour cameras

When working with non-moving scenes/objects multiple images can be captured (shot) with a monochrome camera using different filters to create

colour (multi-spectral) images. This can be done using two different methods, in front of the camera (IFC) (or more complex inside the lens/camera) or in front of the illumination (IFI). In figure 3.8 is shown the layout of an “in front of camera” system with 7 filters and one empty holder. The empty holder is useful (but not necessary) for setting up the system with regard to focussing the lens etc. In figure 3.9 is the transmittance curves for a filter set from Melles Griot (400, 450, 500, 550, 600, 650 and 700 nm) that could be very useful for a system like in figure 3.8. The Melles Griot set also includes one blank (empty) holder.

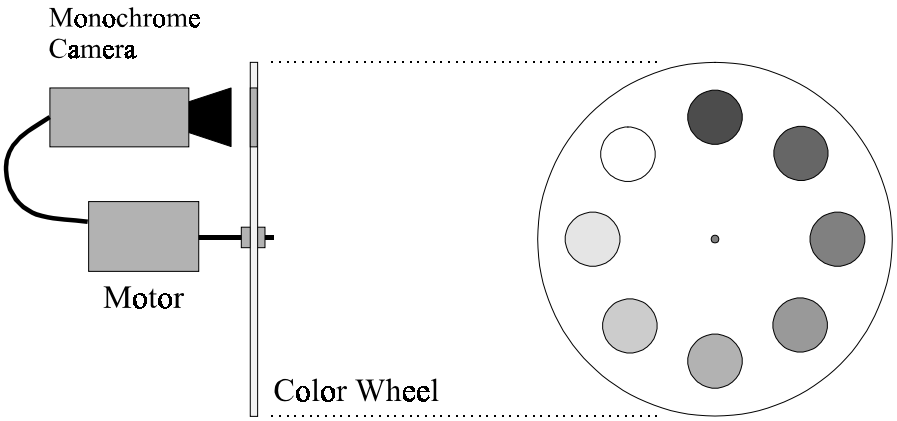


Figure 3.8 The system of a multiple shot color camera with 7 filters and one empty. The motor and the camera must be synchronized to grab the images at the right time.

For an IFC system like in figure 3.8 or an IFI system the colour wheel (or motor) must be very well synchronized with the frame capture on the camera, this can easily be done using the “*Ext. Sync.*” that is present on almost all cameras for machine vision, through an electric circuit connected to some sort of either optical or mechanical marker(s) on the wheel.

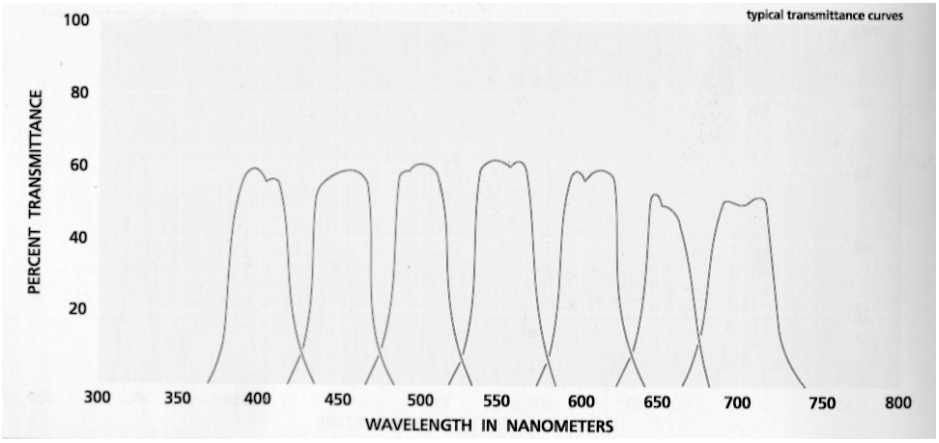


Figure 3.9 Transmittance curves for “Visible-40" which is a 40 nm bandwidth filter set with seven filters and one blank from the Melles Griot 1997/98 catalog page 210.

If working with optically closed systems (controlled light sources) the IFI system has several advantages compared to the IFC, where the main advantage is that there are no distortions and other errors from a filter in front of the camera, allowing for normal camera calibration methods. Dust and scratches on the filters are also more acceptable if they are not placed in front of the camera. Another advantage is that small vibrations from the rotating colour wheel are more likely to influence the camera if placed close to it. When making an IFI system optical fibers are useful conveyers of the coloured light from the colour wheel, this is also the method that has been used in this project.

If working with optical open systems (like in free air or in an office) the existence of other light sources makes the IFI system inaccurate or useless and an IFC system must be used. In such a system an accurate calibration must include some either direct or indirect modelling of the distortions from the filter in front of the lens.

The multi-shot method is very useful for capturing multi-spectral images, through the collaboration with *DELTA Lys & Optik* it has been told that to days highly sophisticated design of filter with many layers of metal and dielectric makes it possible to design very narrow filters. It is e.g. possible to make a good series of eighteen 20 nm filters to cover the 390 nm to 750 nm spectra, or even 10 nm width filters.

When working with closed optical scenes, like an Integrated Sphere, a closed cylinder or a dark room the use of a multiple shot camera is by the author considered the best solution for all non-moving scenes/objects. This is mainly because of the possibilities of capturing multi-spectral images and that the beam-splitter based cameras used in this project have a considerable colour-shading error.

If a multi-channel image captured from an IFI or IFC system is to be displayed in RGB (monitor, printer etc.) or some calculations only are defined for RGB the channels can be converted to RGB using an expression as in (3.2). A conversion as in (3.2) requires a common zero, which is fulfilled with an IFI or IFC system. If the channels come from different cameras and/or grabbers a common zero is no longer necessary given and expression (3.2) must be expanded.

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1N} \\ a_{21} & a_{22} & \dots & a_{2N} \\ a_{31} & a_{32} & \dots & a_{3N} \end{bmatrix} \cdot \begin{bmatrix} C_1 \\ C_2 \\ . \\ . \\ . \\ C_N \end{bmatrix} \tag{3.2}$$

3.2.4 Scanners

Scanners are good sources for obtaining high resolution colour images of planar and non-moving objects, like textiles, paper, wood, X-ray images etc. In many cases they are used in combination with e.g. a high quality reflex camera to obtain high resolution colour images of non-planar objects. In this way photographic 35mm positives/negatives can give good images up to about 6000 x 4000 pixels where the grains on the films set the limits.

Scanners make colour images in various ways, some flash with coloured light using a monochromatic sensor array where the coloured light can come from diodes or a colour-wheel (the Eskofot 2540 uses a multi-spectral colour-wheel),

and some use white light with filters in front of the sensors. The handling of the colours must therefore be decided from the technical specifications of the scanner.

When working with scanners it should be noticed that there is a risk of chromatic displacement between R, G and B channels besides the blurring when the object is not close enough to the flatbed (scanner surface), this has been observed in this project when scanning dia-positives inside slide frames.

3.2.5 Colour Alignments

Most digital colour cameras have the possibility of adjusting the levels of the Red and Blue channels through two gain adjustments or a direct digital control system of the camera. This is used for a *White Balance* adjustment, where the levels of red and blue are fitted to the green level either from a camera specific mode or when pointing the camera at a white surface. The *White Balance* can normally also be set to factory pre-set values, often given in a colour temperature. The existence of *White Balance* also indicates the general problem with colour cameras that has been observed widely in this project. The levels of the channels are not in general very well adjusted and for high quality measurements on e.g. colour textures they subsequently must be aligned using at least an expression as in (3.3). The work with colour cameras in this project has shown that a reasonable alignment of the RGB channels requires such a six parameter expression. The difference in the level of the channels comes from both the digital camera and the subsequent AD-conversion on the frame-grabber if using an analog camera, it is therefore best to do the alignment directly from the scene to the digital images in the computer memory instead of e.g. using an oscilloscope directly on the camera output. The aspects of how the spectral distribution of the light on the scene will influence this will be discussed later.

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} a_{11} & 0 & 0 & R_0 \\ 0 & a_{22} & 0 & G_0 \\ 0 & 0 & a_{33} & B_0 \end{bmatrix} \cdot \begin{bmatrix} R_{in} \\ G_{in} \\ B_{in} \\ 1 \end{bmatrix} \tag{3.3}$$

To estimate the six parameters $a_{1,}$ $a_{2,}$ a_3, R_0, G_0 and B_0 in (3.3) two measurements of each of R_{in}, G_{in} and B_{in} with corresponding known RGB values must be made. At least two different colours must therefore be presented to the colour camera, in this project two different approaches have been used. First two images of a white and a black NCS sheet. Second and more complex capturing an image of a calibration sheet with black disks on a white background and then using an automatic detection routine (described detailed in chapter 4) to detect the black disks and a white area between them.

If the spectral distribution of the RGB filters in the camera is bad for visualization or further calculations (e.g. a red, yellow and blue filter set) the RGB values can be linearly transformed using an expression as in (3.4). To estimate the 12 parameters at least four different colour samples are required.

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix} \cdot \begin{bmatrix} R^* \\ G^* \\ B^* \\ 1 \end{bmatrix} \tag{3.4}$$

The transformation from multi-channel images to RGB images is given in (3.2).

3.3 Chromatic Aberrations

In addition to the monochromatic aberrations described in chapter 2 there are two chromatic ones called transverse (or lateral) and longitudinal aberrations, They are caused by the different refractive index for different wavelengths. These aberrations can be particularly dangerous if the camera is sensitive to IR

and/or UV light since these wavelengths (usually) are the extremes. It is possible to get lenses with e.g. a 400-1000 nm range optimization. Common for the chromatic aberrations is that they cause blurred images with a monochrome camera in multi-spectral light (like white light). It is also important to notice that e.g. yellow light (for the human eye) can be both monochromatic or a blend of red and green, causing the yellow light to have either no chromatic aberrations or significant chromatic aberrations.

3.3.1 Transverse chromatic aberration

The transverse (lateral) chromatic aberration (TCA) is caused by the rays of different wavelength not being focussed at the same point in the image plane. This is shown in figure 3.10 where a violet and a red ray is drawn. The TCA can also be explained as the focal length being different for different wavelengths. If the TCA can be estimated from e.g. a calibration target, it can easily be calibrated for using e.g. a warping. The TCA is zero at the principal point and increases towards the borders. The TCA can be seen near the borders in a colour image where edges will be coloured in fake colours (rainbow-like at black/white edges).

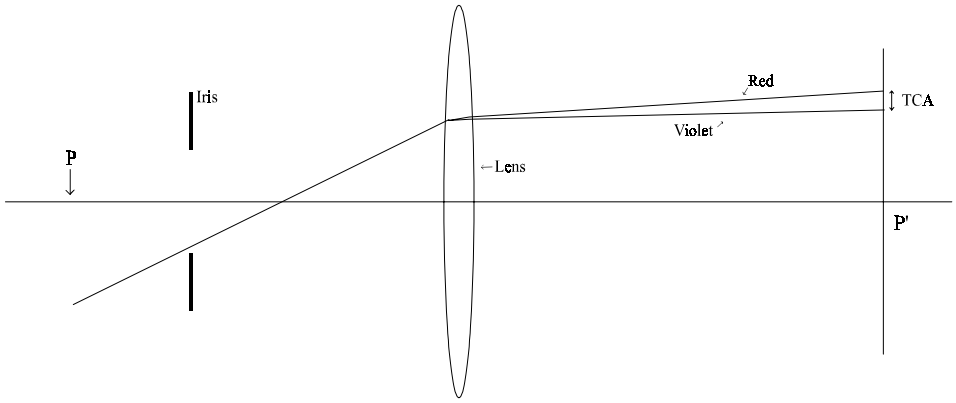


Figure 3.10 TCA for red and violet ray.

When using beam-splitter based cameras it can be very difficult to determine if the spatial misalignments between the channels come from the beam-splitter or the TCA from the lens. The practical work made in this project has shown that the beam-splitter based TCA is more significant than the TCA from the lens. The TCA from the lens has only been the largest when using lenses with extension tubes in close range images.

3.3.2 Longitudinal chromatic aberration

The longitudinal chromatic aberration (LCA) is caused by the rays of different wavelength not being focussed in the same plane. This is shown in figure 3.11 where a violet and a red ray is drawn. When using a lens with significant LCA it will be impossible to have all channels in focus at the same time.

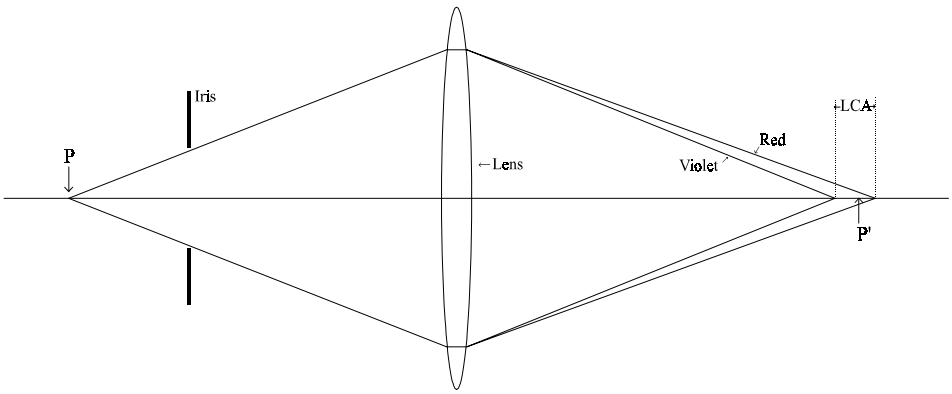


Figure 3.11 LCA for red and violet ray.

The chromatic aberrations have especially been observed with cheap lenses and lenses where the focus has been changed using an extension tube. The problem with extension tubes is that they move the characteristics of the lens away from the designer’s optimizations. In this project the work with colour texture

measurement at distances from 200 to 500 mm has given problems with the lenses not being able to focus at the small distances. For this purpose the normal procedure is to use extension tubes, but the experience from this project has shown that this is not always a good idea because of significant amounts of TCA and LCA, where only the TCA can be corrected by e.g. a warping. In this project the TCA has been observed up to a maximum of 2-3 pixels between the red and the blue channel. If using good lenses designed for these short distances very small amounts of chromatic aberrations have been observed.

Diffraction is also wavelength dependent and has a chromatic difference. In most cases this should not cause any problems since it is symmetrical (different sizes of the Airy disk).

3.4 Full camera calibration

In the classical models from photogrammetry the issue of colours and chromatic aberrations is not considered. If however, colour images are to be used in high accuracy measurements there are some aspects that are useful to consider.

If working with the linear models (DLT) there should not be any reason to consider the chromatic issues with regards to the estimation of the external position of the camera since the accuracy of the DLT normally is low compared to the chromatic errors. It should therefore be sufficient to e.g. calculate an initial guess for the non-linear search from a DLT based on e.g. the central channel.

If the purpose of a non-linear search solely is to get the external position, rotation and the focal length of the camera the best is to perform the search on the positions of calibration objects found in the central channel (e.g. green), since it with most colour cameras and lenses is the most accurate. When this is

combined with fitting the other channels to the central using e.g. warping this is in many cases the most attractive method.

When using single chip or multiple shot cameras it is possible to make the non-linear search for all channels in one step with the constraint that the six external parameters (translation and rotation) are equal for all channels. This should give the best estimates since the position and translation obviously are equal for all channels. Another thing is that this ensures that the differences in focal length for the channels are not absorbed by the strong correlation between the focal length and the distance between calibration object and camera. With a three channel system like RGB such a search would include six external parameters and e.g. seven intrinsic parameters for each channel, this gives a total of 27 parameters to be estimated which is quite reasonable.

Using a beam-splitter based camera (e.g. 3xCCD) it is in general not possible to assume common external parameters for each channel since the chips mounted on the beam-splitter do not necessarily have the same rotation and translation. With a beam-splitter camera the best is therefore either to make a full non-linear search for each channel individually or instead only to make the full non-linear search on the central channel (e.g. the green) and then afterwards to fit the other channels to this one using e.g. warping (see next section).

3.5 Warp based colour calibration

As described in chapter 2 on monochrome calibration there are scenes or setups that make it impossible to gain 3D information from a calibration target. This is often the case in close-range images where the depth of field is very low or when using devices like microscopes, scanners or special lenses (e.g. telecentric). In these cases the classic nonlinear calibration cannot be used and the use of warping (geometric transformation) is a very good alternative.

A warp which is a 2D transformation from one coordinate system to another

can be defined through the use of 2D-polynomials, the general transformation of order N is given in expression (2.36) in chapter 2.

Warping colour images can be done in several ways depending on the task. For the simplification the following is mostly regarding RGB images, but it can easily be extended to multi-spectral images.

If the task only is to get images without TCA aberration and beam-splitter misalignments (or IFC filter based misalignments), warping can be used to transform all channels to one central channel, that includes the wavelength that the optical system is optimized for, typically the green channel in RGB images. This does not involve any perspective transformation but only the spatial difference between the channels and can therefore successfully be modelled with a polynomial based warping. See chapter 2 regarding how the normal lens distortions successfully can be modelled with warping if omitting the perspective transformation, the warping can therefore also model the differences in the lens distortions arising from different wavelengths. The practical work done in this project with warping of the red and blue channels to fit the green channel (using 3xCCD cameras) has shown that a 2^{nd} order warping in most cases is enough and that only a little improvement is achieved when using a 3^{rd} order. In the table below is the displacements (average and maximum) in pixels between the red, green and blue channel from the image in figure 5.3 using a 1^{st}, 2^{nd}, 3^{rd} and 4^{th} order warping to fit the green channel (based on 494 points). As seen in the table there is no exceptional improvement (<8%) between a 2^{nd} and a 4^{th} order warping and in order to avoid fitting to the measurement errors a 2^{nd} order warp must be considered the best. Notice that the maximum displacement is almost fixed at about 0.3 pixels when the warping order is above 1, which most likely originate from the measurement errors.

Channels	**Average** (pixels)					**Maximum** (pixels)				
	Orig.	1st	2nd	3rd	4th	Orig.	1st	2nd	3rd	4th
Red/Green	0.23	0.12	0.1	0.1	0.1	0.46	0.38	0.29	0.31	0.3
Red/Blue	0.41	0.14	0.1	0.1	0.1	0.91	0.53	0.27	0.23	0.23
Blue/Green	0.54	0.12	0.11	0.1	0.1	0.88	0.35	0.3	0.27	0.26

Table 3.1 Average and maximum displacement in pixels between RGB channels in the original image and the 1st, 2nd, 3rd and 4th order warped image (displacement is the Euclidian distance between the centres of the 494 calibration objects).

When using warping in a transform involving perspective transformation it can be done in two ways: warping each channel individually or warping the central channel combined by first warping the other channels to the central channel. The advantage of the latter is that a higher order (more coefficients) warping can be used for the central channel and lower order (less coefficients) warping for the spatial alignment of the other channels. This could be good when either the computational speed is important and/or the positioning of calibration objects in the central channel are better than the other channels. As described in chapter 2 a warping with too many coefficients will fit to the measurement error of the calibration object, which is a good reason for using many coefficients with the warping of the central channel (which in many cases is the best) and less coefficients with the warping of the other channels to fit the central.

3.6 Uneven intensities

When obtaining an image of a uniform surface like a white wall, a paper sheet etc the intensities in the images are in general non uniform. Such uneven intensities in an image plane can either come directly from the scene or inside the camera or a combination of both. In many cases uneven intensities in an image are unwanted and prevent further image analysis. This section will describe these issues with respect to the cause and effect and how to calibrate for it.

Some issues of the uneven intensities are (almost) common for all wavelengths and must be considered with both colour and monochrome images. These are therefore topics that also are interesting for the use of monochrome cameras. The uneven intensities that are wavelength dependent (different unevenness between the channels) also causes problems with monochrome cameras but they are almost impossible to correct unless a significant amount of knowledge of the scene, light sources and objects is present.

This relatively large section has been placed in this chapter since the calibration of colours, as it will be explained in this section, in many cases is closely related to the efforts to get evenly intensities in all channels. In the systems for colour texture measurements that have been worked on in this project, it has due to the complexity been impractical or impossible not to assemble the calibrations for uneven intensities and the calibration for correct colours in one common calibration.

3.6.1 Camera and lens based

Even when a completely homogeneous planar object is uniformly illuminated with completely diffuse light, the image (using an ordinary camera and lens) of the object will still not be uniform in the intensities. This phenomena is normally caused by two different effects: light falloff ($\cos^4$law) and vignetting.

Light Falloff

Light Falloff by the 4th power of cosine to the angle of view (α) is a very well-known and fundamental optical effect that is a combination of several effects, one of the reasons is that an off-axis object sees a foreshortened apparent aperture (entrance pupil) so less light is collected from an off-axis object, this gives a *cos(α)* effect. The lens collection C for a single element lens can be expressed as in (3.5) where a is the diameter of the lens, f is the focal length and it is assumed that the object to lens distance is much larger than the lens to image plane distance [Haralick93].

$$C = \frac{\pi}{4}\left(\frac{a}{f}\right)^2 \cos^4(\alpha) \tag{3.5}$$

The light fall-off is wavelength independent, but a small and in most cases negligible linear difference can come from the difference in the wavelength dependent focal length.

In figure 3.12 is a plot of the angle of view versus the light falloff, as it is seen the falloff is very large for wide angle lenses, when using a 65° wide angle lens the light at the edge will be 50 % lower than at the center.

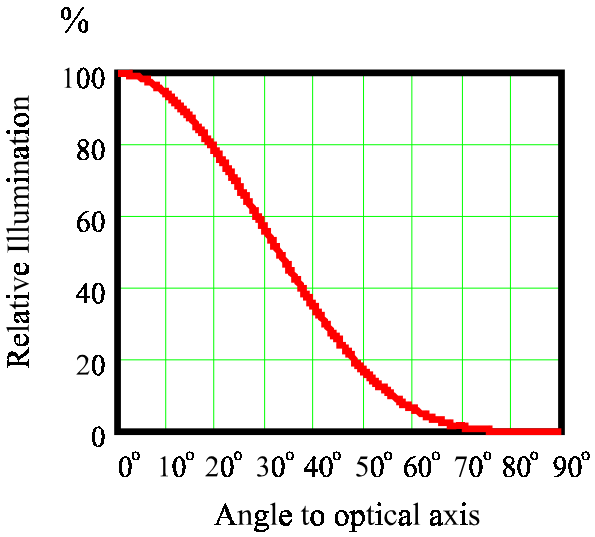


Figure 3.12 Non-uniformity of illumination caused by the cos⁴law.

When using a lens with a diaphragm between two lens groups, it can be optimized for no Light Falloff [Walter95] but the condition for this optimization is in contradiction with the Abbe sine condition and some aberrations must be accepted if the Light Falloff should be reduced. Another method is to use a filter in the lens that spatially reduces the light accordingly to the 4^{th} power of cosine.

In an experiment where the chip and viewport are of fixed sizes, the Light Falloff can be reduced by using a lens with higher focal length and moving the lens/camera further away from the object. This is shown in table 3.2 with a 1/3" CCD chip and a wanted viewport in object space of 100mm x 75mm, the result is a reduction of the Light Falloff effect from 6.9% to 2.9% when using a 25 mm lens instead of a 16 mm.

Lens (focal length)	Lens to object distance	Angle of View (deg)	Maximum light falloff (diagonal).
16 mm	327 mm	21.7^o	6.9%
25 mm	510 mm	14.0^o	2.9%

Table 3.2 Light Falloff using 16 mm and 25 mm lens (1/3" CCD) with a 100x75mm viewport.

Vignetting

Vignetting is a reduction in the light going through the lens caused by physical obstructions, this is typically caused by a bad composition of diaphragms and other lens elements, well designed lenses can be without significant vignetting. For a lens with vignetting it can be reduced by using a smaller aperture. Figure 3.13 shows an illustration of the effect of vignetting using a single lens with a diaphragm in the object space. This figure also gives an idea of why a smaller aperture reduces vignetting, if imagining the bottom and top rays are removed by a smaller diaphragm.

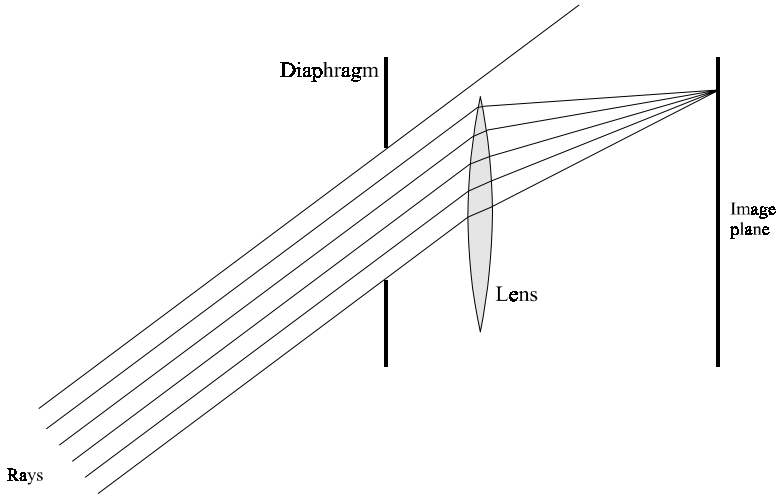


Figure 3.13 Vignetting in a single lens system.

Vignetting is of course more complicated for lenses with the diaphragm between two lens groups than shown in the simple example in figure 3.13, but the principle is the same. Vignetting has been observed with many different lenses in the work done in this Ph.D. project, and the experience is that there is a high correspondence between the price of the lens and the amount of vignetting. An expensive Schneider-Kreuznach lens used in this project has showed no measurable vignetting.

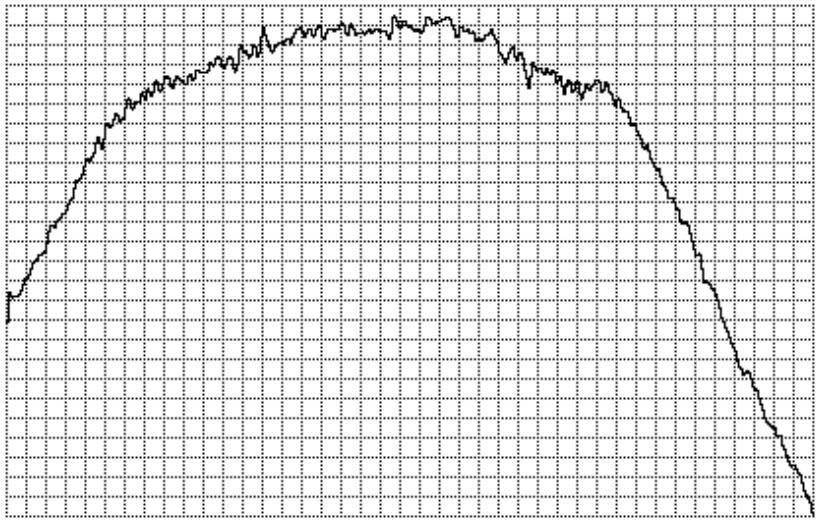


Figure 3.14 Stretched horizontal intensity profile of image of an almost complete uniform surface in an integrated sphere. The lens was a no-name Japanese 50mm TV lens.

In figure 3.14 is shown a horizontal intensity profile using a monochrome JAI M50 camera with a no-name Japanese 50 mm TV lens. The object, a NCS 5000-N sheet, was placed inside an integrated sphere with a diameter of 360 mm, the distance between the camera and the NCS sheet was 400 mm. The NCS sheet was through the use of the integrated sphere completely uniformly illuminated compared to the other error sources. As seen in the figure, the stretched intensity profile is not uniform, in order to make sure that these non uniform intensities did not come from the sphere illumination the camera was moved which made no difference in the intensity profile.

The profile in figure 3.14 is a special case where the light fall-off ($\cos^4$ law) is overlapped with a vignetting effect which starts at a certain angle (distance from optical axis). This is particularly seen with this no-name Japanese 50 mm TV lens where the intensity profile reveals an obvious combination of two error sources. When using a high quality lens in the same system or reducing the aperture the vignetting effect disappeared.

A beam-splitter can also cause uneven intensities in the image, which usually are different for each channel, this is called colour shading see page 72 for more details. In figure 3.15 and 3.16 are the red, green and blue channels from a JVC KY-55 3xCCD camera blurred and stretched between black and white. The RGB images were captured in an integrated sphere similar to the one used in figure 3.14. As seen in the figures the colour shading causes significant non uniform intensity images, it is the high difference in intensities between top and bottom that causes the visible colouring (mostly red/green) of e.g. a white surface that legitimate the name colour shading. The $\cos^4$ light fall-off is insignificant for non wide angle lens compared to the colour shading, while vignetting can be quite significant.

Figure 3.15 R, G and B Color Shading with Cosmicar 25 mm lens and JVC KY-55 3xCCD

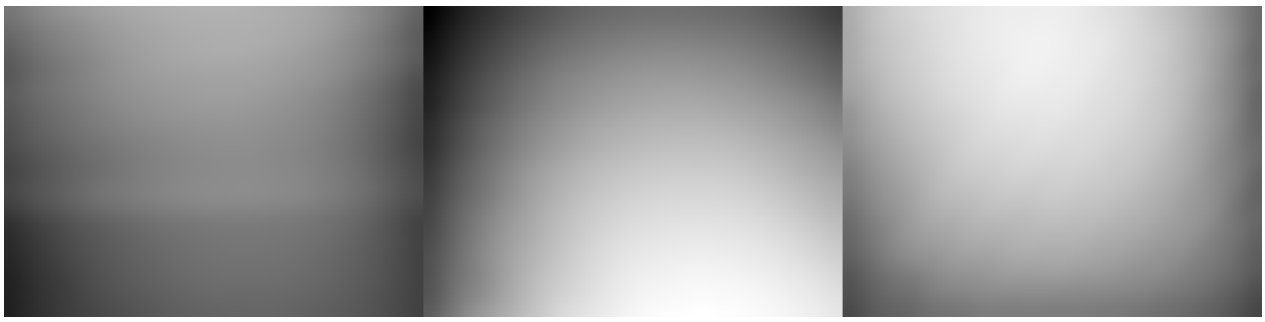


Figure 3.16 R, G and B Color Shading with Jena 25 mm lens and JVC KY-55 3xCCD

As seen in the figures 3.15 and 3.16, where the only difference is the change of lens, colour shading is very dependent on the lens. As seen (can be hard on the printed version of this document) the figures also contains some minor errors in the red and blue channels, the red channel in 3.16 has some degree of “prism” flare with the physically longer Jena lens and the blue channel has some electronically generated stripes.

3.6.2 Scene and light source based

In most cases the uneven intensities in an image of a uniform surface are due to lighting issues like the positions/sizes of the light sources and the optical (reflectance/absorption) properties in the scene.

In practical image analysis many efforts are made to obtain an even distribution for the irradiance. Examples are optical light rings and plates illuminated

through optical fibres, another common method is to use many small sources around the objects or a light source behind white transparent material. Commercial products from e.g. Fostec Inc. have planar plates up to 20x20cm for background illumination that are guaranteed to be uniform within ± 5%.

The description of these aspects is divided into two different optical cases: open scenes and closed scenes. Although there is no clear definition of the boundary between an open and a close scene (almost all scenes are optically closed in some sense), this division has been the most natural seen from this project.

Optically open scenes

In this project the definition of optically open scenes is:

1. Scenes where a substantial part of the light rays reflecting/emitting from the object of interest will not be reflected back on the object itself.

And

2. Scenes where external (uncontrollable) light sources can illuminate the object.

Common examples of such optically closed scenes are outdoor scenes like a camera mounted on a tractor driving on a field measuring the plants for e.g. dosing of pesticides. Although, driving in a cloudy night, when the moon is down or using strong (compared to the moonlight) local light it is possible to have full control over the light sources but it is not granted that the clouds will go away etc.

Whether or not a closed scene can be regarded as an open scene depends, of course, on the relation between the size of the object/camera and the scene. E.g. looking at 20x20 cm textiles in a normal room/laboratory could be considered an open scene, even if there is no windows in the room.

The main disadvantage with open scenes is that there is not full control over external light sources. An advantage is that the object is not significantly illuminating itself through the light reflected/emitted from the object (self-illumination).

The open scenes are very common in image analysis, even with systems requiring high accuracy, because of physical limitations that e.g. makes a shield against external light sources practically impossible or too expensive. As an example of an open scene case can be mentioned a project at AUC (Aalborg University) where a 3xCCD camera is used for measurements on barley in a field in Denmark. Here they say in their paper [Andersen99] *"A problem often limiting or causing colour image analysis methods to fail is the lack of ability to handle images acquired under varying illumination conditions"* which also is one of the conclusions in this project. Another project is the beet reception at Danisco Sugar, where an initial vision system based on outdoor stereo 3xCCD cameras has been tried unsuccessfully on loads of beets on trucks [Frydendal98], the CCD cameras has been moved inside for better shields against uncontrolled illumination.

In an open scene the main part of the scene based on uneven intensities comes from the light sources or local objects that reflect the light towards the object of interest. This is illustrated in figure 3.17 (a) where an object is illuminated from a single point light source (the lightbulb), the left side of the object will then be brighter in the image than the right side due to the smaller distance to the light source. In figure 3.17 (b) is illustrated the same situation where a wall reflects more light down on the object, this reflection consists of both specular and diffuse reflection (see figure 4.34), where only the specular is drawn in the figure. The diffuse reflection will also give a contribution to the uneven illumination in figure 3.17 (b) since the distances between the light source, wall and object not are equal for the entire object.

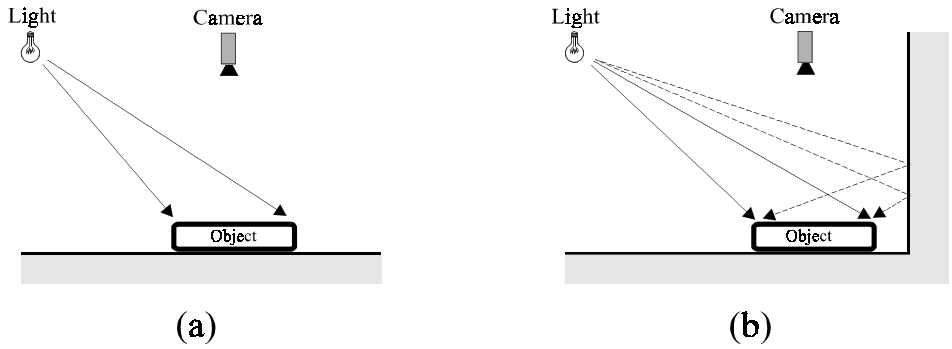


Figure 3.17 Uneven light on object from single light source. (a) free air scene (b) scene with reflecting wall.

With a totally open scene, like in figure 3.17 (a), two or more different light sources can beside the uneven intensities cause an uneven spectral illumination if the light sources are not totally equal in their spectral distribution, even two equal light bulbs can have a small difference in the spectral distribution arising from small differences in the production. With an open scene like in figure 3.17 (b) a single light source system can also cause an uneven spectral distribution if the wall does not reflect/absorb all wavelengths equally.

Since planar surfaces have been the most used in this project (and are in many other projects), the next pages will discuss a very simple “ray-trace” of the irradiance of a planar surface from one and two single point radiators (light sources), this is done under the assumption that there is no specular reflectance.

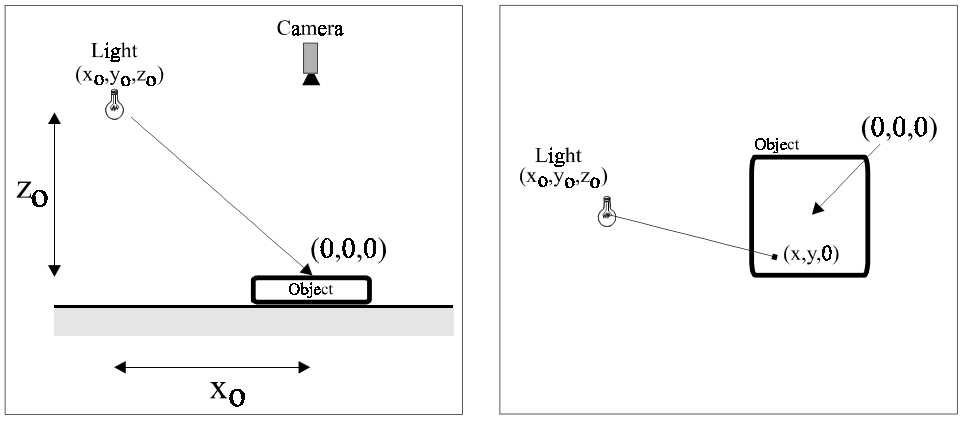


Figure 3.18 Single point radiator and planar surface, seen in two planes.

Using the notation giving in figure 3.18 the distance *d* between a point *(x,y,z)* and the single point radiator at (x_o,y_o,z_o) will be as in (3.6).

$$d = \sqrt{(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2} \tag{3.6}$$

In (3.7) we have the surface irradiance I from a single point radiator I_0 , where θ theta is the angle between the surface normal and the point source [Haralick93]. This shows that the irradiance varies inversely as the square of the distance between illuminated surface and the point source.

$$I = \frac{I_0 \cos(\theta)}{d^2} \tag{3.7}$$

Using the setup and notation from figure 3.18 θ can be calculated as in (3.8).

$$\theta = \cos^{-1}\left(\frac{z_0}{\sqrt{(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2}}\right) = \cos^{-1}\left(\frac{z_0}{d}\right) \tag{3.8}$$

Inserting θ from (3.8) in (3.7) the irradiance *I(x,y)* on a planar surface from a single point light source will follow equation (3.9).

$$I(x,y) = I_0 \frac{z_0}{d^3} = \frac{I_0 z_0}{((x - x_0)^2 + (y - y_0)^2 + z_0^2)^{3/2}} \tag{3.9}$$

In figure 3.19 is an example of an irradiance on a planar surface following equation (3.9), where (x_o,y_o,z_o)=(500mm,0mm,600mm) and the surface is 1200x1200mm.

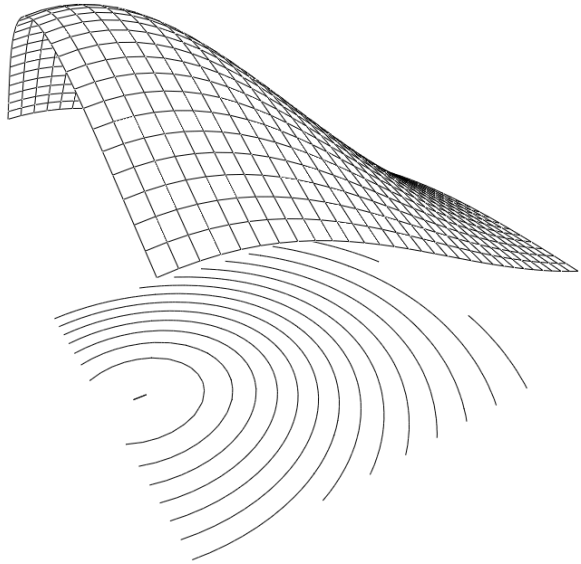


Figure 3.19 Irradiance distribution at planar surface from single point radiator (plotted as 3D surface and contour below).

If applying two or more light sources the resulting irradiance is the sum of the irradiance contributions. In figure 3.20 is the result of applying two light sources symmetrically across a planar surface where one of the light sources is the double intensity of the other.

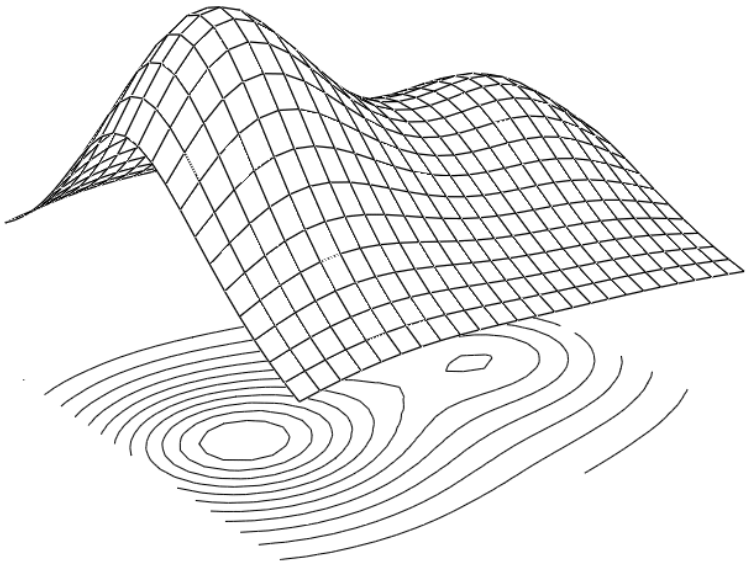


Figure 3.20 Irradiance on a planar surface from two symmetrical single point radiators, one is double the intensity of the other.

As seen from equation (3.9) an almost uniform irradiance on a planar surface requires a single point radiator at a very long distance. Since a radiator at a long distance requires very high energy to illuminate the surface another common method is to use multiple radiators or large (non single point) radiators.

The model for planar surfaces described above is only concerning the distribution of the light falling on the surface (irradiance) and not the light emitting from the surface to the camera. If the camera is placed so that no specular reflection can reach the lens the above model can be used for the relative distribution of the radiance.

In many cases it is wanted to model the intensity distribution using a least a square polynomial fit. This has been done for the single radiator distribution given in figure 3.19 and the dual radiator distribution in figure 3.20 using the software *MathCad* from Mathsoft. In table 3.3 is given the coefficient of determination (R^2) for order N=1..6.

N	Single radiator (fig. 3.19)	Dual radiator (fig. 3.20)
1	0.83068	0.17069
2	0.95957	0.84438
3	0.99781	0.8977
4	0.99904	0.97096
5	0.99994	0.97978
6	0.99999	0.99516

Table 3.3 Coefficient of determination R^2 for a polynomial fit of order N to the irradiance using one and two radiators.

Notice that with more than one radiator a 3rd order polynomial can easily be inefficient. That is also one of the reasons why an estimation of a intensity image can be more useful than a polynomial fit.

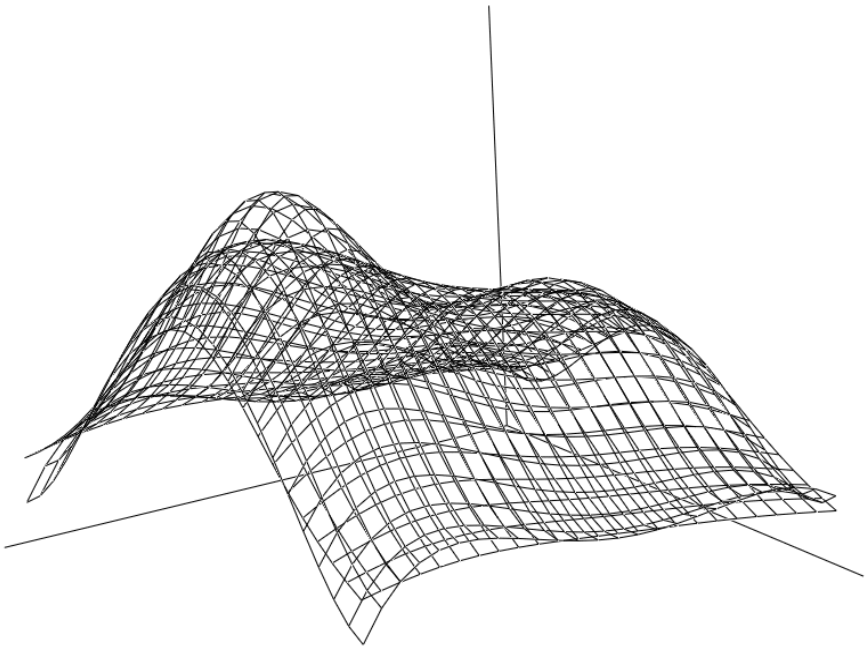


Figure 3.21 A 4th order polynomial fit to a two radiator irradiance distribution on a planar surface (same as figure 3.20)..

In figure 3.21 is shown the 4th order polynomial fit to the two radiator irradiance, as seen the fit is not very good especially at the two tops, the coefficient of determination R^2 was about 0.97. In order to get a good visual fit at least a 6th order polynomial was necessary.

In many cases the radiators (light sources) are placed in such a way that there are no tops in the area of interest (active image area) and a smaller order can be used.

The above model is, of course, very simple, without shadows and specular reflections. It is possible to make more advanced models (ray-tracing) that include specular reflections, shadows based on the knowledge of the scene setup, although this is beyond the scope of this project.

In many cases the environment is much more complicated than illustrated in figure 3.17 and modelling through e.g. ray-tracing cannot be done. In such

cases it is only possible to capture images of well-known even coloured calibration targets and then calculate a calibration for that specific scene. When the camera or object is moving it becomes even more complicated and an online running calibration is required (see next section). In a complicated environment it is normally not possible to differentiate the various contributions to the uneven intensities in the image, unless camera specific factors like vignetting and colour shading are determined in a laboratory.

For special applications like a tractor on a field driving in an illumination which varies through the day there is a method for making images that approximates images captured using fully controlled light. In order to do this two images are captured immediately after one another (like 1/25 sec with a PAL camera) where one is illuminated by the natural background illumination (global) and the other image is illuminated by both the background and a local light source. Then the image without the natural background illumination can be estimated from the two images. If the image intensities do not contain any nonlinearities and saturations and the offset is zero this can be done with a simple subtraction. In figure 3.22 (a) is the image of an office lamp with the normal room lighting and in (b) is the exact same scene with a small local lamp, notice that the local lamp does not influence the appearance of the background.

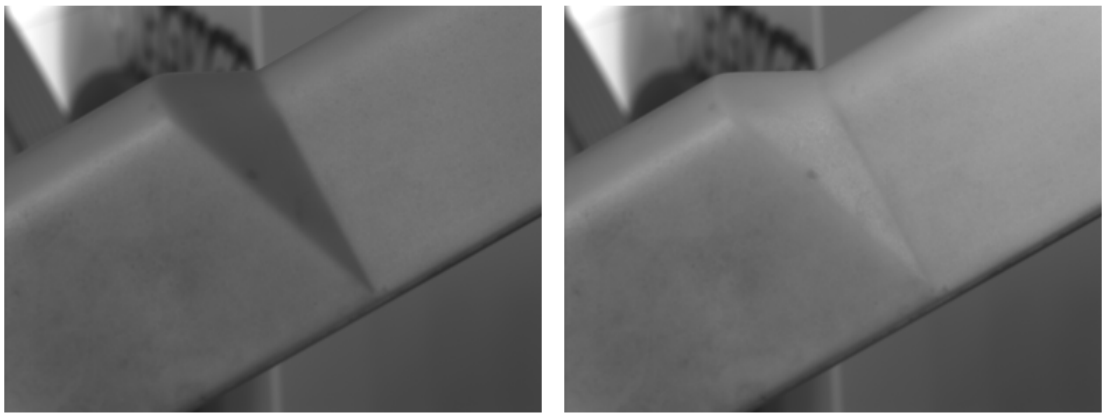


Figure 3.22 An office lamp. (a) illuminated by the normal office lighting and (b) illuminated with an extra local lamp in a steep angle.

In figure 3.23 (a) is the same scene as in figure 3.22 but only illuminated with the same small local lamp as in figure 3.22 (b). In figure 3.23 (b) is the difference between the two images in figure 3.22, notice the visual good similarities between (a) and (b). In order for the subtraction to work the images have to have an offset of zero, this was done by placing a cap in front of the camera and the “zero-image” was captured as an average of 50 frames, this had a mean value of 5.95. This “zero-image” was then subtracted from the images to give the common zero offset. The mean value of the difference between the two images in figure 3.23 was -0.12483 and the variance was 0.88756, which is very good for a small test made in an office with an ordinary CCD camera. If looking closely at these images it is noticed that there are small systematic differences between figure 3.23 (a) and (b) these could be explained by some sort of non linearity in the measurement of the radiance, vibrations/movements and perhaps some interactions between the light sources and the materials.

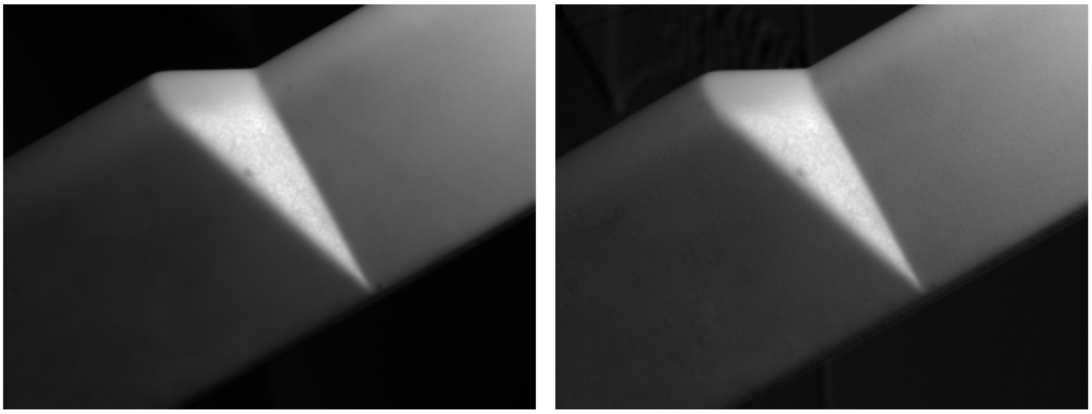


Figure 3.23 An office lamp. (a) only illuminated from a small local lamp at a steep angle. (b) the difference between the images in figure 3.22.

This technique for omission of the background illumination has many usabilities, especially with structural light (like lasers) and in outdoor scenes or rooms with windows. A disadvantage is that it requires either the average of multiple frames or a low noise and high sample rate to generate good images, since the subtraction (within the dynamics of the camera/grabber) enhances

noise. The big advantage is the possibility to estimate a simple and controlled illumination that e.g. can be ray-traced from the known positions of the scene elements (light source, camera, object etc.).

Optically closed scenes

In this project the definition of optically closed scenes are:

Scenes where there is full control over the light sources that illuminates the object and no significant amount of light can escape the scene without being absorbed by material within the scene.

In such scenes there can be setups that cause a substantial part of the light rays reflecting/emitting from the object of interest to be reflected back on the object itself (self-illumination). The main reason why self-illumination is a problem is that a calibration based on a reference object not necessarily will calibrate a subsequent object correctly.

Examples of closed scenes are shielded outdoor scenes (e.g. a tent around a camera setup on the back of a tractor in a field), a laboratory room with shielded windows, a closed box with a camera and a light source inside etc. In figure 3.24 is an example of a closed scene in form of such a box. One light ray emitting from the single point radiator is drawn with its specular reflection and the diffuse reflections where the ray hits a surface. If all the light emitting from the light source with all reflections should be modelled it becomes quite complex even for a simple scene like figure 3.24 and a computer generated ray-tracing is the only (known) possible way to model the radiance from the object to the camera lens. Such a ray-tracing is only possible if the radiometric characteristics are known for all the surfaces and objects in the scene.

If ray-tracing is not possible due to e.g. missing information regarding the materials or the scene is to complex closed scenes have the advantage that it is possible to calibrate for the illumination of one or more reference objects (e.g. colour reference sheets like NCS) and then use that calibration for subsequent

measurements. The disadvantage of this method is that the reference object must have approximately the same characteristics as the objects to be measured, since the self-illumination otherwise could introduce an error, if no calibration is made.

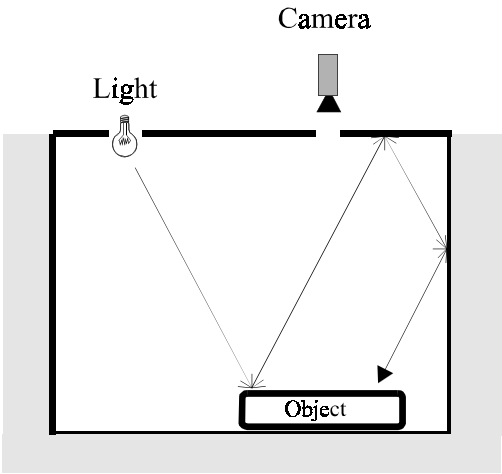


Figure 3.24 Single light source in optical closed scene, the line is the specular reflection. The self-illumination is caused by both the specular and the diffuse reflection.

With closed scenes the materials that make the surroundings have a big effect on the self-illumination. These materials can vary from black velvet to diffuse white painting and even to mirrors. With almost totally absorbing materials like black velvet the self-illumination can be zero if the geometry of the object does not allow self-illumination without the presence of other objects/walls. The reflection characteristics of the material also has a big influence on the self-illumination, it can be complete diffuse (white diffuse paint) or completely specular (mirrors) or something in between. The self-illumination also has a spectral characteristic that follows the reflection/absorption in the scene (including the object). As an example can be mentioned a box with walls that tend to absorb wavelengths below 650 nm (red) more than others which will give a reddish self-illumination. A textured object that absorbs certain wavelengths placed in a diffuse white box will also give the self-illumination a spectral displacement. This could be the case with a red textile with small green stripes, where a reddish self-illumination would cause an intensity displacement between the red and green areas in the images.

It is therefore very important for an accurate colour/intensity measurement on

an image made in a closed scene to calibrate for the self-illumination, since it not only can affect the intensity distribution and the general level but also introduce a difference in the relations between differently coloured areas in the image.

3.6.3 Correction and Calibration

The calibration for uneven intensities in an image is in many situations an important task of the image analysis, and descriptions can be found in many places in the literature, e.g. [Russ90]. A simple example is if a threshold must be applied to get binary objects for e.g. counting and area/perimeter/center of mass measurements, in such cases uneven intensities can cause an incorrect binarization. Using a local threshold (based on percentage of the neighbourhood) can make it better but not correct. As described in the next chapter uneven intensities can give systematic displacements in detection of e.g. a center of mass, which is particularly serious with images of calibration objects.

Cos^4 Light Fall-off

As described previously the cos^4 law light fall-off can be calibrated using a specially designed filter but this is very rare and probably also very expensive. In most cases it must be calibrated from the knowledge of the camera and lens. The classic camera calibration (either DLT or full non-linear) gives the principal point (u_o, v_o) and correct focal length f of the lens from the physical size of the film/chip $sx \times sy$, that normally can be obtained from a data sheet. In (3.10) is given the expression for intensity I for the cos^4 fall-off at a point (u,v), where the intensity is I_o in the principal point and that all elements must be in the same unit (e.g. pixels or mm). This expression has been reduced to the lower part without the $cos^4()$ and $atan()$ using MathCad, which is an much easier expression that can be calculated much faster on a computer.

$$I(u,v) = I_0 \cos^4\left(\operatorname{atan}\left(\frac{\sqrt{(u_0-u)^2+(v_0-v)^2}}{f}\right)\right)$$
$$= \frac{I_0}{\left(1+\frac{(u_0-u)^2+(v_0-v)^2}{f^2}\right)^2} \qquad (3.10)$$

In figure 3.25 is a plot of expression (3.10) with a 25 mm lens and a 1/3-inch format (6.00 x 4.96mm) chip, the principal point is set to the center of the chip.

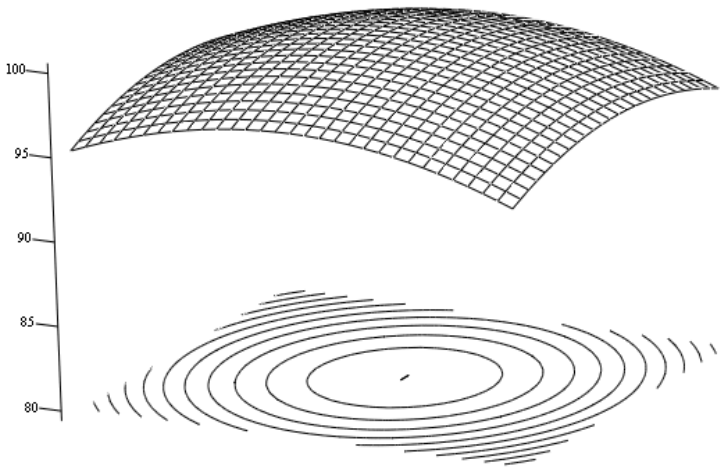


Figure 3.25 Cos^4 fall-off with 25mm lens and a 1/3-inch CCD chip. The intensity is set to 100 in the principal point.

Vignetting

As described previously vignetting can not normally be modelled directly since it is very dependent on the size of the aperture, the best way to “calibrate” for vignetting is to buy a high quality lens or to use a small aperture. If a setup requires a lens system that has vignetting, the calibration can be made from an image of a reference object. In such a calibration the contribution from vignetting can not necessarily be extracted from all the other contributions to the uneven intensities and a calibration must therefore include all (or some) contributions in a common model.

Colour Shading

As described previously the use of beam-splitter based cameras introduces an intensity error called colour shading that is different for each channel. The colour shading that has been observed in this project has been overlaid with other error sources giving a very unstructured uneven intensity, see figure 3.15 and 3.16. Is has been tried to calibrate the colour shading and the other contributions using a polynomial fit, but the conclusion has been that this is not sufficient. It is therefore necessary either to use a much more complex model or as implemented in this project to use images of uniform reference objects. Such images also calibrates for all other sources for the uneven intensities.

The calibrations made in this project have used images of both very dark, grey and bright NCS-sheets in a very uniform light to estimate the Colour Shading and $\cos^4$ light fall-off and subsequent images calibrated from these images has shown very small variations of about 1-2 grey levels (< 1%). This is described in more detail in chapter 5 (Implementations).

Generic Camera/Lens based

In many cases it is easiest and best to make the calibration for the camera/lens based uneven intensities (vignetting, $\cos^4$ fall-off and colour-shading) as one common calibration for a specific camera/lens setting. This can be done in a laboratory or calibration box using a target with uniform radiance. Such a target could e.g. be an integrated sphere with diffuse painting inside or a planar background illumination plate. As with the classic calibrations it is best to grab an average image of multiple images but unlike geometric calibration it is best to move the camera to smooth small variations in the target. In theory it should be possible to move/rotate a camera more or less randomly around in a scene while averaging the images continuously to get a uniform image to model the sum of the camera/lens based uneven intensities, but small experiments made in this project have shown that this requires a LOT of images to give a good result.

The detection of the unevenness from a target image can be done in three ways:

1. Fitting a suitable 2D function (e.g. a polynomial least squares fit) to the pixel values.
2. Calculate an image (average or median of a sequence) that contains the unevenness
3. Estimating the parameters in a complete parametric model that models the necessary error sources.

Pixel variations

A CCD/CMOS chip can also have variations in the sensitivity for the cells (pixels) and some cells might even be defect. Such variations in the chip can be detected by capturing an image of a smooth radiance image. Such an image can be made by moving/rotating the camera and/or the reference target while averaging the images continuously. The reason why the movement is needed even with a very uniform radiance, as in an integrated sphere, is that scratches,

dust etc on the reference target will overlay the pixel variations. A problem (or advantage) with moving/rotating average images is that they also capture scratches, dust etc on the lens/chip. Another and more simple approach is to adjust the lens to very unfocused, this also has the advantage of not including the scratches, dust etc on the lens and including those on the chip.

One simple method to detect the pixels variations from the smooth radiance image is to apply an local mean/median filter to the captured image, which then is subtracted from the captured image. With colour images this is done separately for each channel.

Scene based

The scene based uneven intensities can either be calibrated simultaneously with the calibration for the lens/camera based uneven intensities or alone. The latter requires that the uneven intensities from the lens/camera are known/measured, see above. This can be a big advantage because it keeps the complexity of the remaining unevenness low, which especially is useful with curve-fitting and modelling.

In both cases the detection of the unevenness can be done by capturing an image(s) of a reference target and fitting a 2D function (curve fitting), estimate an image (or images) that contains the unevenness or make a complete model using e.g. ray-tracing combined with the models for the lens and camera based uneven intensities.

A calibration with a reference target will only work in scenes where the variations in light sources only vary a little bit in small time steps and where there is no significant self-illumination. If the variation in time (moving objects in the scene, conveyor belt etc) makes it impossible to place a calibration target in the scene that can be used for subsequent calibrations of measurements, it can in some cases be possible to make an online running calibration, see e.g. [Marszalec97] for further details on online colour calibration.

This can be done by e.g. placing a border of calibration material (NCS etc) in the scene or simply stripes of calibration material that are visible in the camera view but are acceptable for the measurements, see figure 3.26.

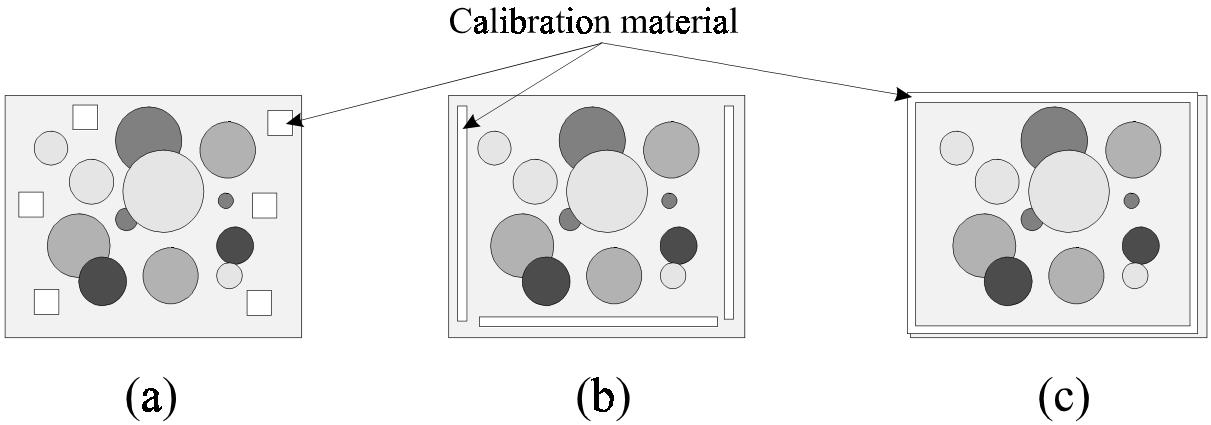


Figure 3.26 Various setups for the calibration material for an online calibration for uneven intensities. (a) using small samples (squares) (b) using stripes (c) using a complete border. The circles illustrates the objects to be measured in the image.

The idea of using stripes like in figure 3.26 (b) and (c) is that the radiance can be spatially followed which can improve the estimate of illumination on the objects compared to using single point marks as in figure 3.26 (a).

In some cases it is possible to use the prior knowledge of a scene or an object to make a calibration for uneven intensities, an example is a project at AUC (Aalborg University) where a 3xCCD camera is used to measure barley on a field in varying daylight conditions, here the calibration is based on the prior knowledge of the spectral distribution of the sunlight (clear sky, clouded, shadows etc) and the plants [Andersen99]. Another example is a geometric calibration target with e.g. black disks on a white background. Here it is possible to detect the uneven intensities based on the knowledge of the target, see chapter 4.

In order to make a proper calibration from one single coloured reference target (like one NCS sheet) it is very important that the real zero is known. If the real zero point is above the pixel value zero it can be measured by covering the camera lens and capturing an black image (the best is an average of multiple

captures). If the real zero is below the pixel value zero such a black image will be zero and the use of multi coloured reference targets is needed (like using two NCS sheets or a mosaic of multiple colours on a single sheet).

Self-illumination

From the experience gained in this project the calibration for self-illumination when using closed scenes like spheres can be made in the following ways;

- By placing separate light sensors inside the closed scene (this could also be an extra camera). The sensors can be online calibrated from e.g. two different known objects placed inside the scene.
- By using small reference coloured objects that always are inside the camera view (see figure 3.26).
- By using small reference coloured objects outside the normal camera view (or the walls of the closed scene) that e.g. a mirror can turn the camera view against or by moving the camera . This requires two image captures, one for the object and one for the reference object(s).
- By using specialized calibration sheets to calculate the coefficients in a self-illumination calibration that can calibrate an image from the contents of the image itself (estimating the current self-illumination from e.g. the mean value of the image).

These methods and the specialized calibration sheets are described in more details in chapter 5.

Temporal

Another common type of intensity variation in images are the temporal variations. In open scenes these could come from moving illumination or reflectors and these require an online running calibration, as described above. In optical closed scenes the temporal variations can be temperature dependent,

in both the light sources and the camera. The variation often occurs in both intensity level and relative difference between the channels (temperature shifts). In many cases the variation is largest immediately after the system is turned on, due to starting fluctuations. Such intensity variations in time have been observed in this project using a highly calibrated system with a camera in an integrated sphere (see chapter 5). The variations were typically of 3 - 4 pixel values (8 bit = 256 levels), some were fluctuating while others started when a window or door to the lab. room was opened. The calibration for such temporal variations can either be found as a drift in a measurement series, re-calibrating for e.g. every tenth measurement or by adding a small calibration mark (e.g. a small white plate) inside the camera view that each measurement is corrected for, such a small mark can also give information about the self-illumination.

3.7 Summary

There has been given an introduction to colours and the chromatic system of the human eye, where it has been stated that the use of RGB (or any trimulus colour system) in many machine vision colour measurement tasks only is an unnecessary simplification based on the human eye. There is no reason to apply a human limitation to a machine vision system without any specific reasons to do so. Various camera based devices for capturing colour or multi spectral images have been described with their respective limitations and advantages. The chromatic aberrations and the pixel based colour misalignments have been described with respect to the calibration of colour cameras. The use of the classic camera calibration (linear and non linear) and warping with colour images has been described, where the essence is to apply the monochrome models directly on each channel separately or only the central (best) channel. There has been given a substantial description of uneven intensities (illuminance and radiance) which is a common problem in image analysis and especially important in order to get correct colour/spectral measurements in an image.

Chapter 4

Calibration Targets

When calibrating cameras or other image capturing devices it is necessary to use a calibration target with clearly identifiable objects (marks). These objects can be of any geometrical shape as long as they either manually or automatically can be identified in both a world coordinate system and image coordinates using geometric entities, like points, lines, corners or circles. This chapter will give a description of some important elements to consider when using such calibration targets. The targets described in this chapter are primarily for use with calibration for image distortion (or geometrical transformations). Although many of the calibration targets and detection methods can be used with various devices the main concern in this thesis is cameras.

4.1 Different Devices

When designing or using calibration targets it is important to consider the type of capturing device and how to use it. In this section is a short overview of the devices that were used in this Ph.D. project. Often a calibration target can be used with almost any device but there are small differences that are useful to consider.

4.1.1 Common for Almost all Devices

If the capturing device and scene is mounted secure and free of vibrations it is possible to capture a sequence of images and generate the average image in order to reduce the noise/signal ratio. This will increase the accuracy of the detection of the individual objects (marks) on the target and through that the accuracy of the entire target, if making a least square fit on a high number of objects this could be unnecessary but it will never harm if the time allows it. In this project image averages of about 25 images have been used extensively. Especially with noise/signal ratio sensitive object detection algorithms the use of averaged images is a very useful method for improving a camera calibration.

Another approach which can be used in surroundings with periodic vibrations or other distortions is to capture a series of images with the same conditions and then repeatedly exclude the image which contains the worst measurements of the calibration objects until some set condition is reached, this approach has been used successfully in [Melen94] with monochrome cameras.

4.1.2 Monochrome Cameras

This is probably the most widely used type of camera in vision systems, due to the fact that they are fast, accurate, simple and cheap compared to the colour cameras that are used in machine vision system. The monochrome cameras are also the devices that can use almost any visible calibration target. The traditional calibration targets for monochrome cameras are (partly) planar and have a white background with black objects, different types of these targets are described later in this chapter.

The black and white targets are not always the optimal type when considering the actual dynamics of the system, which can be influenced by many parameters in a vision system like brightness, contrast, gain, iris, black level etc. When camera, lens and grabber are adjusted for a specific scene/lighting the white parts of the target may be saturated in digital/grabber or CCD cell sense and the

black parts may be darker than the digital black level on the grabber causing a digital saturation. In order to avoid this a calibration target could instead of black/white be created in two or more grey tones that provide good dynamics in the image without any pixels being saturated. If some parts of an image of a calibration target are saturated it could influence the detection of the objects and this error could easily be systematic in some direction. If the dynamics of the system are made so that it is close to saturation and the lighting conditions of the system can vary a bit, this could cause an error that could vary from calibration to calibration. The temperature of the camera and light source could also have some influence in this case so it is a good idea to make room for some variations in the grey levels in the images. These considerations are also useful with colour cameras.

Another consideration is the situation where the camera is calibrated in another light than it should be used for after the calibration. This could be the situation where the camera only can be calibrated in white light and later should be used for precise detection of laser dots/structures which is monochromatic light. In this case is would be very useful to use a calibration sheet with a background colour that exposed to white light is close to the wavelength of the laser in order to reduce the effect of the chromatic aberrations in the lens.

Beware of some camera controls, e.g. sharpness, which might apply a systematic error (displacement) when detecting calibration marks.

When using analog cameras like CCIR or RS-170 the length and quality of the cable between camera and frame-grabber (and the connectors) may also be an error source for automatic detection algorithms, figure 4.1 shows an example with a JAI M50 PAL monochrome camera and a Matrox Meteor grabber with *a)* a 13 metres low quality cable and *b)* a two metres high quality cable. As it can be seen from the horizontal profiles the long cable (13 metres) causes under- and overshooting in the electrical signal in the coax cable caused by induction. There is of course no problem in the vertical direction when using CCIR/RS-170. The under- and overshooting is only a problem with specific detection algorithms (e.g. template matching and parametric fitting) and if applying a local filter to the image before the detection which will cause a

systematic displacement of all calibration objects in the horizontal direction. The problems with under- and overshooting also apply to colour cameras.

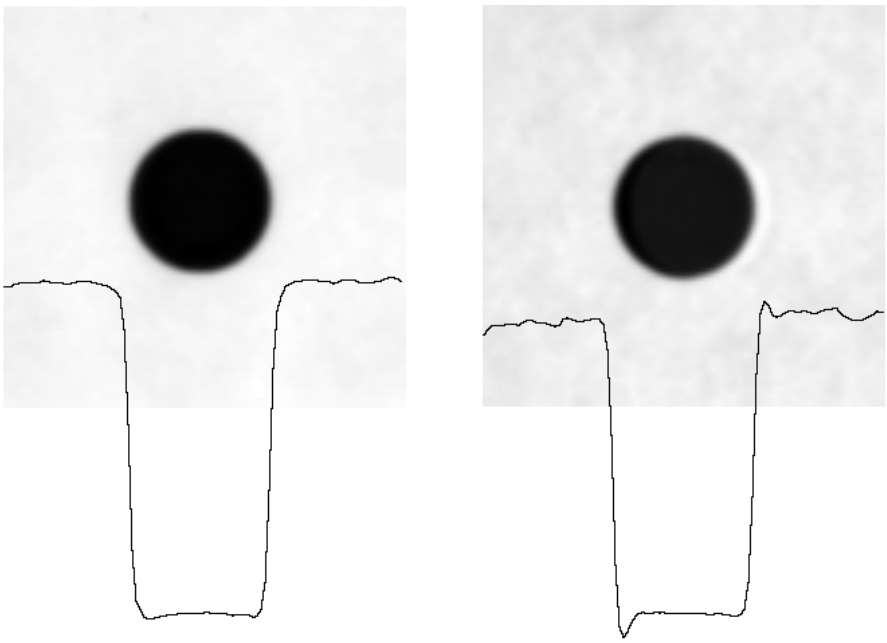


Figure 4.1 a) Grabbed with 2 m cable b) grabbed with 13 m cable.

Another problem with analog cameras occurs when the synchronization signal is overlapped with the image signal where the frame grabber must synchronize with the camera, if this is done inaccurately a horizontal image shift depending on the contents (intensity level) of the image signal can appear [Heikillä97b], this common phenomena is called line jitter. There has been no evidence of significant line jitter in the hardware used in this project, in appendix B we show a practical example of a measurement of line-jitter which shows a change in horizontal shift of about 0.007 pixels.

4.1.3 Colour Cameras

When using calibration targets with colour cameras, the experience from this project is that it is mostly the same targets as when using monochrome cameras

and many of the problems from monochrome cameras also apply to colour cameras. With multichannel or RGB cameras it is very useful to calibrate and resample the channels individually to the common coordinate system. When calibrating each channel separately it is of course very useful to use grayscale (often black and white) calibration targets to get a high degree of uniform intensity levels and comparison between the channels.

In some cases it is practical to use different colour combinations for specific objects on the calibration targets in order to get a distinction between objects in different channels. A simple example could be a planar target consisting of black disks on a white background with the center disk coloured green and two of the neighbours in red and blue. This use of coloured objects on targets could be used for aligning channels with a very low degree of spatial correspondence or for a simple automatic orientation detection of the target.

As described in chapter 2 and 3 the two most significant errors with images grabbed from an ordinary colour camera are the transversely chromatic aberration TCA and the barrel/pincushion distortion. In figure 4.2 is a computer generated drawing of the upper left corner of an RGB image of a calibration target with 3.8 mm disks, the center of the circles represents the found center of mass in the original image and the radius correspond to the found area of the disk. In this example the TCA error was about 2.2 pixels between the red and the blue channel in the upper left corner and 0.9 pixel in average (50 mm lens with a 10 mm extension at a distance of about 300 mm).

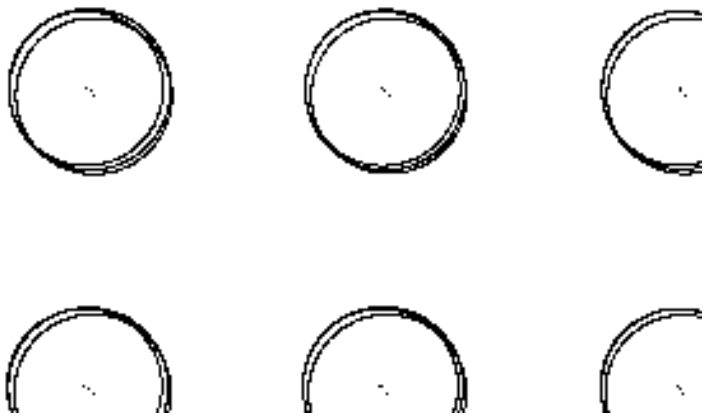


Figure 4.2 TCA error on RGB image of calibration target with black disks on a white background.

Another important thing about colour cameras is that the levels of red, green and blue in the image are not always aligned, this can be caused by many different sources in the long signal way from the lens to the digital data in memory, for further details see chapter 3. When making targets that also calibrate for these differences between red, green and blue levels at least two levels are required for a high quality calibration. A target that has been used for practical purposes in this project is the use of a grid of dark disks on a bright background, which at the same time can be used for calibration for distortion and TCA. Because of the good spatial distribution of the disks this can also be used for the calibration for colour shading in a 3xCCD camera and even for an uneven distribution of the light. In this way only one calibration target is needed for four calibrations at the same time and the algorithm that detects the center of the dark disks simply used the space between two dark disks as a bright area. This has been implemented in the Surfix texturemeter device with very good results, see chapter 5.

The problem with under- and overshooting described in the previous subsection is also a problem with analog colour cameras, in this project it has been observed with a 3xCCD JVC camera connected with two 2 metre 4xCoax-cables joined with four female BNC junctions to a Matrox Meteor/RGB grabber. Long cables would probably cause the same phenomena.

4.1.4 Scanners and Line-scan Cameras

Line-scan cameras are single array CCD’s which instead of an image give continuous horizontal lines, an image can then be generated by moving the camera or the object in the vertical direction. They use an ordinary camera lens and the lens distortion is therefore a combination of radial and tangential distortions, the radial distortion introduces an error where the horizontal line is non-linear measured from the principal point and the tangential gives an error where the left and right side of the lines are curved a little bit up and down. When moving the camera or the object the speed may vary cyclically (for continuos grabs) or in the startup phase introducing an error in the vertical

direction. It can be very difficult to place the camera in the precise position/movement where the scan-line is perpendicular to the moving direction, introducing an error where the images actually are parallelograms but are grabbed as rectangles. When calibrating line-scan cameras it is therefore in most cases necessary to obtain information of coordinate displacement in both directions. For this purpose a calibration target consisting of a planar grid of disks (with colours optimized for the camera system parameters, see above) is very optimal. Another useful calibration target could be black lines on a white background grabbed in both directions, this would give a visual notion of the displacements.

Like the line-scan cameras a scanner is often some sort of array that is moved perpendicular to the line, but with a scanner the object being scanned is usually planar (or nearly planar), the calibration targets for scanners is therefore usually planar (sheet). With most scanners there is no lens in traditional camera sense and therefore no radial/tangential distortion. The line-array is typical very straight and since it is factory mounted it is also very perpendicular to the moving direction.

Today's (1998) traditional scanners use a CCD chip which captures image data through a lens-and-mirror optical system. In such a system the optics can become misaligned giving manufacture specific displacement errors. These displacements can be in both directions and a calibration sheet that can obtain coordinate information in both directions would be the best (like the dark disks on bright background). The work with colour CCD scanners (most common today) made in this work has made it clear that the calibration sheet must be able to give information about misalignment of red, green and blue channels since this was the most significant error (on cheap scanners up to 2 pixels in 600 dpi).

Calibrating a CIS (Contact Image Sensor) scanner can be made a little simpler than the CCD based since the sensor-array is straight and free of optical misalignments from lenses and mirrors. A calibration should only calibrate for misalignments from the moving array including if it is totally perpendicular to the moving direction. Calibration sheets for a CIS scanner could in a simple

calibration be vertical lines to check if the array is perpendicular to the moving direction and a more complex calibration could be dark disks on a bright background to calibrate for a possible small misalignment between the red, green and blue channels.

Another type of scanner is the triangulation 3D laser-scanner, which consists of a laser-line and a camera. In order to calibrate a 3D capture device it is best to have some 3D objects to calibrate against, in the work done in this Ph.D. project with laser-scanners a flat board with cylinders mounted in a regular grid has been used. Other good objects could be real 3D objects like a precise cube or spherical cap, these would also be very good for a final quality evaluation of a laser-scanning system. Another problem with laser-scanners are the errors due to reflectance discontinuity (colour or gloss change). There is no easy way [Sternberg97] to calibrate for these errors but the design of a smooth target (like a spherical cap) with different colours and degrees of gloss can give some results of the evaluation of the errors.

4.2 Miscellaneous Targets

The most elaborately, common and practical useful calibration targets are planar or partly planar calibration targets. For an easy detection the planar targets have a background colour with differently coloured objects making a step edge. Black and white are two very useful colours for this purpose since they can be used for all wavelengths. In figure 4.2 is an illustration of one of the reasons why step edge objects are the best, beside the fact that they are much easier to construct in practice. When applying unevenly distributed light to a non step edge object (like the cone in the figure) almost all detection algorithms will make a systematic displacement of the found object centres.

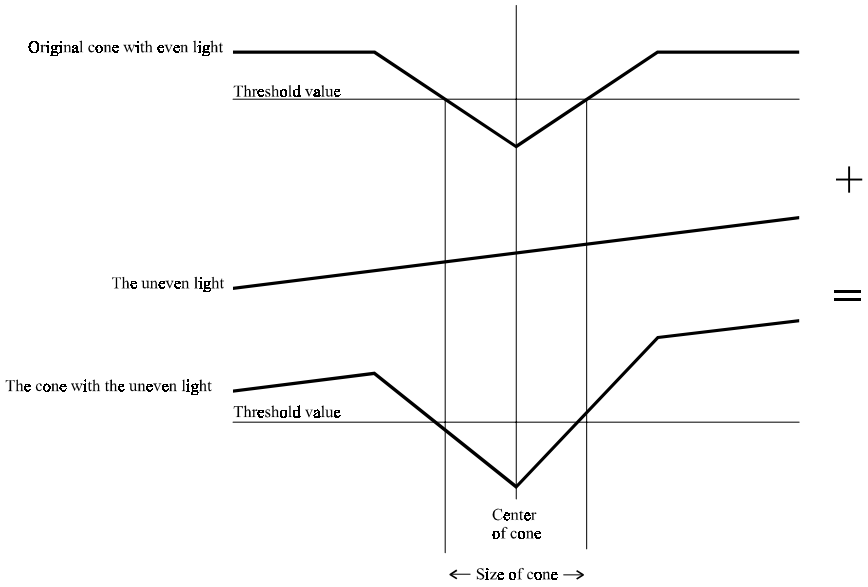


Figure 4.2 Profile of a dark intensity-cone on a bright background added with some uneven light which can give a displacement of the cone detection.

For an object on the calibration target it is often used to identify it as a single point or a group of points. For circles this could be the center, triangles the three corners and for squares it can be the cross between the diagonals or all four corners. In figure 4.3 is two examples, a) is a planar calibration target with squares, they give either four calibration points for each square using the corners (see [Weng92] or [Tsai86] for a practical use of such targets) or one point using the cross of the diagonals, b) is a cube (piecewise planar) with 16 disks on each side this can be used for obtaining 3D points.

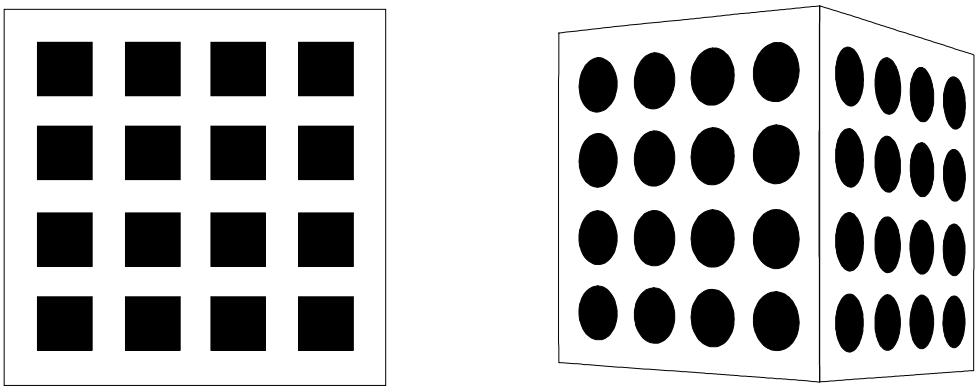


Figure 4.3 a) Planar calibration target with squares b) Piecewise planar 3D calibration (cube) target with circles.

It is also possible to design a real 3D calibration target using multiple transparent planar targets in parallel layers e.g. with equal distance between the layers, see figure 4.4. It is of course possible that some of the objects will be overlapped from the used viewpoint dependent on the object size and the number of objects. The effect of two objects overlaying a little bit can cause a misinterpretation of the two objects being observed as one with a slightly different shape, causing a possible displacement of the object in both layer number and image coordinate. This can be corrected using some sort of colour coding of the layers and grabbing of images with colour filters and one image without filters for the final detection.

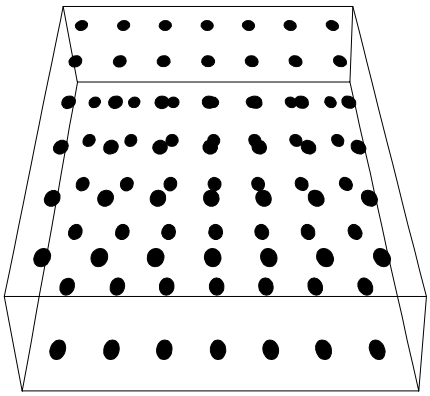


Figure 4.4 Two layer 3D calibration grid.

Another commonly used calibration target is a checkerboard like in figure 4.5, this can be identified/detected using horizontal and vertical lines which shift between a black-white and a white-black edge. Then the crosses between the horizontal and vertical lines can be used as calibration points. Because of the shift between edge direction, displacements on the edge will neutralize each other which makes this a quite robust method. One disadvantage with this method is that the lines due to distortion are not necessarily straight and they must therefore be some sort of snakes/deformable lines perhaps with the constraint that they can be bent according to a normal lens distortion. An advantage with this sort of calibration target and other corner/line/edge detection methods is that they do not have any perspective projection displacement (PPD) see more details later in this chapter.

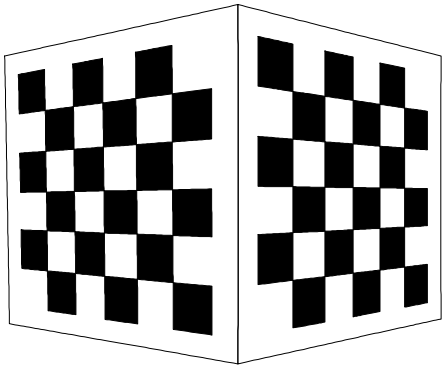


Figure 4.5 Checkerboard 3D calibration target.

Another line based calibration target is a grid made of horizontal and vertical lines. The use of lines for calibration is called plumb line calibration and it can be very well suited for determining lens distortion if the proper model is applied [Melen94]. This sort of target is not the most useful in practice since the subpixel detection of a cross or individual lines is not very accurate compared to e.g. a disk of the same size and a detection of the lines has the same limitations regarding distortion as with the checkerboard.

In many applications, like stereo-vision, it is useful to get the same unique positions for the individual objects in a regular grid. One way to do this could be to make a coding of a reference object (like the center object). In figure 4.6 is a coding for dark disks on a bright background invented by me where the center disk is larger and three (or all four) neighbours are smaller than the normal disks. An algorithm for finding the reference disk must then look for a disk that is significantly larger than three (or all) of its neighbours, this is a very robust method and works for all angles and distortions that have been used in this project. A weight function like the area of the actual disk divided by the area of the four neighbours would be a good way to do this. As a case study in this Ph.D.-project the use of such calibration targets has been implemented in a 3D PIV system (a stereo system) at Dantec Measurement Technology A/S with very good results (see chapter 5). In the Dantec MT system the center disk is 30% larger and the neighbours are 30 % smaller than the normal disks.

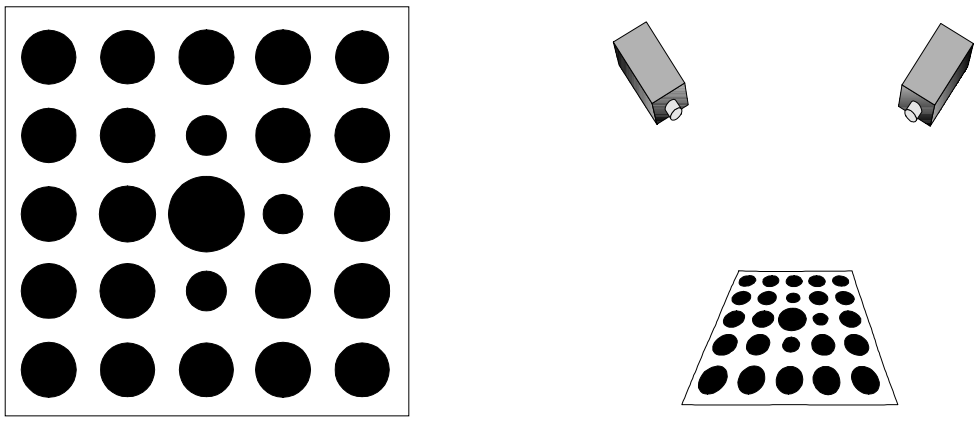


Figure 4.6 Reference coded calibration target. Useful for matching images with stereo-vision or a calibration using only one planar calibration target.

In some situations the calibration objects must be suitable for both manual detection in e.g. a stereo-comparator and for automatic detection. When using disks as objects a small hole with a cross in the middle of each disk is one example of such objects, see figure 4.7.

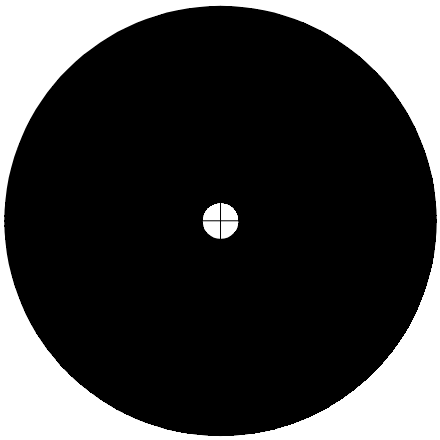


Figure 4.7 Calibration object suitable for both manual (the small cross) and automatic detection.

Materials

When designing calibration targets the choice of material is of course very important for the obtainable quality.

Paper is cheap and easily obtainable but since it is a natural product made of wood it is very sensitive to temperature and humidity and paper should therefore for optimal performance be glued with a water resistant glue to a stabile surface. In the calibration work done in collaboration with Dantec Measurement A/S a calibration target made of a paper (size about 100x100 mm) that only was fixed with clips to a metal plate was giving some initial unexplainable errors in the camera calibrations, it turned out that the reason was that the paper was a little bit bumpy, less than ½ mm. After the paper was glued to the metal plate the error disappeared.

Another commonly used material is mat plastic sheets that can be used directly in plotters or printers (glossy plastic can have some unwanted reflection). They are more resistant for humidity and temperature than paper and they are also easier to keep clean in an industrial environment.

When fixing a calibration target to a metal plate (or any plate), the temperature dependent expansion should be considered if not all measurements can be done at the same temperature. This can be a considerable error if e.g. a light source is pointing directly at the calibration target, direct sunlight or a camera on a robot in a heavy industrial environment.

Other common methods are to stitch planar objects directly onto a hard material or to draw directly on the hard material.

When using plotters or printers to make calibration targets the quality is normally quite high. But there are some error sources to be considered. Since most printers uses rolls for paper feeding an error source in the paper feed direction (normally vertical direction) could be the thickness of the paper (the pressure on the rolls is not the same which could influence the speed). The

easiest way to check the horizontal / vertical resolution is to print a large object and then measure the height and width. Plotters are usually very precise, but again it is a good idea to measure the vertical and horizontal resolution.

In the table below is some measurements of a 150 x 150 mm square printed on four different printers, measured with a 0.05 mm precision device. The measurement was done on several prints from the same printer without any measurable difference except when changing paper thickness.

Printer	Used Paper	Measured vertical	Measured horizontal	Vertical deviation	Horizontal deviation	Relative difference*
HP4000TN	A4 200g/m^2	150.1 mm	149.7 mm	0.07 %	0.20 %	0.27 %
HP4000TN	A4 80g/m^2	150.1 mm	149.8 mm	0.07 %	0.13 %	0.20 %
HP4000N	A4 80g/m^2	150.5 mm	150.2 mm	0.33 %	0.13 %	0.20 %
HP1120C	A4 80g/m^2	150.3 mm	150.5 mm	0.20 %	0.33 %	0.13 %
HP ColorJet5M	A4 80g/m^2	150.1 mm	150.1 mm	0.07 %	0.07 %	0.00 %

Table 4.1 Measurements of prints of 150 x 150 mm square on four different printers.
* Relative difference between horizontal and vertical direction.

The problem of the small deviations using ordinary printers is normally not so much the size of the objects on the calibration targets but more the distance between the objects. Using a target of 300 x 300 mm gives a maximum error of about 0.6 mm or when using a PAL camera equivalent to about 1 pixel which for many applications is quite unacceptable.

Another possibility is to make a complete measurement of a calibration target, although this requires that the objects (marks) are suitable for manual detection in e.g. an optical stereo-comparator (device for photogrammetry). In [Sternberg97] is an example of such a measurement of laser printed disks on A4 film, the results from this work are that there is a compression in the paper feed direction (< 0.5%) and that the compression has a small variation from top to the bottom of the page caused by a little difference in paper feed speed from start to end. Another important conclusion from this work was that the size of the errors was too small to influence significantly on the shape of object but it was large enough for influencing on the distance between objects.

In some automatic detection cases like when using a template matching or a parametric fitting the accuracy of the object detection is dependent on the number of edge pixels. In this case objects with many edge pixels can be designed, in figure 4.8 is an example of such objects.

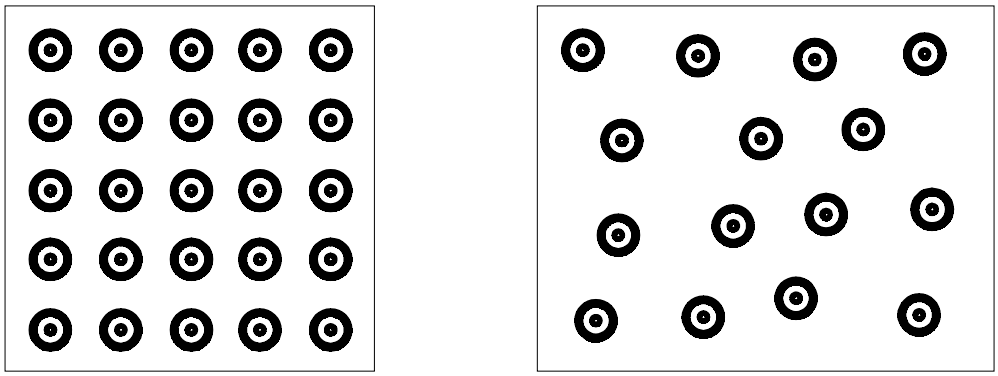


Figure 4.8 Calibration targets with objects made of disk with a white circle to give more edge pixels, a) objects in an ordered grid b) objects placed in a non ordered way.

As illustrated in figure 4.8 it is not always necessary/possible to use calibration objects in a regular grid. When the calibration objects cannot be automatically integrated in a calibration target and instead must be placed manually in various places in the scene (sometimes the only places where it is possible) a very precise measurement of the positions of the objects must be performed. This is done in a 3D framework at the Department of Planning DTU where about 100 planar objects are wired up in a non-regular 3D grid, the position of these 100 objects are then measured with very high accuracy using a theodolite. Another place is the Odense Steel Shipyard in Denmark where calibration objects similar to those in the figure above are placed manually on large steel objects [Knudsen98]. A disadvantage of objects placed in a non ordered way is that it makes it difficult to make a completely automatic detection algorithm, since it must make a unique identification of each of the objects. With three known objects it is though possible to get an initial good guess of the position of the remaining objects.

When designing calibration targets an important task is to determine the best

size and number for the objects. For most calibration methods a large number of measured points are desired for an accurate calibration [Heikkilä97b]. However many calibration objects also require a lot of space giving a contradiction with the request for large objects for high (subpixel) accuracy measurements of the individual objects. The practical work in this project indicates that the radius of e.g. disks should be larger than 4 pixels and the distance between object borders should be larger than about 8 pixels in order to get acceptable results when using semi-automatic and automatic detection algorithms. In [Sternberg97] is a simulation of centroid estimations of disks in pixels, the results from this work states that the error is about 0.1 pixels when the radius is about 4 pixels and about 0.06 pixels at a 10 pixels radius. These errors are very small and if they are non systematic and a calibration is made over many points they should not cause any problems. The practical work in this project also confirms these statements, the only found limitation in the size of objects (when using disks) was concerning automatic detection algorithms being able to find the objects in the first place.

Another sort of calibration targets are real 3D targets for 3D scanners. They do not necessarily need to be textured or coloured in different colours since they normally are used for the depth/3D information. They must not be glossy when using laser based 3D scanners because of some common reflectivity problems. When calibrating/testing pairs of cameras for stereo vision the calibration target must of course contain some sort of texture/colours suitable for stereo matching. When designing 3D targets the possibilities for (3D) shadows should be considered carefully, a good idea could be a planar plate with some sort of known objects mounted on top of it, in this way the objects will not shade for each other.

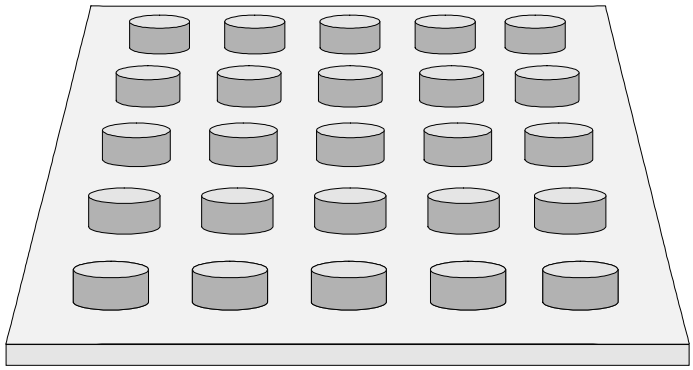


Figure 4.9 Example of 3D calibration target for laser scanners. Actually used at Danisco Sugar for testing.

In figure 4.9 is an illustration of a target for a laser 3D measurement device designed and used in the collaboration with Danisco Sugar, the target was 1000mm x 1000mm, 25 cylinders was fixed to the plate with height 43 mm and radius 50 mm, the distance between cylinders was 200 mm. The cylinders are easy to create accurately in a turning lathe and the fixation can be done with glue or screws in accurately drilled holes from the bottom. When using a target like this one with a height measurement device from above many properties can be extracted; variations in height measurements from the measured height of the cylinders, spatial displacements from the center of the top of the cylinders, variations in ground height from the plate and variation in measurements in various heights when grabbing multiple height-images at different height.

4.3 Detection

The detection of objects on calibration targets can be divided into detecting images of targets for camera calibrations and detecting objects directly on the targets for measuring the targets. The positions of objects on a target must of course be measured correctly and with high precision (or at least with non-systematic deviations). The detection can in both cases either be manual, semi-automatic or automatic. In some situations the objects must initially be measured manually and afterwards detected using semi- or fully automatic methods making it important to design the objects in a way that is suitable for both methods.

4.3.1 Manual

In many cases the only possible or suitable method for detecting and measuring the positions of the objects on a calibration target is to use some sort of manual device directly on the target. These devices can vary from rulers to sophisticated equipment for photogrammetry. Also when detecting the objects

on an image of the target the manual approach can be necessary on negatives, positives, paper or even on screen with digital images.

A simple case can be a regular grid of calibration objects (e.g. made on a printer) where a normal slide gauge which gives an accuracy of ±0.05mm can be used for measuring the distance between the outer objects directly on the targets. When making the same measurements on images of targets (negatives, positives, paper) this approach is very slow and very time consuming because of the distortion.

With digital images of targets many software packages make it possible to zoom in on an object and subsequently place a cross (mouse pointer) on the place representing the calibration point for that object. Using a large zoom and allowing the user to move the cross/pointer manually in subpixel steps actually makes it possible for the well trained user to make a coarse subpixel position detection of the calibration objects, see figure 4.10.

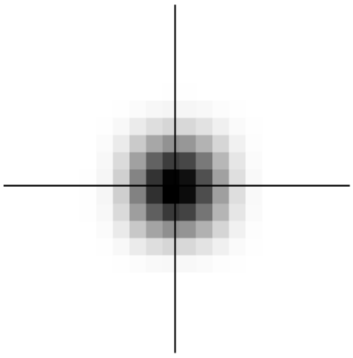


Figure 4.10 Manual subpixel detection of calibration object.

When using a photogrammetry device like a stereocomparator or a monocomparator it is possible to measure calibration points manually directly on negatives or positives with an uncertainty of ±1µm [Krauss93], with the condition that the design of calibration objects makes it visually possible. These devices can also be used with most planar calibration targets for measuring the exact positions of objects.

At the Department of Planning - Surveying is a 3D framework with about 100 planar objects wired up making a full 3D calibration target, see figure 4.11, each of the objects is measured in world coordinates with a precision of less than 1 mm. A framework likes this requires periodic (annual) inspection and measurement of the objects in order to serve as a high precision calibration tool. This framework is primary designed for manual detection of calibration marks using a stereocomparator or monocomparator and the layout of the calibration marks are made for an easy detection at many resolutions. This makes it easy to find and detect objects at a low resolution and the layout of the center of the objects make it possible to do a very precise detection when using a high resolution.

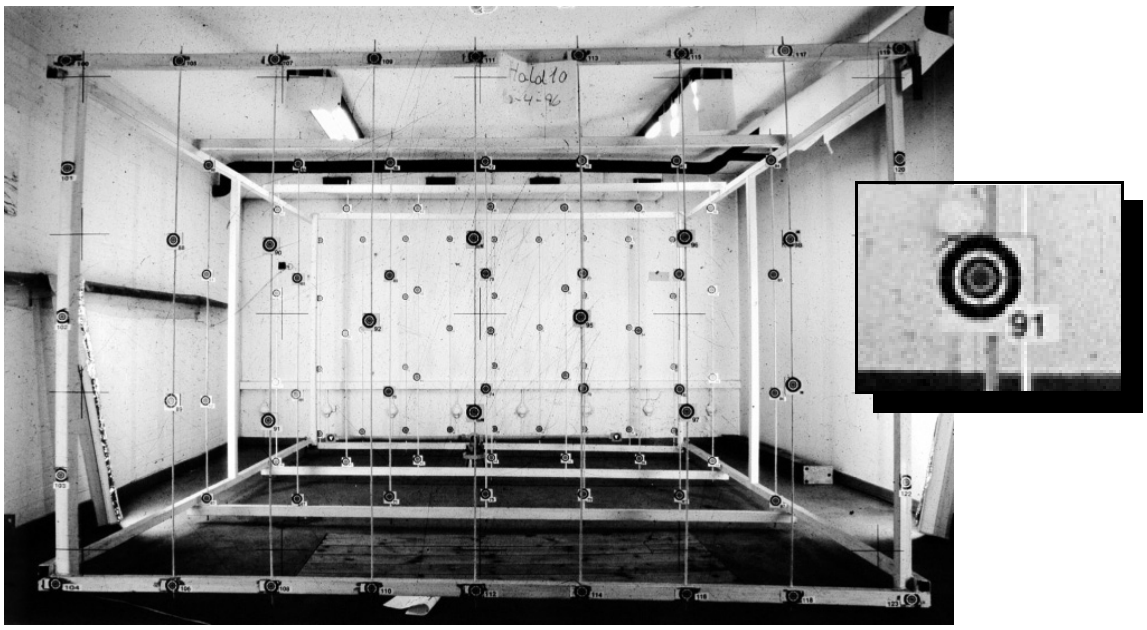


Figure 4.11 Framework at Department of Planning, Surveying (35mm film)

In figure 4.12 is an illustration of a typical calibration object used for manual detection, is consists of concentric black and white circles overlapped with lines getter thinner and thinner towards the center of the object, making the center very detailed. A problem with this sort of calibration objects is that it can be hard to detect the center manually if the image is blurred, this is normally not a problem in photogrammetry because it mainly is used for landscape or building

measurement where focus is set to infinity. But in machine vision where the objects can be close to the camera and at different distances it is not always given that all objects are in focus.

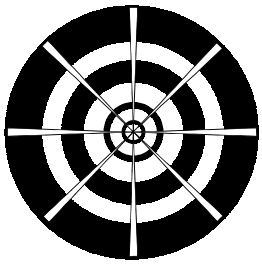


Figure 4.12 Typical calibration mark for manual detection.

For objects suitable for both manual and automatic detection see figure 4.7 in the previous section in this chapter.

4.3.2 Automatic and Semi-Automatic

The primary task for an analysis of a calibration target is to determine each of the individual objects in the image plane. In this Ph.D. thesis the only thing that differs in the definition of a semi-automatic method from a fully automatic is the manual specification of the approximate coordinates of each object. The simplest semi-automatic method is to display the image on screen and by using a mouse pointer to click out the approximate positions of each calibration object and either enter the world coordinates or let the software make an estimation based on the prior knowledge about the positions in world coordinates and the permitted perspective transformations (with distortions).

For a fully automatic method the software must look for the objects using some sort of prior knowledge about the shape/properties of the objects. The needed extent of this knowledge is of course dependent on the method.

Template matching using e.g. the correlation coefficient can be useful when looking for objects in a complex world, if successful, there will be a highlight at the positions giving a match between the template and the object. Since the template matching is not rotation invariant the design of objects for this purpose must be rotation symmetrical. Another problem is the lack of scale invariance and that the approximate size of the objects in the image plane must be known. In images of 3D calibration targets or images grabbed at a steep angle the size of objects can vary considerably making this method useless. At Odense Steel Shipyard template matching (using the correlation coefficient) has been used with some success for an automatic detection of objects consisting of concentric black and white circles [Knudsen98].

When looking for objects in a simple world, like black disks on a white background, the use of a threshold on a background intensity corrected image for finding the objects is possible. This method has been used much in this Ph.D. project and also in different modifications through the industrial cooperation with Dantec Measurement A/S, Department of Fluid Mechanics and the design/development of the VMX2000 texturemeter, see chapter 5.

When using threshold and labelling for detection of objects in a real world application one of the first tasks is to correct for a possible uneven intensity distribution in the image plane caused by e.g. uneven lighting, shadows, colour shading or vignetting. In the following is an example from a single row of black disks on a white background (white paper) illuminated from the left side, the example starts with the original image in figure 4.13.

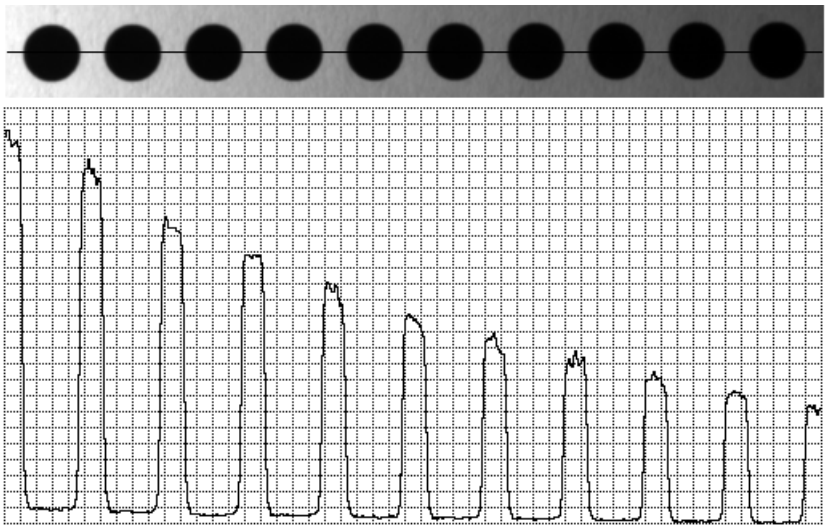


Figure 4.13 Original image of black disks on white paper illuminated from the left side. Below is the intensity profile of the marked line

In order to find an estimate of the intensity profile in the image the used approach was to use a local 3x3 minimum filter (morphology erode) first in order to remove possible peaks and then run a large (33x33) maximum filter (morphology dilate) in order to remove the black disks, the size of this filter must be at least the largest diameter of the disks in the image. In figure 4.14 is the result of this operation.

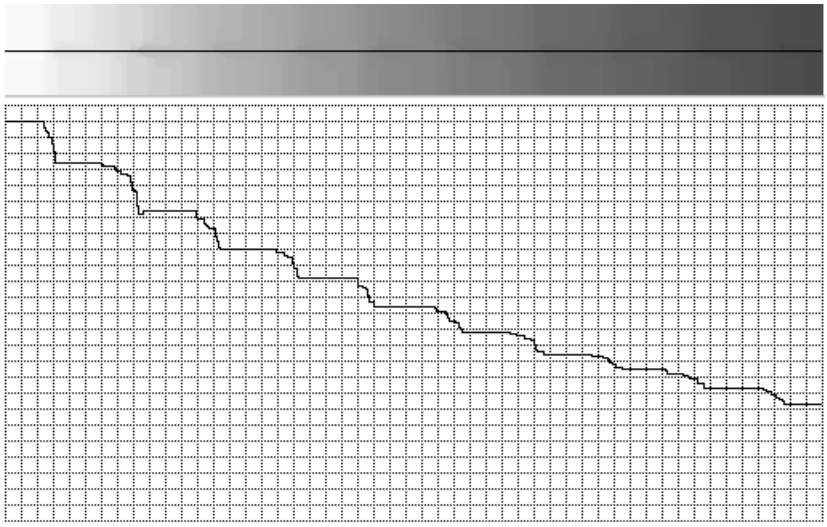


Figure 4.14 The result after first a 3x3 erode and then a 33x33 dilate.

After the morphologic operations the resulting intensity estimate has a significant staircase appearance. In order to reduce this a large local mean filter is applied see figure 4.15. In the initial texturemeter (named *Surfix* see chapter 5) the background correction (during the camera calibration routines) was made as a two dimensional second order polynomial fit to the observed intensity which of course reduced all effects like the staircases in figure 4.14.

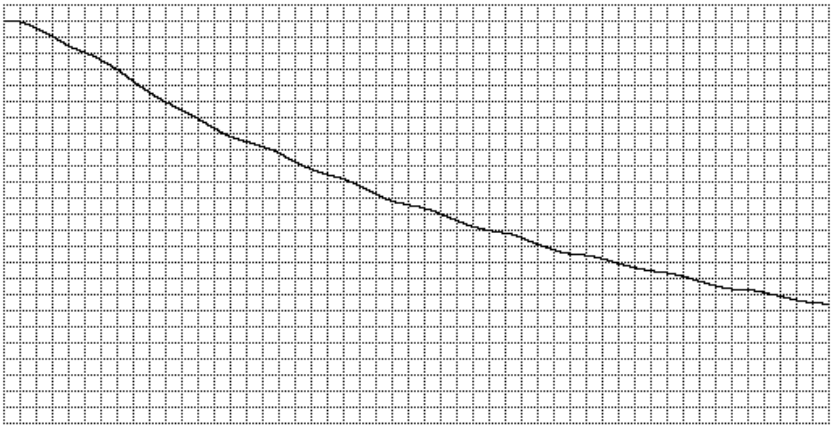


Figure 4.15 The intensity profile after the morphologic operations and a large 33x33 local mean filter.

By using the estimated intensity image from figure 4.15 it is possible to construct a new image without intensity trends. One way of doing this is to multiply the original image pixel by pixel by the reciprocal intensity image, see figure 4.16, this is not entirely correct since it does not precisely correct the darker areas because of the lack of a correct zero, but it is sufficient for most applications. In the work made in this project with the first texturemeter *Surfix* the values of both the black disks as well as the white areas in between was used to make a linear stretch between black and white [Folm-Hansen97] making the background correction correct with all grey tones. Another solution can be to use a large morphological erode to let the black disk grow to the complete image and then estimate the dark intensity trends similar to the previous background correction. Then with a simple linearly stretch between the dark and light areas to make a more correct background correction.

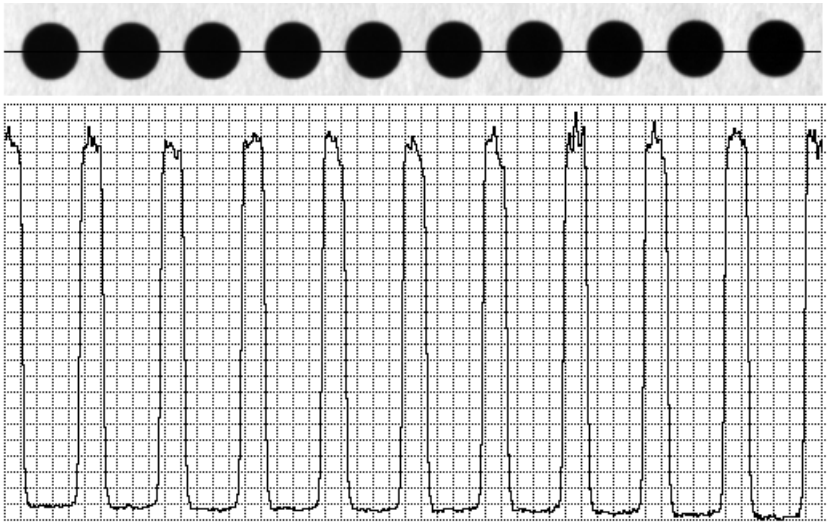


Figure 4.16 The result of a reciprocal based background correction. Notice the small trend in the dark areas (black disks) originating from the lack of correct black level.

When the image of the calibration target has been background corrected the search for a correct threshold is a simple task, in figure 4.17 is the histograms before and after the background correction. As it can be seen, the search for a valley between the two tops is very easy in the corrected image but quite complex in the original.

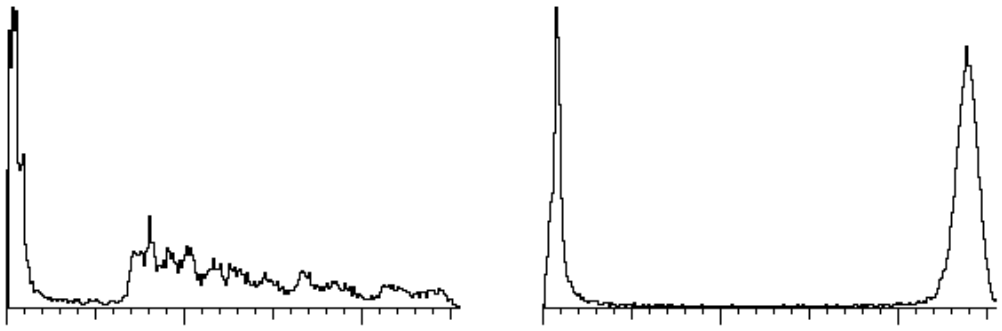


Figure 4.17 Original histogram (left) and background corrected image histogram.

A way of determining the threshold value in a background corrected image could be to assume dark values equal the mean of e.g. the lower 10% grey tones and light values to the higher 10% and then use the mean value of these two mean values as the threshold value.

In figure 4.18 is the thresholded image using the average of the minimum and maximum value as the threshold value of the same rows of black disks as in the previous figures.

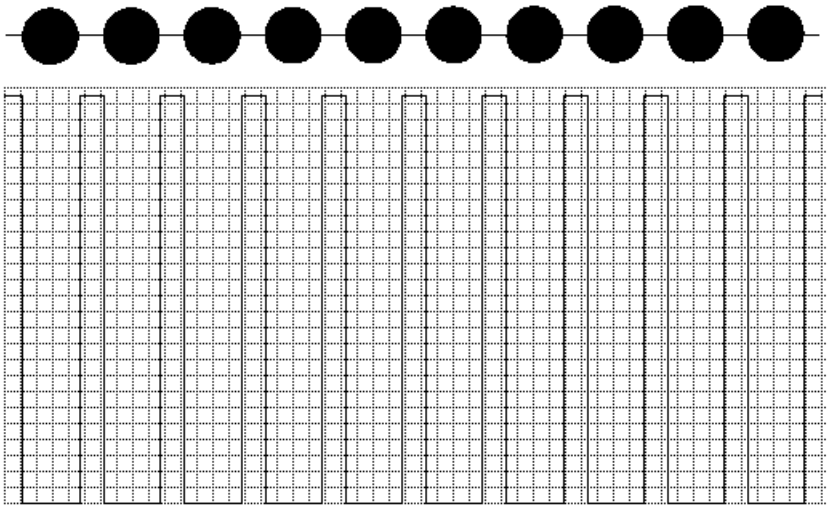


Figure 4.18 The background corrected and thresholded image with its binary profile

Labelling

After the thresholding each of the objects in the image must be uniquely segmented in order to get information about area, position etc. for each of them. For this purpose the ordinary image processing method called labelling or colouring can be used, this creates an integer image where the background e.g. is zero and each integer value is assigned to only one object, the two main procedures for this are recursive and sequential labelling. The recursive is fast with small objects but with large objects there is a significant risk of making an overrun in the computer registers. The sequential is very robust with all sizes of objects but requires at least two passes through the image and is therefore not as fast as the recursive, the method used in this project is the two pass sequential presented in [Sonka93].

Moments

When the objects in an image of a calibration target are easy to separate and can be identified as a single point given by the center of mass a useful method for detecting the objects is to use the moments for each label (or region containing an object), see [Sternberg97], [Melen94] and [Folm-Hansen97].

From [Gonzalez87] we have given a two-dimensional continuous function f(x,y) the definition of moments of order (p+q) :

$$m_{pq} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^p y^q f(x,y)\,dx\,dy \tag{4.1}$$

for p, q = 0,1,2, ...

For a digital M×N image f(i,j) (grey tone as well as binary) this becomes:

$$m_{pq} = \sum_{i=0}^{M-1}\sum_{j=0}^{N-1} i^p j^q f(i,j) \tag{4.2}$$

for p, q = 0,1,2, ...

When applied to an integer label image the function f(i,j) must be zero for background and one separately for each label. In a binary or label image m_{00} denotes the area. The estimate of the center of mass is then given by *c*:

$$c(\bar{x},\bar{y}) = \left(\frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}}\right) \tag{4.3}$$

In [Heikkilä96b] is an example of an extended center of mass with a weight on each pixel in the binary (or label) image according to the original pixel value. This approach suffers from small systematically errors when applied to images with uneven illumination (see detailed description later in this chapter) and is therefore not recommended or used within this project.

Mapping

After the labelling and calculation of moments for each label the mapping between world coordinates and the image coordinates must be established. This is also the case where other methods like template matching or parametric fitting has been used. The methods for this mapping can roughly be divided into two categories; objects arranged in some sort of regular grid (often planar targets) and objects placed in a non-regular way (often non-planar targets).

For objects placed in a non-regular way the known world coordinates can be used to establish all physical possible and plausible perspective projections of the calibration target in order to try to obtain the correct mapping. In many practical cases e.g. the distance between camera and calibration target and camera angle can be set to a limited interval which in many cases makes the search for a solution more possible, this could e.g. be that the distance between camera and target always is between 0.2 and 2.0 metres. If the possible mapping after these limitations still has more than one solution the limitations must either be further constrained or a rearrangement of the calibration objects must be considered.

If the calibration objects are arranged in a regular grid either completely planar or partly planar like a cube the methods for establishing the mapping between world and image coordinates can be quite unproblematic especially with a coded center and direction as described in the previous section. In the following is a description of the methods used in this project together with some remarks on other solutions, the used calibration targets where all black disks on white background and in some cases with a coding for the reference object (which was 30% larger than the normal disks with neighbours 30 % smaller).

In the methods used in this project (only planar targets) the first task is to find the center or reference object. For a non coded regular grid this can simply be the center object in image coordinates and for the reference object coded grids this was done by calculating the mean area for objects and then looking for an object larger than the mean with all four neighbours smaller than the mean area. When the reference/center object is found a least square estimate of the two

perpendicular lines going through the object and the four adjacent objects is calculated, this is used for estimation of the orientation of the target.

From this point two methods have been used. The first and simplest one was to expand the grid directly from the found two perpendicular lines with the object space found from between the center/reference object and the four adjacent objects and then allowing half an objects space deviation from objects to the grid, see figure 4.19. This method only works with images with approximately parallel image plane and target plane and the distance between objects must not exceed the image distortion. This is implemented in the VMX2000 texturemeter prototype and has only failed in a few cases with lenses combined with expander producing a large image distortion.

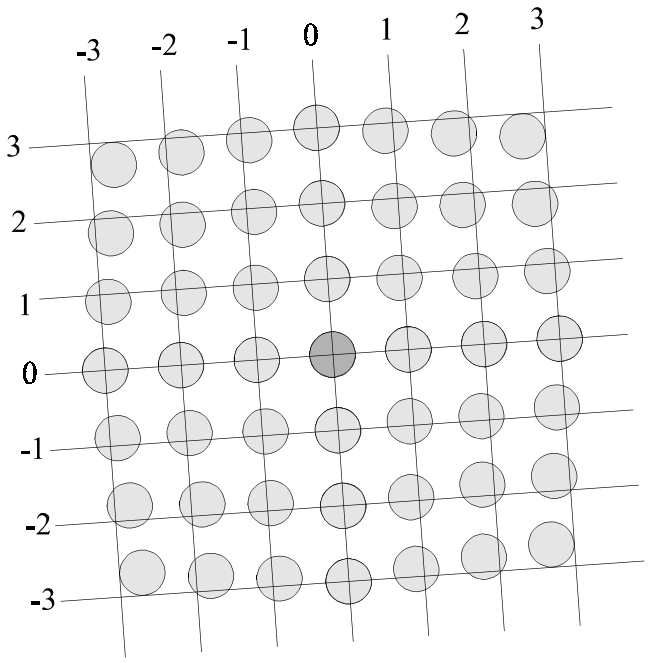


Figure 4.19 An illustration of the detection of disks in a rotated and distorted image using a grid starting at the center object.

The second method for estimating the regular grid is a recursive search method that is suitable for both images with parallel image plane and object plane as well as steep angle images. The recursive search starts from the center/reference

object (with coordinate (0,0)) and moves to all four object neighbours while transferring the numbering and coordinates according to the estimated angle and distance between objects. Each new search allows a change in angle and distance between objects making it possible to follow objects in a very distorted and perspective image. A search will end if it is close to the border, the object has already been visited or all four adjacent objects do not exist. In figure 4.20 is an illustration of the recursive search in two levels (second level is dotted). The result of this recursive search method is a very robust and fast method that has shown promising results in the practical work made in this project. This method has also been implemented at Dantec Measurement A/S from a software prototype developed in this project.

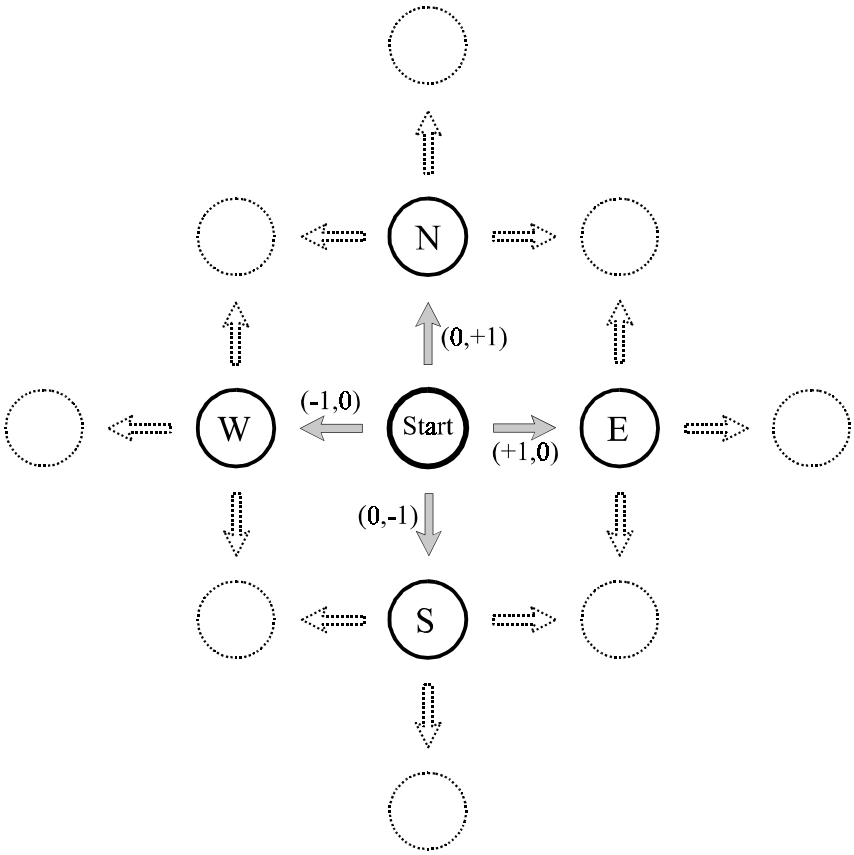


Figure 4.20 A new search at *start* in a recursive search algorithm. The coordinate list is updated for each new search. The dotted disks/arrows present the next level searches. A search ends at an already visited object or a border.

As a summarization the method for automatic detection of planar calibration with black disks on white background used in this project can be described with the following 11 points:

- Correct for uneven illumination
- Threshold image from histogram
- Label image
- Calculate moments (area and center of mass)
- Remove all objects too close to border and all very small objects.
- Remove all objects *50 % smaller or larger than the average object.
- Rearrange labels for removed objects/labels
- Find center/reference object
- Find directions of target through the four neighbours.
- Start recursive search in four directions from the center/reference object
- Save result

* The 50% is based on the experience from Dantec Measurement with images in a 45^{O} angle, this value should of course be adjusted to the actual camera setup, with parallel image plane and target plane it can be a very low value like 5%.

When using 3D calibration targets like multiple parallel and transparent layers the camera calibration is best made where the image plane and grid plane are parallel so that the size of objects can be used to group the found objects into layers. Objects in the grid with wrong size will then be overlays between objects and must be ignored.

In the following are descriptions of some other useful methods for semi-automatic estimation of objects.

Edge based methods

A commonly used alternative to the moment methods for detecting calibration objects is to detect the edge(s) of the object and then fit a suitable parametric function to the found edges. This could e.g. be fitting an ellipse to an edge image of a circle (disk), this method gives a high precision but is

computationally heavy alternative to the moment methods. Another advantage compared to e.g. template matching is that it is independent of scale and variations in perspective transformation.

The first problem is to detect the edges (or parts of the edges) with a correct and unbiased method. In figure 4.21 is an illustration of an edge profile from a blurred step-edge object, the figure illustrates the two main edge detection models, the first order and the second order derivative. Using the first order derivative the local maximum and minimum values can be used for edge detection and with the second order derivative the zero-crossing is the edge position. The zero-crossing method offers a reliable alternative in cases where the edges are blurry or when a high noise content is present because of the possibility of smoothing the image to reduce the effects of noise [Gonzalez92].

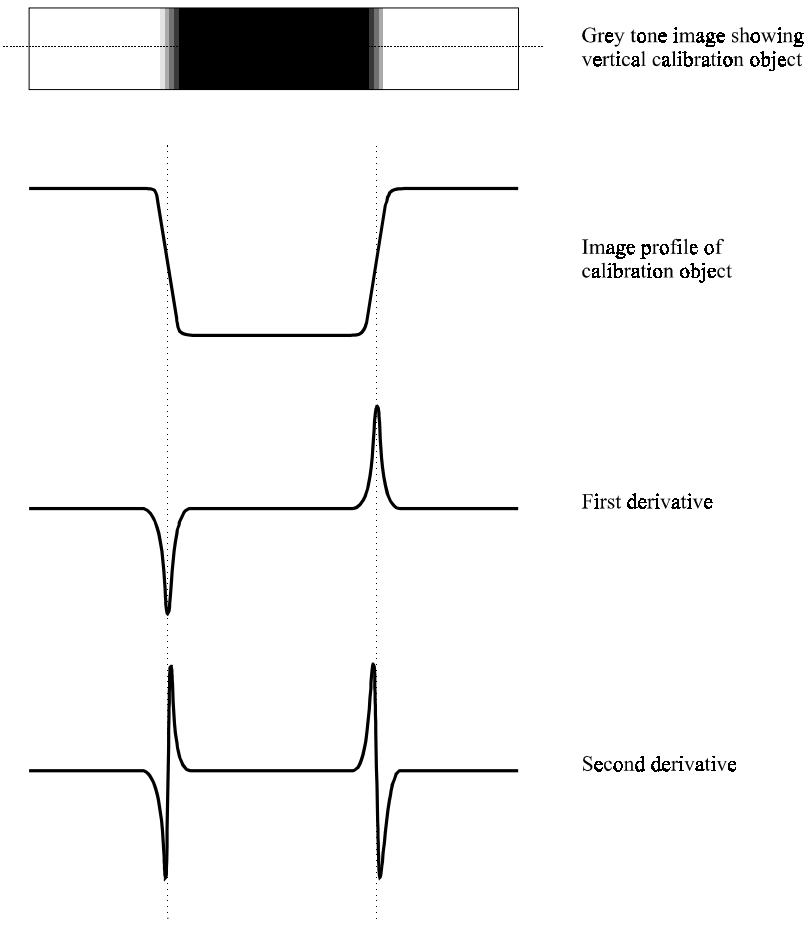


Figure 4.21 An illustration of an edge profile with first and second derivative, notice the zero-crossing in the second derivative.

When working with discrete images the first and second order derivatives can be estimated in various ways. One group of methods is to make a parametric fit (polynomial or 1-D hyperbolic tangent) to the local pixel and then calculate the derivatives directly from the found parameters, this approach works well with blurred edges but has problems with step-edges [Heikkilä97]. Another group of methods which are more coarse is to use gradient based edge operators like the Sobel (first order) or Laplacian (second order) operators, see [Gonzalez92]. The Sobel operator is normally made as the amplitude $(G_x^2 + G_y^2)^{½}$ of a vertical (G_y) and a horizontal (G_x) component (see figure 4.22 (a)+(b)). In (c) is an example of the 3x3 Laplacian. The use of gradient based edge operator with one pixel accuracy does not necessarily give a bad localization of the whole object if the number of edge pixels is high and the edge localization is unbiased.

-1	-2	-1
0	0	0
1	2	1

(a)

-1	0	1
-2	0	2
-1	0	1

(b)

0	-1	0
-1	4	-1
0	-1	0

(c)

Figure 4.22 (a) Vertical Sobel component (b) Horizontal Sobel component (c) Laplacian 3x3 mask.

After an unbiased detection of a series of edge pixels from a calibration object the next step is to make a parametric fit to the found edge pixels. This step can be performed in various ways depending of the model that should be fitted, most common are the least squares fit. See [Heikkila97b] for more details about various methods and aspects of line and quadratic curve (ellipse) fitting.

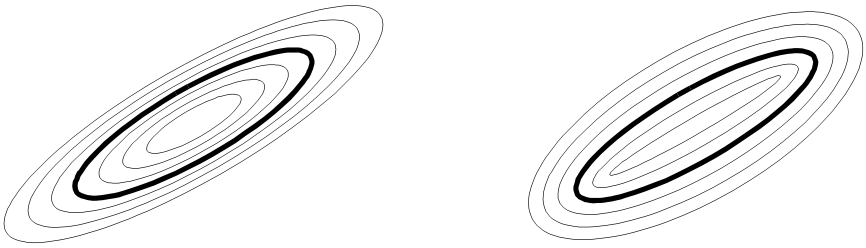
When fitting an ellipse (the common calibration object a circle is through a perspective transformation transformed into an ellipse) to a series of edge pixels there are some very important factors to be considered, first a distance measure from a point to the ellipse must be established. The algebraic distance is given in formula (4.4) where d(x,y) is the “distance” from a point (x,y) to the ellipse with center at (x_o,y_o) and major/minor axis $2\cdot a$ and $2\cdot b$. If using the algebraic

distance for a least square fit there is a direct linear solution based on using a pseudo-inverse, this is unfortunately not sufficient for a high accuracy detection but can be used for an automatic initial guess of the ellipse [Heikkila97b].

$$d(x,y) = \frac{(x - x_0)^2}{a^2} + \frac{(y - y_0)^2}{b^2} - 1 \qquad (4.4)$$

In figure 4.23 (a) is given the equidistant curves from an ellipse using the algebraic distance, as seen this weights points near high curvature different from points near the low curvature (high curvature bias).

The most correct would be to use the geometric distance where each data point is weighted on the orthogonal distance to the ellipse, see figure 4.23 (b). The solution for the geometric distance is quite complicated and requires the solution of 4^{th} order equation, this is why many people has used approximations instead, see [Heikkila97b]. The geometric distance can also be used in a least squares fit but this is highly nonlinear and the minimizing requires an iterative procedure like the Levenberg-Marquardt, this has not been implemented nor further analysed within this project. The algebraic distance can be used as a good starting guess for this iterative search.



(a) (b)

Figure 4.23 Equidistant curves from ellipse (bold) using (a) algebraic distance and (b) geometric distance.

Template Matching

One of the most used methods is a template matching with an artificial object in the image plane. This is typically done by calculating the correlation between a small image segment w(x,y) of size J×K and the examined image f(x,y) of size M×N like illustrated in figure 4.24 [Gonzalez87].

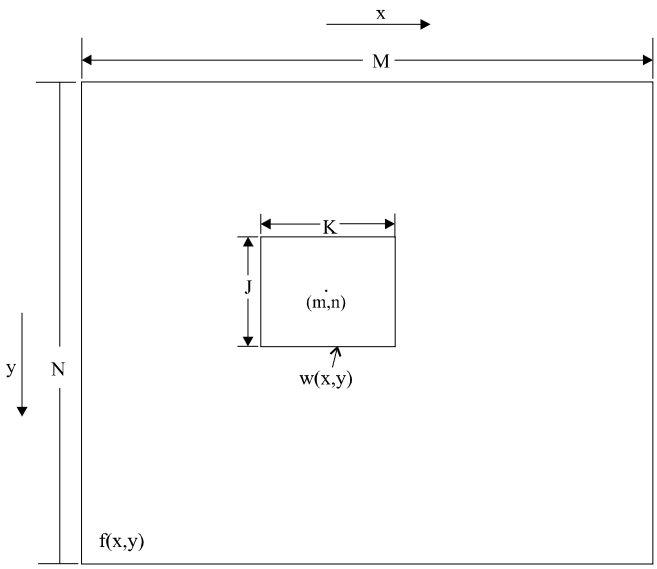


Figure 4.24 Arrangement for the correlation between f(x,y) and w(x,y) at the point (m,n).

To overcome problems with scale changes in amplitude between f(x,y) and w(x,y) the correlation coefficient r(m,n) is often used.

$$r(m,n) = \frac{\sum_x \sum_y [f(x,y) - \bar{f}(x,y)] \cdot [w(x-m, y-n) - \overline{w}]}{\sqrt{\sum_x \sum_y [f(x,y) - \bar{f}(x,y)]^2 \cdot \sum_x \sum_y [w(x-m, y-n) - \overline{w}]^2}} \qquad (4.5)$$

where $\bar{f}(x,y)$ is the average of f(x,y) in the region coincident with w(x,y) and $\overline{w}$ is the average of w(x,y).

The result is an image of correlation coefficients where the peaks are possible candidates for the match. This method has two major disadvantages, the size and the perspective deformation of the objects must be known and almost equal

in the entire image. Furthermore, the accuracy of an object estimation is not subpixel based.

In some applications using this method a two step method is used, first an initial guess in order to estimate the approximate angle and position of the camera and then resample either the image or the template according to the estimated perspective transformation and then redo the template matching.

The template matching can be enhanced to perform an accurate subpixel estimation e.g. by fitting an analytical function to the autocorrelation peak and then move that function in subpixel steps to minimize the squared difference to the original correlation coefficient image, for further details see [Knudsen98].

Another sort of template matching is using a morphological hit and miss transformation, which does not provide an exact match but can be used for e.g. initial detections. The advantages of this method is that it can find objects in a fairly complex scene and the disadvantages is that it works on binary images so the image must first be thresholded, for more details see [Heikkila97b].

Parametric Fitting

A more complex variant of the template matching is called parametric fitting and has been used at the Odense Steel Shipyard for detection of calibration objects made of concentric black and white disks. This method was initially made for straight lines [Blaszka94], but has been improved for ellipses by [Gramkow98]. The idea is to project the image pixels to a calibration object coordinate system using the initial guess of the camera orientation and rotation, in this system the calibration objects have a known shape and size and the image data can then be used for a parametric fitting to a function suitable for the calibration objects. The disadvantages of this method is the need for initial good guesses about the camera position and rotation. The advantage is a very high precision subpixel detection.

Bayesian

Another complex method is the *Bayesian Point Matching* approach based on the Bayesian restoration methodology, first introduced in [Carstensen96] and further developed and described in [Hartelius96]. In this method a grid of disks is regarded as a deformable template modelled by a Markov Random Field (MRF) which explains the interaction between a disk (grid node) and the four adjacent disks in the grid. An observation model describes how the disks are depicted in the images. This method has been used with very good results for automatic analysis of images of hybridisation filters used for DNA experiments [Hartelius96]. The advantages of this method are that it can be used with a wide range of grid configurations and that it is possible to model images with high background variation and other kinds of signal noise through the observation model. The disadvantages are that the initial guess of the grid and objects must be quite good and the perspective transformation small and on a local scale the estimation is only down to one pixel accuracy.

Another use of templates modelled by MRF is for detection of lines after an initial guess based on e.g. a Hough Transform (see later in this chapter). Then the template can be a grid of curves which can curve according to e.g. an ordinary radial lens distortion.

The Hough Transform

When looking for straight lines or line segments in images a useful method is the Hough transform, this could e.g. be the case where the calibration target is a grid of lines or the checkerboard target combined with an edge detector. The Hough transform is a method for detecting straight lines or curves from discrete binary images [Gonzalez87], it works on a parametrization of the form it should detect. In [Fisker98] is an example of detecting ellipses with great success using the parametrization of an ellipse. When detecting straight lines the common parametrization in $d=u*\cos(\alpha) + v*\sin(\alpha)$ is used, then for each image coordinate (u,v) only a finite number of discrete (α,d) can satisfy the

parametrization and they are added into an accumulator image (in the implementation made in this project this is done according to the actual grey level). In figure 4.25 is an example of a Hough parameter space with a 180x180 quantization from a single line rotated at a 45°angle, as it can be seen in the image the form is a butterfly shape.

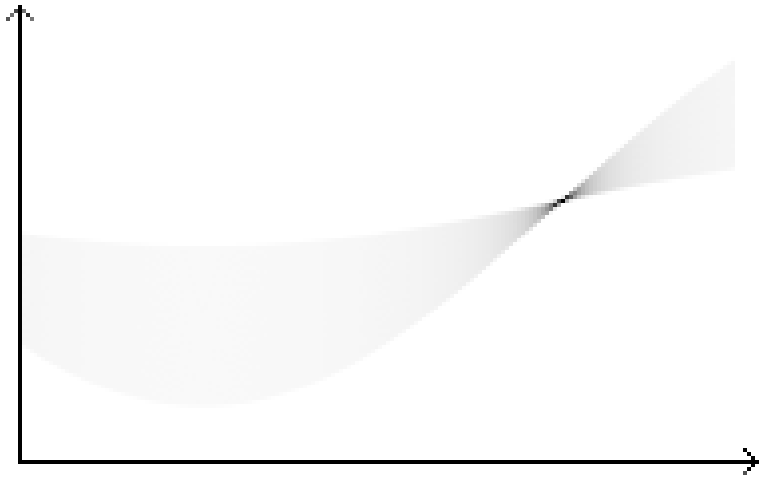


Figure 4.25 The Hough parameter space of a single 45 deg angle line. The quantization is 180x180.

After all pixels in the image have been processed all peaks in the accumulator image represent a possible line in the original image and the equation for the line can be obtained from the accumulator image coordinates. In figure 4.26 is an example of a 4 x 5 grid of lines with the corresponding Hough transform, notice the beautifully found 4 and 5 peaks at two different angles (horizontal axis).

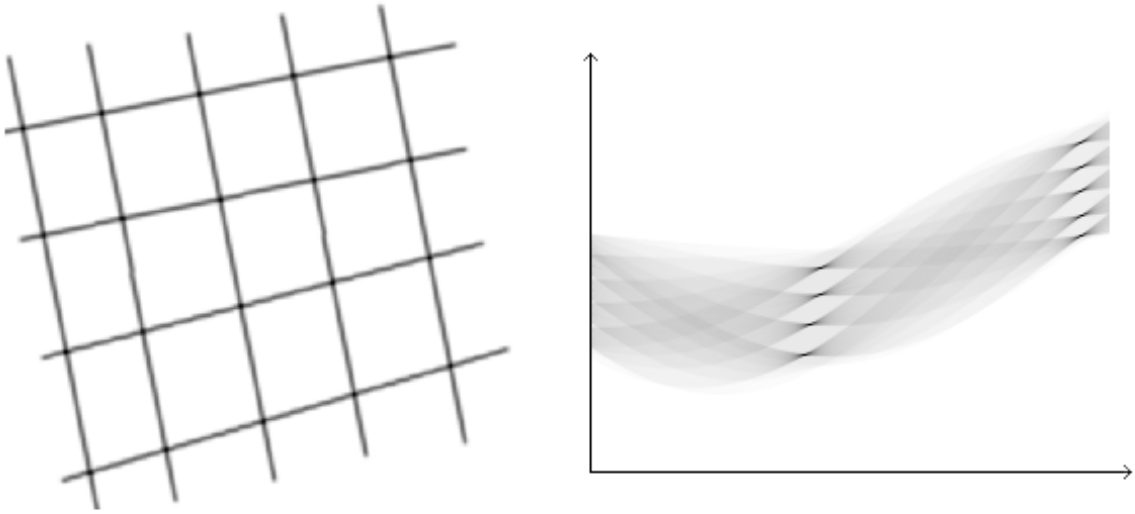


Figure 4.26 A grid of lines rotated about 11 deg. with the corresponding Hough transform at a 256 x 256 quantization. Notice that the number of lines is correct.

Since the method is based on an accumulator that needs a quantization of the parameters the accuracy of the parameter estimates can be rather low. If using a more dense quantization the peaks in the Hough space can be difficult to detect [Heikillä97b] and if using a less dense quantization the accuracy of the parameter estimates gets lower. As a result of this the method is not always suitable for an accurate subpixel detection of lines and further postprocessing using e.g. deformable templates can be needed after the Hough transform line detection. Another disadvantage is that straight lines from a calibration target with normal lens distortions are curves in the image plane and the Hough Transform can therefore cause trouble in images that are heavily distorted. One big advantage is that it is very suitable for a fully automatic detection of straight lines or line segments.

The Hough transform implemented in this Ph.D. project is based on the method found in [Gonzalez87] with the useful extension that it works directly on a grey scale image on a floating point accumulator image. This is done by finding the maximum I_{max} and minimum I_{min} values in the image either global or locally around the actual pixel being processed (if the image is very disturbed) and then

for each pixel value in the image I_i to add the value $A_{acc}= (I_i - I_{min}) / (I_{max}-I_{min})$ (which is between 0.0 and 1.0) to the floating point accumulator. The main advantage of using grey levels instead of simply the found points that possibly are on a line (from e.g. a threshold) is that all intensity values on the lines are used. In figure 4.28 is an example of the Hough parameter space using grey levels on an image of a piece of squared paper (see figure 4.27) with slightly uneven light. The quantization was 512 x 512 giving an accuracy on the angle estimation of about 0.4 deg if the peaks can be detected with one pixel accuracy. As it can be seen (print quality may destroy this) almost all lines are detectable and a simple threshold is able to find 14 lines in each direction, another method for automatic detection of the peaks is a butterfly fit see [Lassen94] for more details. The length of the lines is highly correlated to the level of the peak as it also can be seen in figure 4.28 where the peaks/butterflies at the bottom and top (representing the lines near borders) are weaker, in [Lassen94] is a method for calibration for this phenomena.

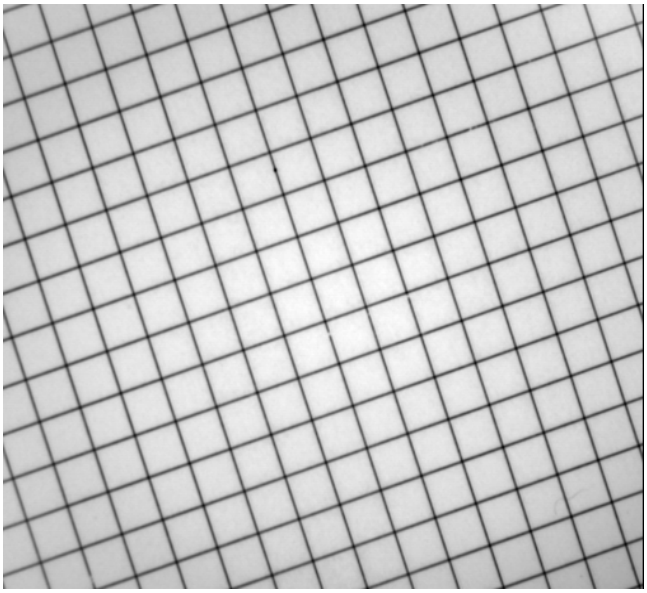


Figure 4.27 A grabbed image showing a piece of a squared paper.

As described in the previous pages the number of different calibration targets and detecting methods are large and the choice of target and detection method must therefore be considered carefully for each problem with respect for the wanted accuracy and speed.

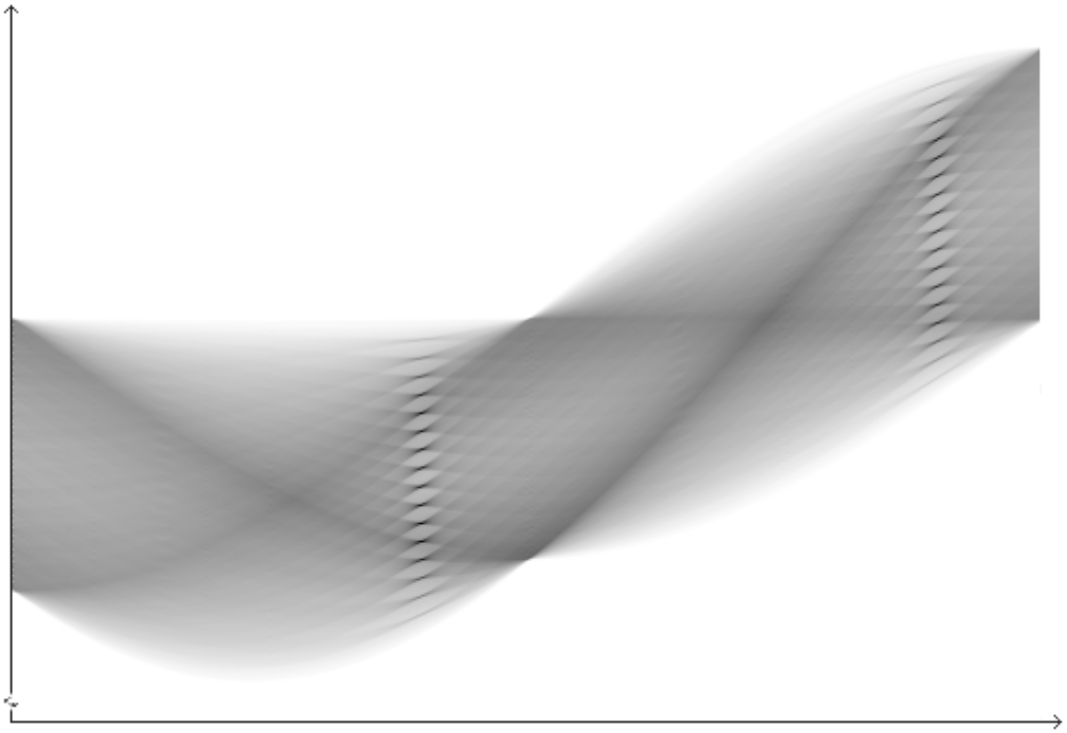


Figure 4.28 Hough parameter space of the image of a piece of squared image, made using the grey level accumulation.

4.4 Uneven Condition Displacements

When the conditions for the image plane or the calibration target plane have some degree of unevenness miscellaneous errors may appear. The design and detection method of these calibration targets have high influence, and correct consideration about the calibration target and camera equipments can reduce or remove these errors.

4.4.1 Uneven Distributed Light

Unevenly distributed light over the calibration sheet can give a small but systematic error when detecting objects. In figure 4.29 is an illustration showing a continuous profile of a step-edge object with uneven intensity (the narrow profile) overlapped with a blurring (the bold line) which could be a normal lens blurring. As seen in the figure applying a threshold value to detect the object position will give a significant shift away from the illumination (the high intensities), the proportions are of course enhanced in figure 4.29 and in practice this will be a small but biassed displacement. The displacement will also occur when using template matching or parametric fitting. In cases with detection of step-edge objects without blurring this error is zero, but absolute zero blurring is almost impossible in real world applications. The use of edge based detection methods like the first or second order derivatives will not give any error with unevenly distributed light when using step-edge objects, but the work with non-step edge objects made in this project has shown a small displacement even when using the first and second order derivatives. This displacement will of course also occur when using threshold, template matching or parametric fitting on non-step edge calibration objects. Again another good reason to use step-edge objects in a calibration target.

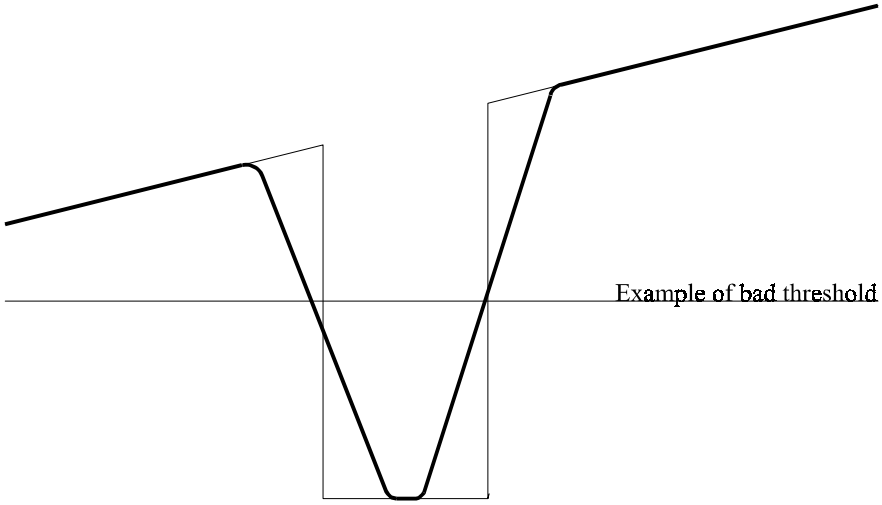


Figure 4.29 Intensity profile of step-edge object illuminated from the right side (narrow line) overlapped with blurring (bold line) giving a displacement in a threshold positioning.

In [Heikkilä96b] are used a grey scale weighted center of mass on images of white disks on a black background, the article describes a systematic shift of about 0.02 pixels in experiments just with the $\cos^4$law falloff, this illustrates that the use of grey scale weighted center of mass estimations also suffers from displacements. This is actually worse when using white objects on a black background (inverse of figure 4.29) than black objects on a white background, because the unevenness (on top of the object) in absolute pixel values is higher for bright pixel values than for dark pixel values, this can be noticed in figure 4.13. Instead of using white disks on a black background digitally inverted images of black disks on a white background can be used for a grey scale weighed center of mass instead.

This problem applies to almost all shapes of calibration object, in figure 4.30 is shown how it affects a step-edge square that has been blurred.

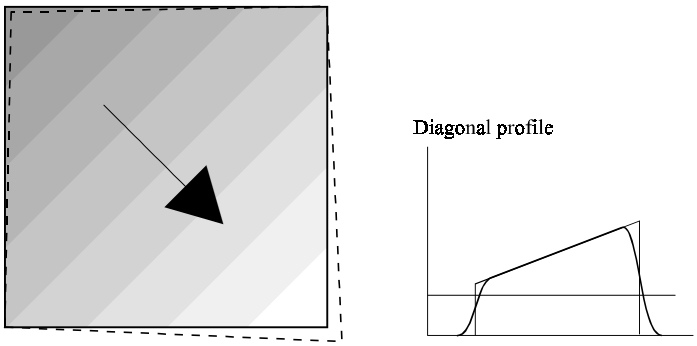


Figure 4.30 A step-edge square, blurred and with uneven light (a). The dashed line is the threshold detected line, (b) the profile of the diagonal.

The displacement from the continuous profile in figure 4.29 has unfortunately a parallel digitalization failure originating from small differences in the grey tone levels on the object edge, this is illustrated in figure 4.31 where a step-edge dark object on a bright background is illuminated from the right side and the digitalization (the dark bars) show a one pixel shift away from the high illumination when using some threshold values.

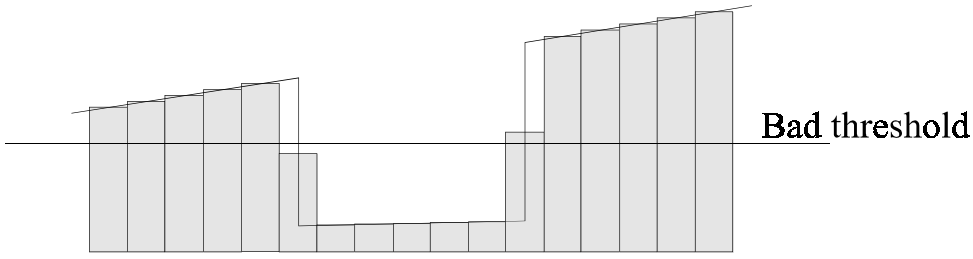


Figure 4.31 Small subpixel displacements stemming from uneven intensities in the image and the digitalization when positioning objects with e.g. threshold.

The effect of this error is that all uneven distributions of the intensities in the image plane will cause a displacement of the positioning of the calibration objects with both step-edge and round-edge objects when using a detection method like threshold, template matching and parametric fitting. When using

edge-based method like first and second order derivatives this error seems to be zero but it has not been tested extensively in this thesis as the problem is more complicated than illustrated in figure 4.29 and 4.31.

In order to illustrate the scope of this error figure 4.32 shows two images of a calibration target grabbed with the same conditions except that (a) is illuminated from the left side and (b) from the right. When applying an automatic detection algorithm like the one described on page 139 (threshold based without correction for uneven lighting) to each of the two images an average horizontal shift of 0.19 pixel and a vertical of 0.009 from (a) to (b) is observed. This experiment has been repeated several times and from different angles with approximately the same result, but since the 0.19 pixels only correspond to about 0.03 mm the setup had to be kept completely free of vibrations which was quite difficult because the light source had to be moved from one side to the other, many of these experiments was therefore unuseable.

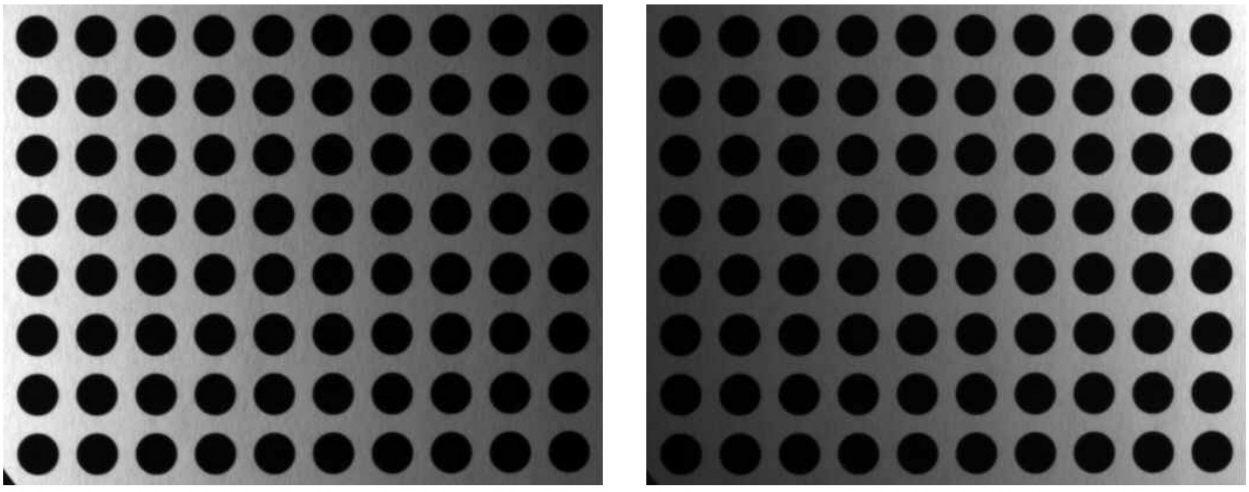


Figure 4.32 Two images grabbed with same conditions except illumination from the left side in (a) and from the right in (b).

In the tables below are the horizontal and vertical shifts between the detected positions of the 80 objects from figure 4.32, as it can be seen the horizontal shifts are all positive and about 0.1 to 0.3 pixels while the vertical shifts mainly are between -0.l and 0.1 pixels.

dx	*28*	*78*	*129*	*179*	*230*	*280*	*330*	*380*	*430*	*480*	avg
27	175	0,236	0,210	0,150	0,130	0,103	0,278	0,244	0,218	0,188	0,193
77	0,275	0,148	0,170	0,261	0,230	0,231	0,316	0,220	0,076	0,165	0,209
127	0,237	0,196	0,107	0,232	0,228	0,186	0,286	0,144	0,242	0,078	0,194
178	0,246	0,157	0,151	0,221	0,185	0,167	0,124	0,104	0,091	0,172	0,162
228	0,252	0,132	0,240	0,202	0,117	0,164	0,119	0,072	0,276	0,154	0,173
278	0,204	0,231	0,241	0,139	0,115	0,126	0,124	0,249	0,259	0,166	0,185
329	0,202	0,218	0,164	0,226	0,231	0,187	0,231	0,331	0,164	0,203	0,216
379	0,089	0,225	0,155	0,203	0,311	0,253	0,200	0,272	0,126	0,139	0,197
avg.	0,210	0,193	0,180	0,205	0,193	0,177	0,210	0,205	0,181	0,158	191

dy	28	78	129	179	230	280	330	380	430	480	avg
27	0,000	0,019	0,041	-0,048	-0,043	-0,085	-0,036	0,010	0,082	0,012	-0,005
77	0,023	0,060	-0,013	-0,028	0,045	0,069	-0,067	-0,072	-0,087	0,090	0,002
127	-0,040	-0,029	-0,013	0,015	0,070	0,030	-0,001	-0,090	0,024	0,005	-0,003
178	0,003	0,027	0,016	0,039	-0,017	-0,015	0,026	0,003	-0,026	-0,010	0,005
228	-0,076	-0,065	-0,016	-0,016	0,017	-0,037	-0,039	0,015	-0,072	-0,078	-0,037
278	-0,058	0,001	-0,015	0,052	-0,046	-0,031	-0,088	-0,065	0,158	-0,014	-0,010
329	-0,025	0,055	-0,044	-0,027	0,020	-0,021	0,011	-0,089	-0,051	-0,074	-0,025
379	-0,038	-0,009	-0,025	-0,017	0,126	-0,022	0,036	0,048	-0,050	-0,004	0,004
avg.	-0,026	0,007	-0,009	-0,004	0,021	-0,014	-0,020	-0,030	-0,003	-0,009	-9

Tables showing horizontal (dx) and vertical (dy) shift (pixels) in positioning of 80 calibration objects between two image illuminated from the left and right side.

In figure 4.33 is a vector field plot of the positioning shift from the right side to the left side light. As seen the difference between the individual positioning is quite disturbed but the 0.2 pixel drift in the horizontal direction is obvious.

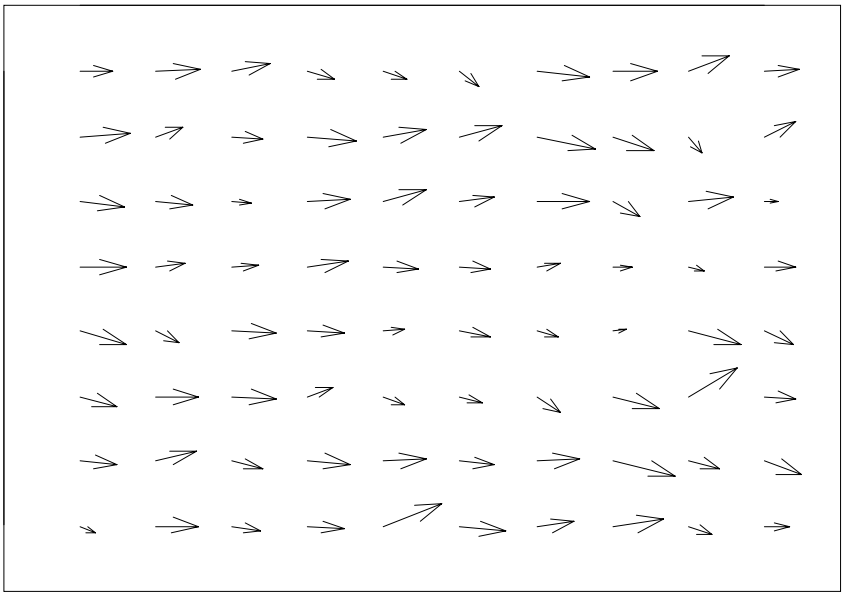


Figure 4.33 Vector field plot of shift in positioning of 80 disks illuminated from right and left side.

Another type of source for uneven lighting (irradiance) across an image plane is the $\cos^{4th}$ law of light falloff from the principal point, this is a very fundamental optical phenomena and is present in almost all lenses (see page 85). For ordinary lenses used in machine-vision this is a very common error and the contribution must be taken into consideration, the $\cos^4$ light falloff is especially critical with wide angle lenses, e.g. with a total viewing angle of 60^0 the irradiance would be 56 % ($cos^4(30^0)$) lower at the image border than the center.

A phenomena called vignetting is also a source for uneven irradiance across the image plane, this a frequent lens system design problem which almost can be ignored with very good designed lenses (normally the expensive lenses). Vignetting is caused by cutting off parts of the rays going into the lens system, for a more detailed description see page 87.

4.4.2 Reflection

In figure 4.34 is a 2D illustration of a light source hitting a surface dividing the non-absorbed light into a specular reflection and a diffuse reflection. The specular reflectance will come out in the same angle as the entrance angle of the light source (angle of reflection) while the diffuse reflection will spread out in all directions. This can be observed when pointing a laser dot at a glossy planar surface where the laser dot can be seen at almost the same intensity in all viewing angles but in the angle of reflection your eyes will be dazzled.

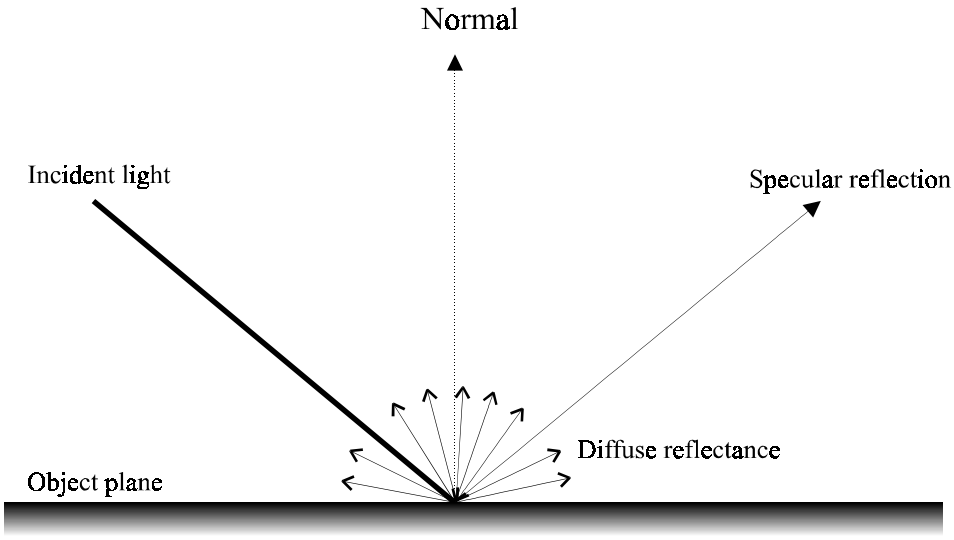


Figure 4.34 Refraction of a light source hitting a planar surface.

The specular reflections can give problems with a calibration target in the sense that it can cause uneven intensities of equal objects/backgrounds in the image which can give displacements of the positioning of the calibration objects similar to the displacements described in the previous section of this chapter. The specular reflections can of course only occur if the target (objects and/or background) has some degree of glossiness and the light source is not completely diffuse. Problems with specular reflection occur because of the camera and/or the light source being relatively close to the target where the possibility of the reflectance being uneven across the image plane is largest. Reflections in an image can give some very high artificial intensities in the

image plane which compared to an uneven illumination, can be much more critical when e.g. detecting a calibration target.

It order to reduce the effects of reflections attention should be paid to the following three items:

- Avoid glossy material for calibration targets, this applies to both background and objects (paint, ink etc.).

- Make the positioning of camera and light source in such a way that specular reflections can be avoided, alternatively use some sort of shields.

- Use light with the highest possible level of diffuseness for illumination of a calibration target.

It is very difficult to make an accurate calibration for the reflection in the image plane since the level of reflections is very dependent on the material. When working with calibration targets it is not unusual that the material used for the background is different than for the objects, this is e.g. the case with calibration sheets printed on a laser-printer where the printed areas (objects) often have a clearly more glossy surface than the paper. Since the exact positioning of objects in a calibration target obviously is unknown and the level of glossiness can be different between object and background there is no obvious way of making an accurate correction of the individual pixels for the uneven intensities originating from the reflections.

Similar to uneven illumination the edge-based methods like first and second order derivatives do not suffer significantly from these effects when using step-edge objects unless a large intensity saturation causes problems or the unevenness is extreme.

4.4.3 Uneven Blurring in the Image Plane

Uneven blurring over the image plane when grabbing calibration targets can give a small but systematic error when detecting objects. In figure 4.35 is an illustration showing a profile of a step-edge object with uneven blurring (the bold line) with three different threshold values. Uneven blurring of the image plane will typically be seen with close range images and lenses not specifically designed for this purpose.

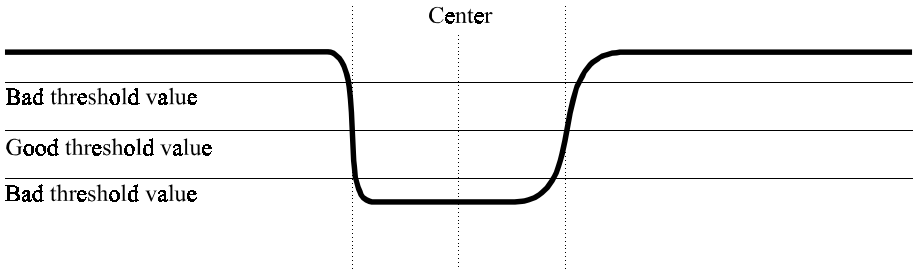


Figure 4.35 Shift in positioning using threshold when blurring is unevenly distributed over the image plane.

Uneven blurring can come into existence with both parallel image and object plane and non parallel when the images are grabbed at close range. With close range images with parallel image plane and object plane the uneven blurring comes from the lens aberrations, in this project a visible uneven blurring has been observed with a 25 mm lens mounted with an 10 mm extension tube at a distance of about 300mm. Another and more significantly uneven blurring will appear when the camera is tilted according to a planar calibration target in such a way that the entire calibration target can not be in focus. This sort of uneven blurring can be very significant in certain configurations and must be considered when detecting the calibration objects, a way to reduce this is to use a Scheimpflug camera system (bent lens) like in the 3D PIV system at Dantec Measurement A/S, see chapter 5.

As in figure 4.35 the threshold method for detecting a calibration target is most likely to have problems with uneven blurring, but other detection methods like template matching and parametric fitting can also have problems depending on

the actual implementation of the method. The use of edge based methods like first and second order derivatives do not have problems with uneven blurring when using step-edge objects.

4.5 Perspective Projection Displacement

When working with calibration targets it is important to realize that the perspective projection not will necessarily conserve a shape. The center of mass (or a template matching using a circle/ellipse) in the image plane of a planar object circle or ellipse will not always correspond to the actual center in the scene. The effect of this phenomena on planar calibration targets is called perspective projection displacement PPD, projection asymmetry or perspective distortion.
When using small calibration objects in a (partly) planar calibration target the PPD is a very small error, it is independent on the camera constant (focal length) and is very dependent on the distance and angle. The PPD is zero when using parallel image plane and object plane.

4.5.1 Illustrative Example

In figure 4.37 is an image of a square with its diagonals drawn on a piece of 5 mm grid paper mounted on a circular and flat surface. Its was grabbed from an approximately 47 deg angle at a distance of about 480 mm using a JAI M50 PAL camera with a 23 mm Schneider lense. The lens was focussed at the 480 mm distance and the corresponding camera constant was therefore about 25 mm. The center of the square is only about 4 pixels from the center of the image and it is therefore very close to the principal point.

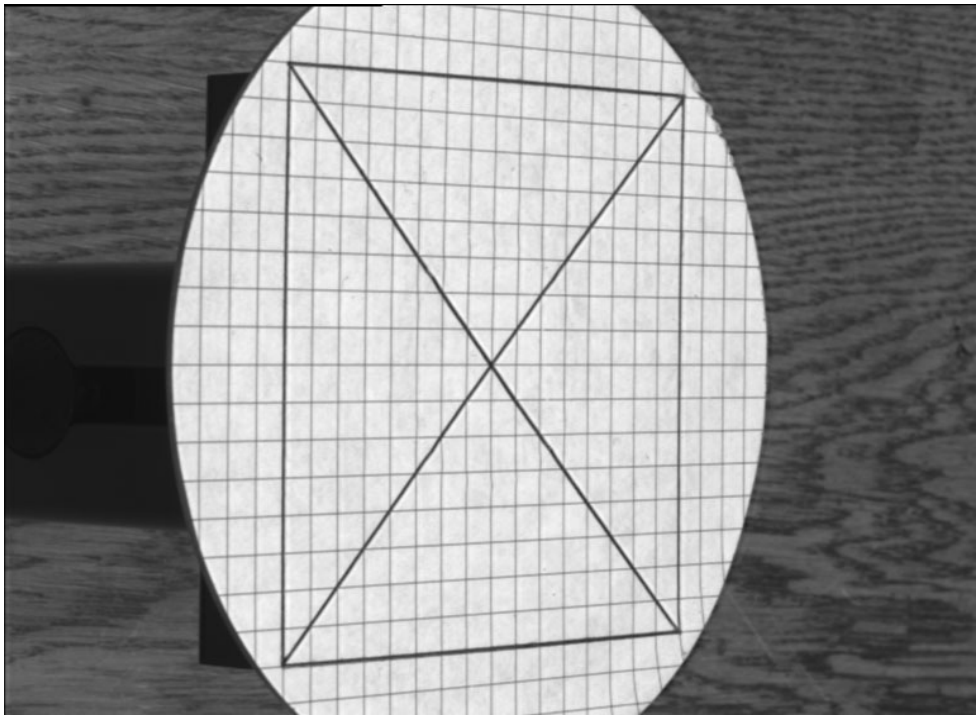


Figure 4.37 80x80mm square target on 5 mm grid, 47 deg angle.

Using an ordinary ruler on figure 4.37 and measuring the distances between the cross of the diagonals and the left (d_2) and right (d_1) side of the square, it is easily seen/measured that the center of the square or the circle in figure 4.37 does not reflect the length divided by two. It is also seen that this phenomena does not occur in the vertical direction (where the camera plane and the image plane are parallel).
When measured on a 197 x 148 mm print of figure 4.37 the distances d_1 and d_2 were 38mm and 42mm, giving that d_2 is 11% larger than d_1. When measuring on the circle in the horizontal direction the matching $d_1´$ and $d_2´$ was 56 mm and 65 mm a relative difference of 16%, notice the higher error for the larger object. It is obvious that the same error would have occurred if the center of mass was estimated from a digital image of the circle or the square. It is therefore very important to consider this error source when using the center of mass or a simple template matching for finding planar objects on calibration targets. In the next section this possible error source will be described in more detail.

4.5.2 A Model for PPD

In order to illustrate the size of this error source and how to calibrate for it in a practical camera calibration this section will give some useful theoretical results. In figure 4.38 below is an illustration of the phenomena in one of the components of the image plane for an object size *2*d* with distance *k* to the center in the object plane corresponding to the principal point. The angle between camera and the planar object is α. The displacement (the difference between d_1 and d_2) is very easy to observe since the camera constant **c** is very large compared to the object distance.

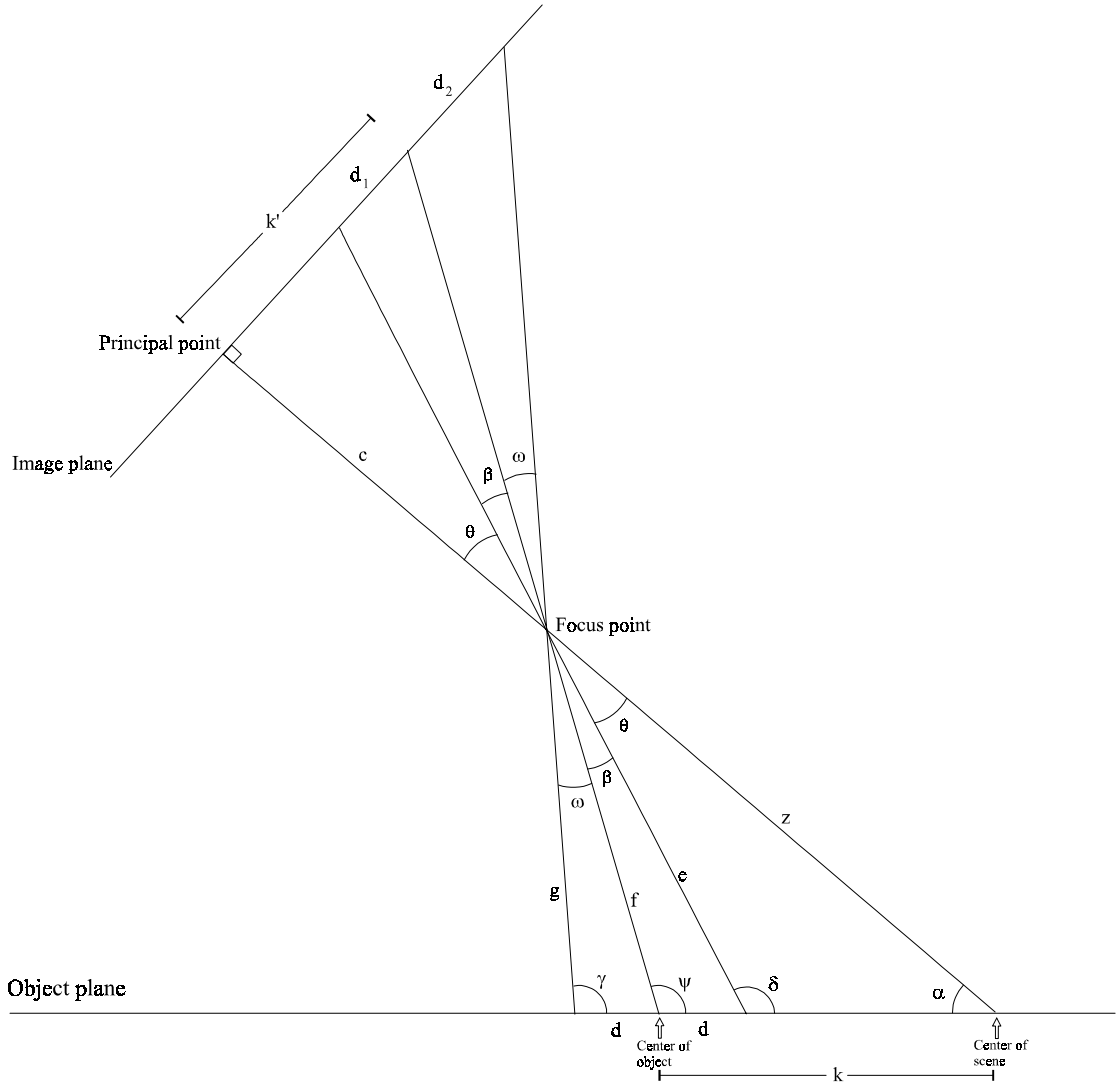


Figure 4.38 Perspective projection displacement

When assuming distance z, camera constant c, angle α, distance k and object radius d known, it is possible to make a simple model of the system illustrated. When using the law of cosines (triangle ABC: $c^2 = a^2 + b^2 - 2ab\cos(C)$) on the lower part of figure 4.38 we get the formulas in (4.6). The formulas in (4.6) are valid for all values of k, if k is positive when the object center is on the left side of the center of scene and negative (with e, f, β and ω mirrored) on the other side.

$$
\begin{aligned}
e &:= \sqrt{z^2 + (k - d)^2 - 2\cdot z\cdot(k - d)\cdot\cos(\alpha)} \\
\delta &:= \operatorname{acos}\left[\frac{(k - d)^2 + e^2 - z^2}{2\cdot(k - d)\cdot e}\right] \qquad \text{for } k \neq d \\
\delta &:= \pi - \alpha \qquad \text{for } k = d \\
f &:= \sqrt{e^2 + d^2 - 2\cdot e\cdot d\cdot\cos(\pi - \delta)} \\
\psi &:= \operatorname{acos}\left(\frac{f^2 + d^2 - e^2}{2\cdot f\cdot d}\right) \\
g &:= \sqrt{f^2 + d^2 - 2\cdot f\cdot d\cdot\cos(\pi - \psi)}
\end{aligned}
\tag{4.6}
$$

Using the law of sines (triangle ABC: $a/\sin(A) = b/\sin(B) = c/\sin(C)$) on the upper part of the lower triangles in figure 4.38 we get the three formulas in (4.7).

$$\theta := \operatorname{asin}\left[\frac{(k - d)\cdot\sin(\alpha)}{e}\right]$$

$$\beta := \operatorname{asin}\left(\frac{d\cdot\sin(\pi - \delta)}{f}\right) \qquad (4.7)$$

$$\omega := \operatorname{asin}\left(\frac{d\cdot\sin(\pi - \psi)}{g}\right)$$

Using the well known trigonometric function *tangent of angle A equals the opposite divided by the adjacent* for the two right angle triangles in the upper part of figure 4.38 and in the lower part the fact that the sum of the angles in a triangle equals π we get the formulas (4.8).

$$d_1 := c\cdot(\tan(\theta + \beta) - \tan(\theta))$$

$$d_2 := c\cdot(\tan(\theta + \beta + \omega) - \tan(\theta + \beta)) \qquad (4.8)$$

$$PPD := \frac{d_2 - d_1}{d_1}\cdot 100$$

As it can be seen from the PPD in (4.8) the relative displacement is only dependent on the distance between the focus point in the camera, the angle between the image plane and object plane and the size of the object (which is very understandable when looking at figure 4.38). This is a very important result since it says that the relative displacement is independent on the lens.

In figure 4.38 is a plot of PPD versus the angle between the optical axis in the camera and the object plane. The object radius *d* was 4 mm and the distance *z* between camera and center of object plane was 480 mm, the three plotted lines are at k=-60, k=0 and k=60 showing a 120mm scene. In this example the PPD is at the maximum 2% which in pixels (using a PAL camera disks would be about

17 pixels in radius) would have been around 0.3 pixels. When working with more normal angles around 90^0 to 45^0 the error would be below 0.2 pixels.

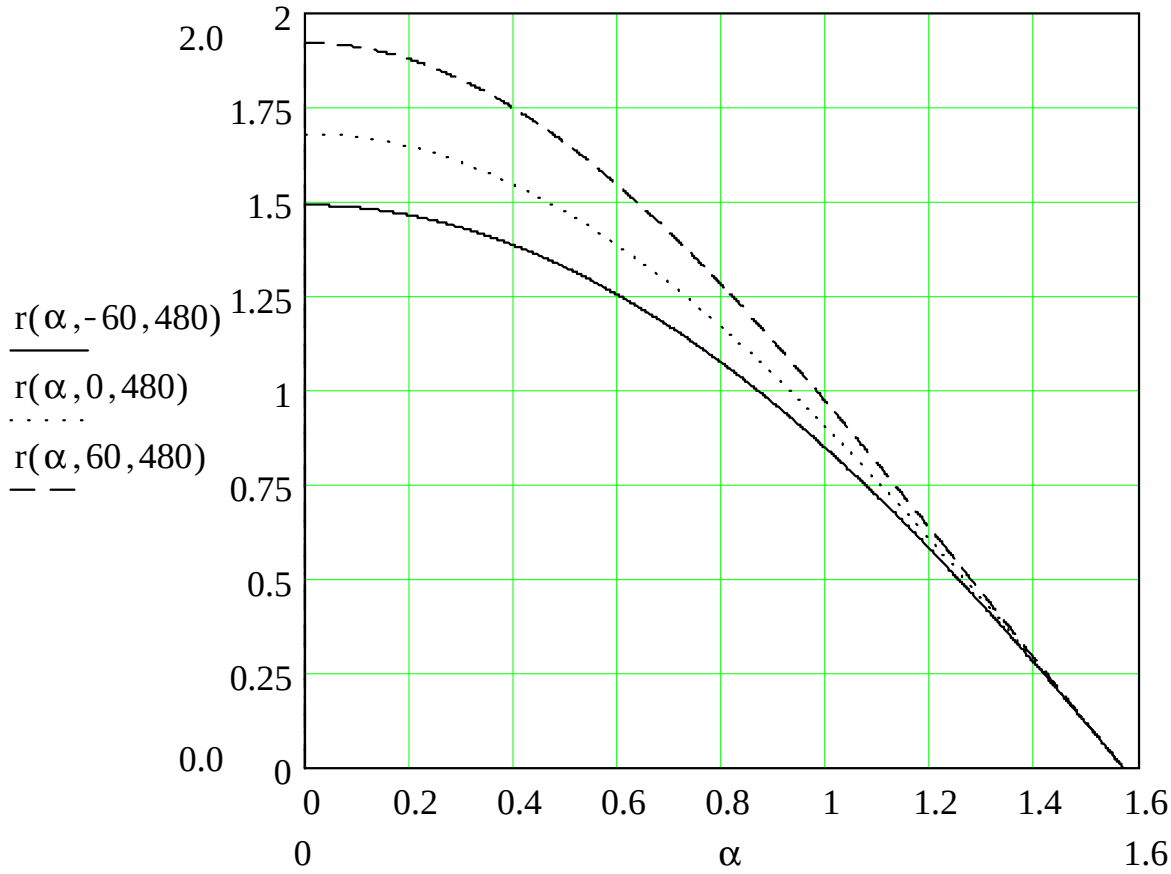


Figure 4.38 PPD from angle 0 to π/2 for 3 different k values.

In figure 4.39 is a plot of PPD versus the distance z between camera and object. It is plotted for three different angles 1.4 (80^0), π/4 (45^0) and π/8 (22.5^0) presenting the most common range of angles, object radius *d* was 4 mm.
At a distance of 1.5 metres the PPD is already below 0.5% (0.1 pixels with disks at radius 20 pixels) when using object radius 4mm. As proven in the PPD model this is independent on the lens (camera constant), and only the size of the object should be considered.
The larger the distance between camera and object the larger the objects must be in order to have a useful pixel size in the image. But in general it is better to

have small objects at a large distance.

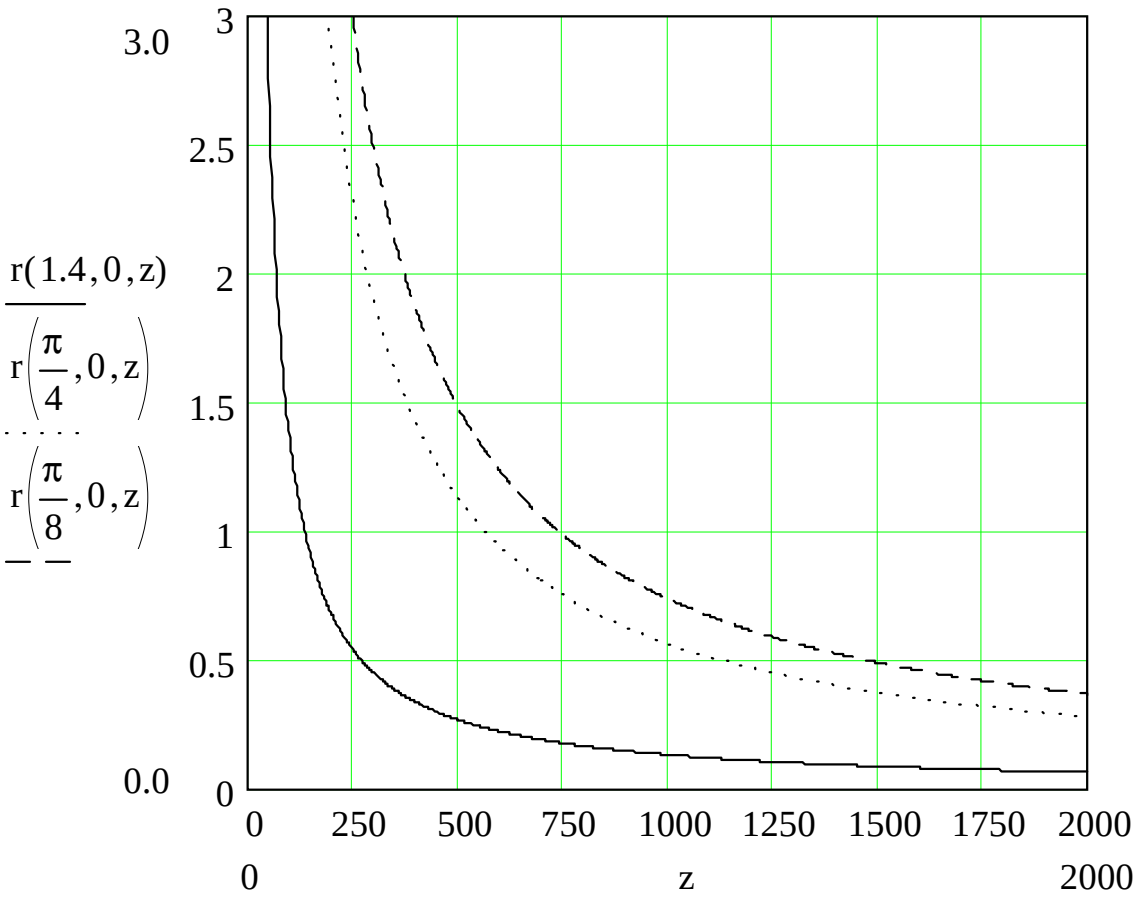


Figure 4.39 PPD from distance 0 to 2m, three different angles.

In figure 4.40 is a plot of PPD versus the object distance k to the center of the scene, showing the change in PPD over the image plane. It is plotted for three different angles, the distance z between camera and object is 480 mm and the object radius is 4 mm. The important result from this plot is that the lower the angle the more effect from change in PPD over the image plane and for almost orthogonal images the PPD is nearly constant.

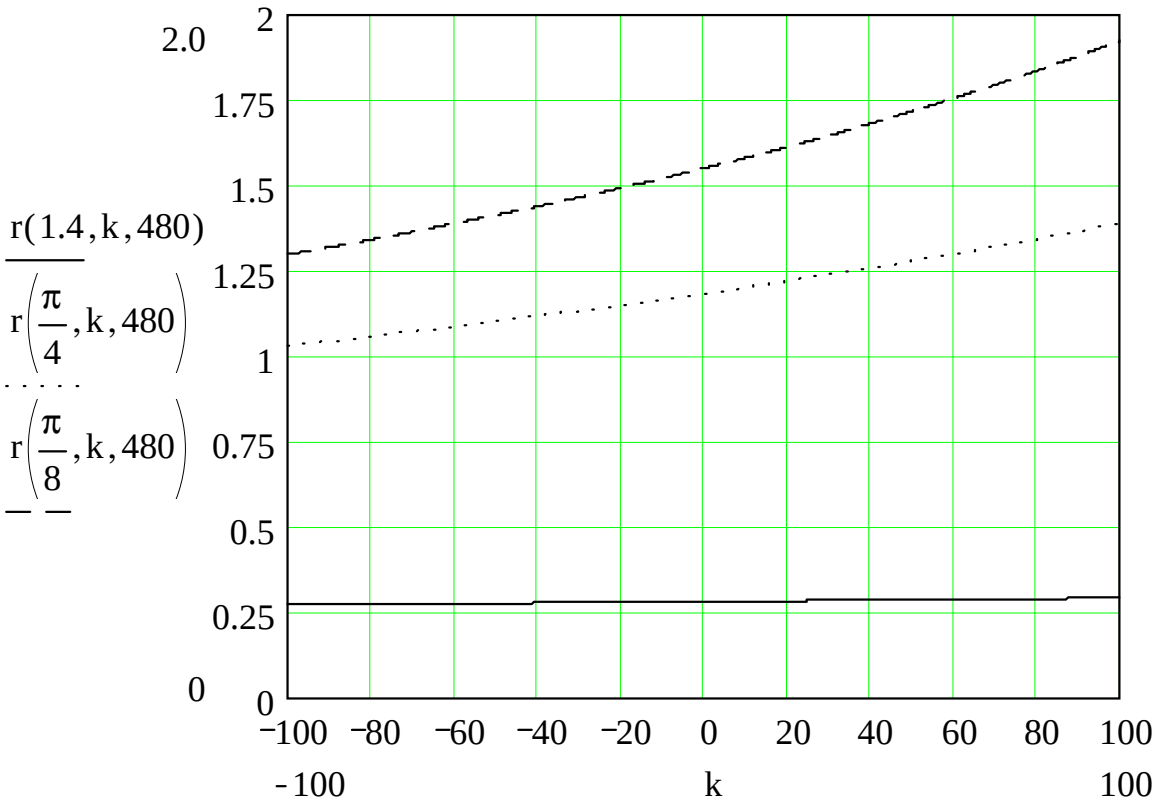


Figure 4.40 PPD for a 200 mm scene, three different angles.

The model for PPD described here is very practical when working with calibration targets consisting of a number of disks (because of the perspective projection of disks being ellipses). One important thing to remember when using this model for adjusting for PPD in an image of a planar target with disks is to find the minor axis of the ellipses in the image plane and rotate the displacements so that they are along the minor axis of the ellipses.

A more complete mathematical model of the perspective projection of a circle in 3D can be found in [Heikkilä97a], this model can either be used directly in a non-linear least squares fit for a camera calibration or used recursively for a correction of a set of center of masses (given an initial guess about the distance and angle between camera focus point and the image plane). The formulas for this model are quite complex and the transformation matrix between the image coordinates and the world coordinate system with origin in the camera focus point must be known.

One important thing to remember when evaluating the requirement for PPD adjustment is that the error is systematic and therefore can be a considerable error. Do not compare the accuracy of the center of mass measurement on one disk with the PPD because the center of mass error is not normally systematic and it is often used for a least squares fit (over a lot of disks) that will absorb this error while the PPD will effect the least squares fit. The PPD should be compared to the wanted quality of the calibration and measured in e.g. pixels. In [Sternberg97] is an evaluation of the quality of a center of mass measurement on a disk versus the size of the disk, the conclusion was that a disk should be over 5 pixels, but when using a least square fit over a very large number of disks this could probably be smaller.

4.5.3 Parallel Image and Object Plane

In figure 4.41 is a simple illustration of figure 4.38 in the special case where the image plane and the object plane are parallel.

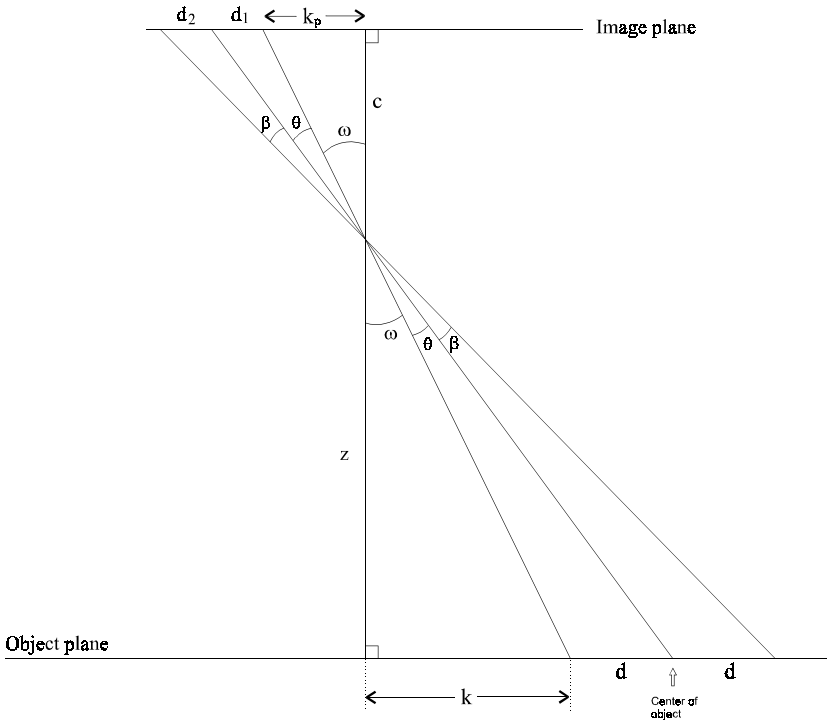


Figure 4.41 Orthogonal perspective

Using some simple geometry on figure 4.41 gives the formulas in (4.9) which solved for d_1 and d_2 gives the result in (4.10) which says that they always are equal.

$$\begin{aligned} \frac{z}{k} &= \frac{c}{k_p} \\ \frac{z}{k+d} &= \frac{c}{k_p + d_1} \\ \frac{z}{k+2\cdot d} &= \frac{c}{k_p + d_1 + d_2} \end{aligned} \tag{4.9}$$

$$d_1 = d_2 = k_p \cdot \frac{d}{k} \tag{4.10}$$

This simple important result proves that the perspective displacement is zero when using parallel image and object plane, which is a very common camera setting in many industrial and laboratory vision system. When using parallel image and object plane it is therefore possible to use the center of mass on e.g. circular objects without any systematic displacement.

4.6 Summary

The design and use of calibration targets is a very important stage in a vision system that requires a high quality camera calibration. This chapter has presented some of the most commonly used (partially) planar targets and has given an example of a robust algorithm developed for automatic detecting of planar calibration targets which has been applied in various implementations through the industrial corporation done in this project.

When working with images of calibration targets with various uneven conditions (like light, focus or reflections) over the image plane small but systematic displacements of the calibration objects may occur depending on the detection method used. There has been given a presentation of these aspects which has not been presented previously in the literature (or not been found in this project).

There has been presented a simple and useful model that can be used in many practical situations for calibrating or evaluation of the perspective projection displacement PPD.

Chapter 5

Implementations

The implementations made in this project all include some sort of calibration issue. The main implementations are made on devices for accurate colour texture measurement (Surfix and VMX2000), this includes both geometric and advanced chromatic calibrations. In this chapter are also described two smaller projects involving laser sheets into moving gas or seeded fluids, these projects involved calibration for uneven distribution of the laser sheets light, removal of unwanted objects in front of the images, automatic detection of calibration targets and programming a Direct linear Transform (DLT) for a stereo system.

5.1 Advanced Colour Texture Measurements

A very important part of this project has been the design of appropriately closed scenes and calibration routines for colour image analysis devices. The following sections will describe the history of the development of a commercial device for colour texture measurements and the important contributions made within this project.

5.1.1 The Surfix Texturemeter

The first attempt on IMM to design a commercial device for colour texture measurements was made in the early nineties by Associate Professor, Ph.D. Jens Michael Carstensen in cooperation with 7-Technologies in Copenhagen. From those ideas a prototype called the Surfix texturemeter, based on a light ring and a 3xCCD colour camera, was created, see the photo in figure 5.1.

Figure 5.1 The Surfix Texturemeter prototype.

The idea of the first prototype was to create a handheld device with controlled light conditions (closed scene) and a high quality colour camera. The light source was an external box (the white box in the right side of the photo) with a halogen lamp connected to a bundle of optical fibers, that transports the light into a light ring inside the device. The external box was chosen to avoid heat problems inside the device, unfortunately this and the heavy weight also made it less mobile and the handheld idea was not really achieved.

In figure 5.2 is a coarse drawing of the Surfix device with both the (a) outside view and the (b) inside view with the main parts. The lower internal walls were replaceable and could either be a diffuse white or a non glossy black dependent on the application. The light ring, optical fiber bundle and the source was from Fostec Inc., the light source included a separable aperture and temperature (voltage) adjustment.

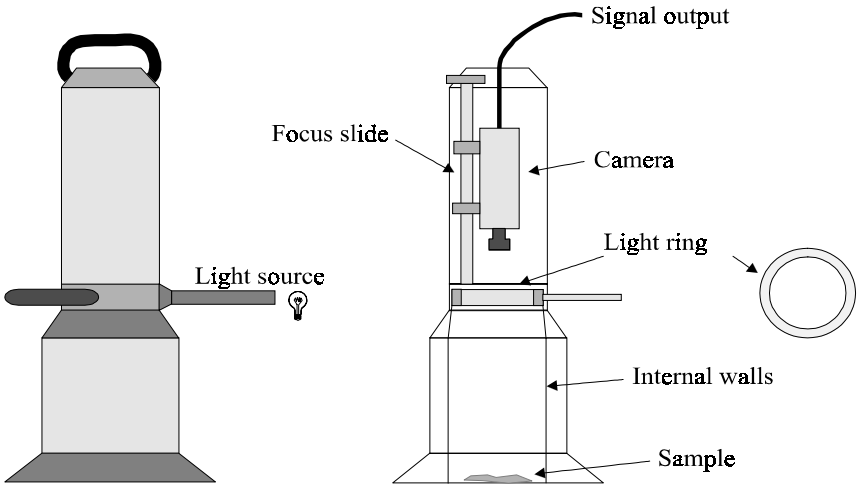


Figure 5.2 Illustration of the Surfix. (a) outside view (b) inside view and main parts. Samples could be about 10 cm.

The camera used in the device was a JVC KY55 3xCCD (a high quality beam-splitter (prism) colour camera) and an Imaging Technology ICPCI frame grabber. The resolution was from about 24 mm to 70 mm in the rectangular 3:4 hole in the bottom. At this early stage in the design of the texturemeter the software was quite primitive, it could grab colour images and through a serial communication with the camera it could adjust the white balance, gain and an optional motor driven lens. This system made it possible to correct the white balance of the camera and to get proper dynamics of the grey levels in the images. However, it was easy to see that the colour shading in the 3xCCD camera was a significant error that needed to be corrected.

It was at this point with a need for some sort of calibration that the author of this thesis on the request from and under supervision of Jens Michael Carstensen was involved in the project. The Center of Information Technology (CIT) also gave a significant financial support to the project at this time.

One of the first tasks was to design a graphical user-interface in a Microsoft Windows® 32bit environment to the Surfix with a new frame grabber (Meteor/RGB) from Matrox. The more user-friendly interface made an explorative investigation easier in this pioneer project. From the previous and new experiments with the Surfix device and the experience with camera calibrations it became clear very quickly that a much more complex camera calibration besides the white-balance and colour shading was needed in order to get repeatable, robust and measurable colour images.

The main problems at that time were :

a) Uneven illumination of the samples from the light ring and the design of the inner of the device.
b) Colour shading, vignetting and $\cos^4$ light fall-off causing uneven intensities in the channels.
c) No absolute pixel values. The possibility to get the same results independent of the time and the physical device (two devices must make the same measurements).
d) Considerable chromatic aberrations up to 4-5 pixels causing false colour on edges in the images.
e) Considerable monochromatic aberrations, causing geometrical deformations with a loss in accuracy of geometrical measurements, that can influence some texture measurements.

In order to make an automatic calibration for the above error sources a planar calibration target with black disks on a white background was designed. This target was placed under the Surfix and an average image was grabbed for the calculation of the calibration parameters. In figure 5.3 is the green channel from such an image (in this case a 16mm lens with 1mm extension tube), notice the heavy barrel distortion and the off-axis blurring that probably is due to heavy

astigmatism. This heavy aberration was due to the use of extension tubes on the lens to make a short distance focus, see the chapter about monochrome calibration. Besides the monochromatic aberrations, the images from the Surfix also had a heavy amount of chromatic aberrations, where the displacement between the blue and red channel in some setups (dependant on lens, extension tube etc) was up to 2.2 pixels near the borders and with an average displacement of about 0.9 pixel, see [Folm-Hansen97].

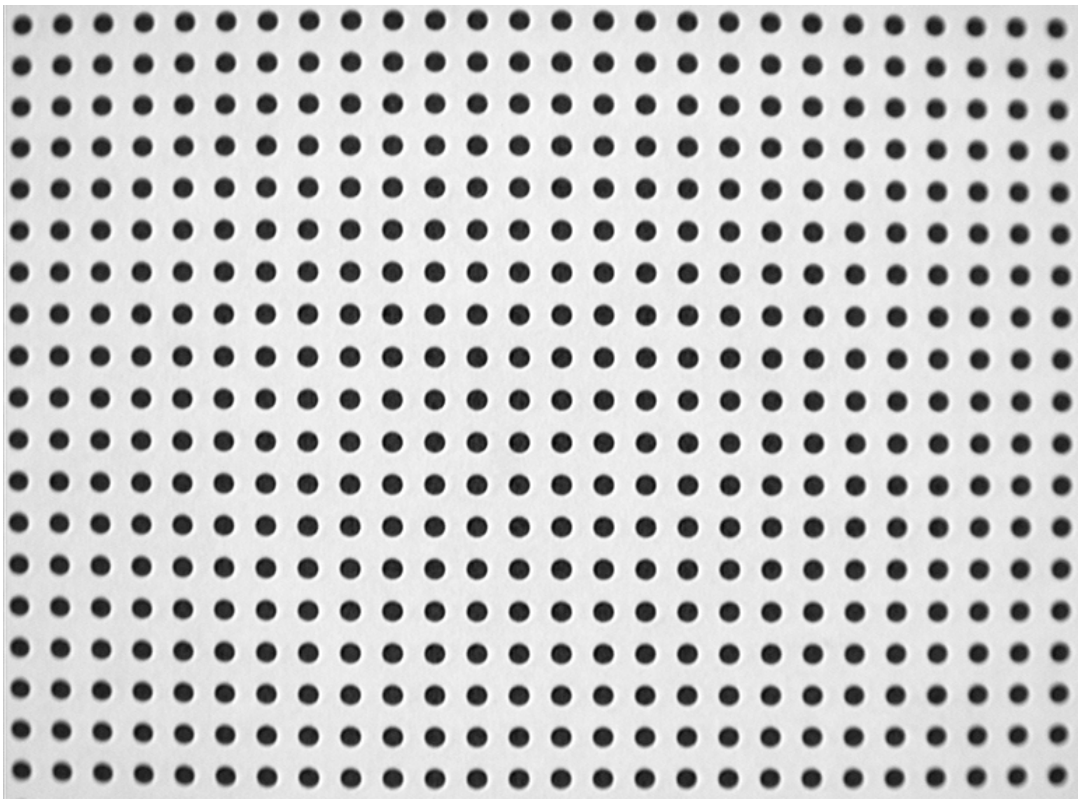


Figure 5.3 A calibration target grabbed with the Surfix device (green channel).

First an algorithm for automatic detection of the black disk was implemented, this was based on the center of gravity of a thresholded image for each channel red, green and blue. The threshold value was found as the value between the two main tops in the histogram and the center of gravity was estimated for each binary object in the label image, see chapter 4.3.2 for further information on this method. Since the image plane and the sample plane are parallel there is no “Perspective Projective Displacement” (see page **160**). From the generated list of center of masses an automatic grid arrangement algorithm was used (see page 137). This also removed outlier objects, based on area and position in the grid.

A polynomial based geometrical transformation (warping) as in equation (**2.36**) was used for the calibration for the monochromatic and chromatic aberrations. The warping was done on each channel separately and the warp parameters were estimated from a least square fit normally based on more than 300 objects (center of gravity of the disks).

In order to get a correction for the uneven intensities in the images (a, b and c), it was necessary to get two reference levels well distributed spatially. Based on the idea that only one calibration target should be used this was done after the above detection of all the black disks, where the mean values of the black 7×7 windows with center in the disks and the mean value of a white 7×7 window between the black disks was calculated.

At this time the uneven intensities (from vignetting, colour shading, uneven illumination etc) were roughly estimated through a 2nd order 2D polynomial as given in equation (5.1) where the $g_Z(x,y)$ is the intensity level for the *Z* material at position (x,y) in the image.

$$\begin{aligned} g_{white}(x,y) &= a_1 \cdot x^2 + a_2 \cdot x \cdot y + a_3 \cdot y^2 + a_4 \cdot x + a_5 \cdot y + a_6 \\ g_{black}(x,y) &= b_1 \cdot x^2 + b_2 \cdot x \cdot y + b_3 \cdot y^2 + b_4 \cdot x + b_5 \cdot y + b_6 \end{aligned} \tag{5.1}$$

The parameters a_i and b_i in (5.1) were estimated through a normal least square fit.

Going from almost no calibration to this method was an extreme increase in quality of the images from the Surfix. The visual improvement was very large and it was now possible to get the same measurements within a few pixel values on the same sample from day to day.

This work was presented in 1997 at both the *DSAGM 1997* conference [Folm-Hansen97] and the *"Vision Day June 1997"* at the Technical University of Denmark (a one day university conference with both industrial and university participants).

The main usage of the Surfix device at this time was textile inspection, especially denim, that was done in cooperation with Novo Nordisk A/S. The requirement for automatic inspection of textiles is quite important for the Novo Nordisk and other companies working with textiles, since the evaluation of the colours of the textiles mainly is done by human visual inspection with very high variations. There have been efforts trying to make an automatic system, based on e.g. a colour flatbed scanner but the results was not good enough.

The further work with the Surfix showed problems with glossy/shiny materials, the specular reflectance from the light ring was to high, the result was that the calibration for uneven intensities did not work correctly in all situations (where the sample material differed too much from the calibration target). This error source was too high to be ignored and another design of the device was initiated.

Almost simultaneously (without our knowledge at that time (1997)) other people at the Center for Industrial Research Applications at West Virginia University was working with a similar product that should be used for measuring colours of blotches and other distinctly-coloured regions on flat surfaces [Pertl97]. Their device looks almost identical to the Surfix device (figure 5.2) and they also compensate for spatial variations in illumination and differences between pixels and RGB channels. Another important thing they have worked on is the ability to assure consistent measurements of differently-coloured regions over a number of samples. They also have a 3 hours warm up before the used light (colour-corrected halogen) is consistent. They also have the conclusion that a colour camera can be superior to a spectrophotometer if proper calibrations are done. They do not consider self-illumination, problems with glossy/shiny materials or advanced geometric calibrations.

5.1.2 The Integrated Sphere/VMX2000

As described in the previous section the Surfix suffered from specular reflections that made accurate measurements with glossy/shiny materials

impossible. The solution to this problem was to avoid specular reflections and at the same time not to have shadow effects on non-planar samples (e.g. textiles).

In figure 5.4 is illustrated the reflection versus shadow problem. If the light comes from above (a) to avoid shadows it causes specular reflections and if it comes in low angles to avoid reflections (b) it causes shadows.

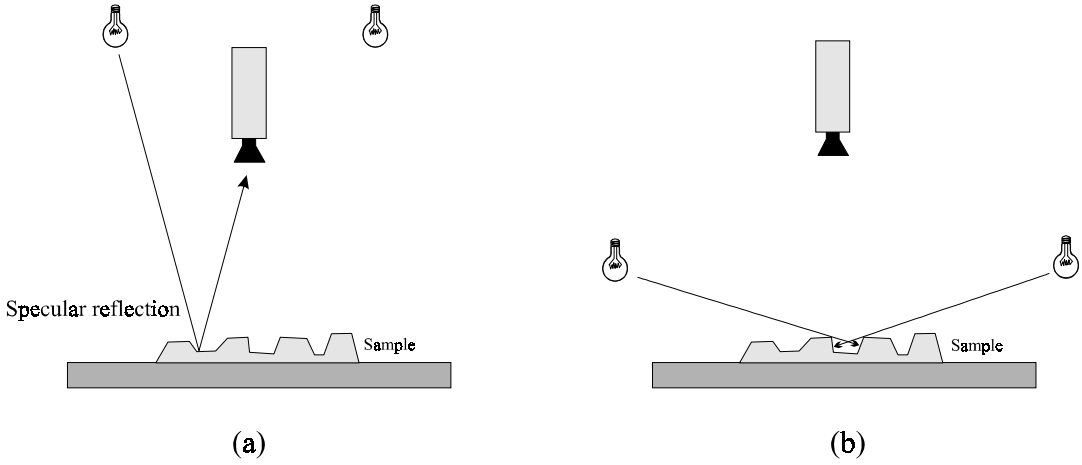


Figure 5.4 Different positions of illumination (a) no shadows but specular reflection (b) no specular reflection but shadows.

The best solution to this problem is a light source where the light comes from all directions at the same time. Such light can be generated using an integrated sphere which is a common device for measuring diffuse light in optics. The integrated sphere normally is a sphere with a white diffuse internal surface, one or more entrance (illumination) ports and one exit (camera/eye) port. The principle of the 360 mm integrated sphere used in this project is shown in figure 5.5.

In the spheres designed in this project four equally spaced entrance ports have been used to generate more uniform light than using one port. The light sources has been placed in the lower part of the sphere and to avoid direct illumination they are sending an oblique light beam at the opposite wall.

The exit port is a 21 mm hole placed at the top of the sphere. To avoid flares from direct light at the camera lens, shields are placed inside the sphere just

below the hole.

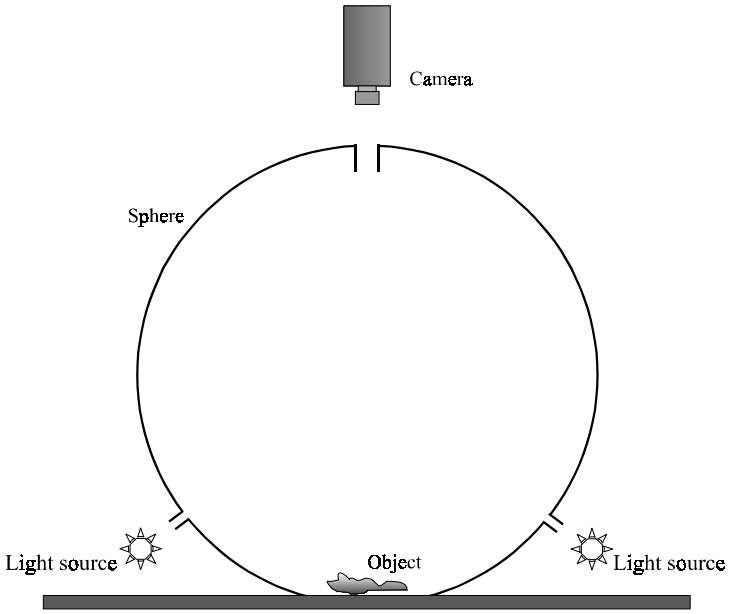


Figure 5.5 Integrated sphere

In order to place the samples inside the sphere a 120 - 140 mm cap has been cut off the bottom of the sphere. In this way the sphere can be placed directly on top of a sample and there is no need for a gate. This also has the advantage that the placement of a sample can be done automatically (it would be quite difficult through a gate in the sphere). With this setup of an integrated sphere a camera can be placed above the small hole at the top and be looking down on the sample placed in the cap, this also gives all the advantages of parallel image and object plane, see chapter 2 on monochrome calibration.

In figure 5.6 is a photo of the sphere used in this project, it is 360 mm in diameter and is made of carbon fiber vacuum cast over a computer made (through a numerically controlled lathe) form. The illumination is generated by the Fostec box in the right side of the photo, the output from the Fostec is split out in the four black cables containing a bunch of optical fibers. Notice the strong light coming out of the cap in the bottom of the sphere and the light emitting at the small hole in the top. As seen in the photo the camera is fixed to the sphere, so that the up/down movement is of both the camera and the sphere.

This was chosen, instead of the simpler method of just moving the sphere, to ensure the same distance between sample and camera regardless of the thickness of the sample and to make it easier to follow the movement of the sphere through the camera.

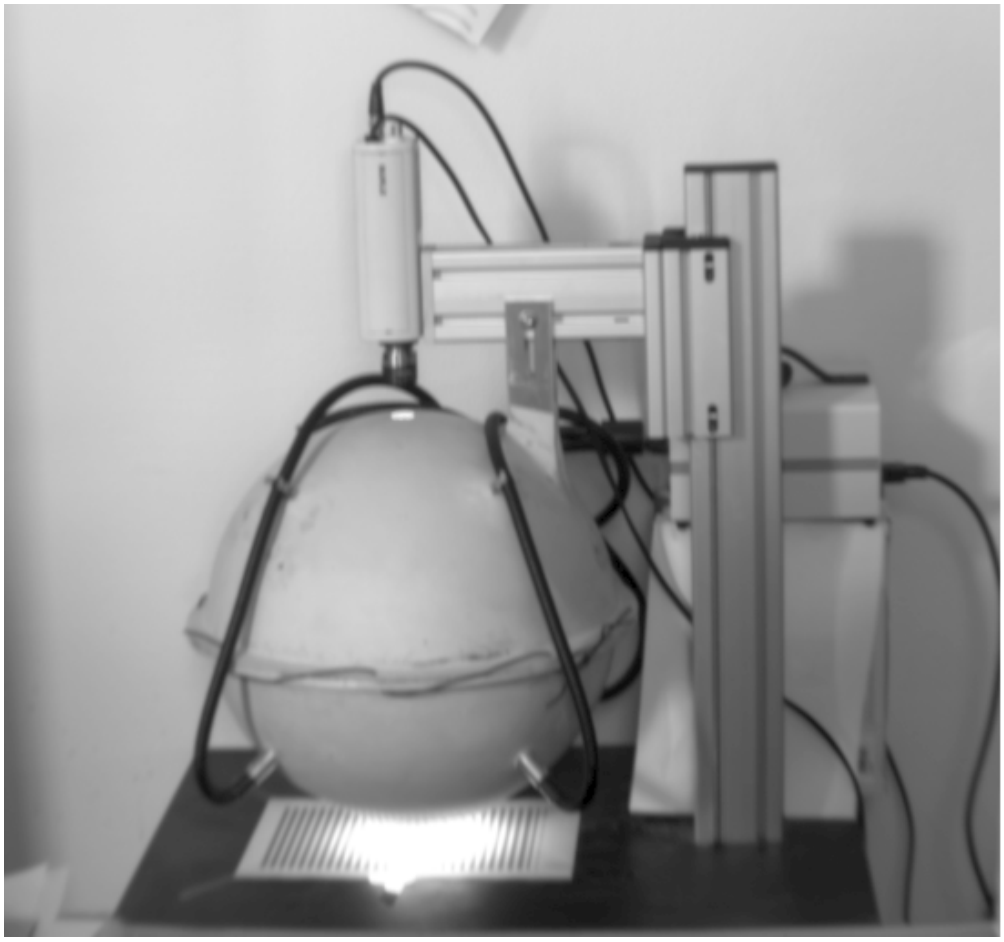


Figure 5.6 The integrated sphere without shields and motor for moving the sphere up and down.

The exit port is a 21 mm hole placed at the top of the sphere. To avoid flares from direct light at the camera lens a small cylinder painted black inside and white outside is placed inside the sphere just below the hole as seen in figure 5.7. The distance between the camera and the top of the sphere is very important for the amount of flare in the images, see figure 5.7. The practical experience with the sphere and the cameras (both JVC 3xCCD and monochrome) has shown very high amounts of strangely formed light stripes that can be removed

by increasing the distance between lens and sphere, the result with this sphere is that a distance of at least 40 mm is necessary to avoid flares.

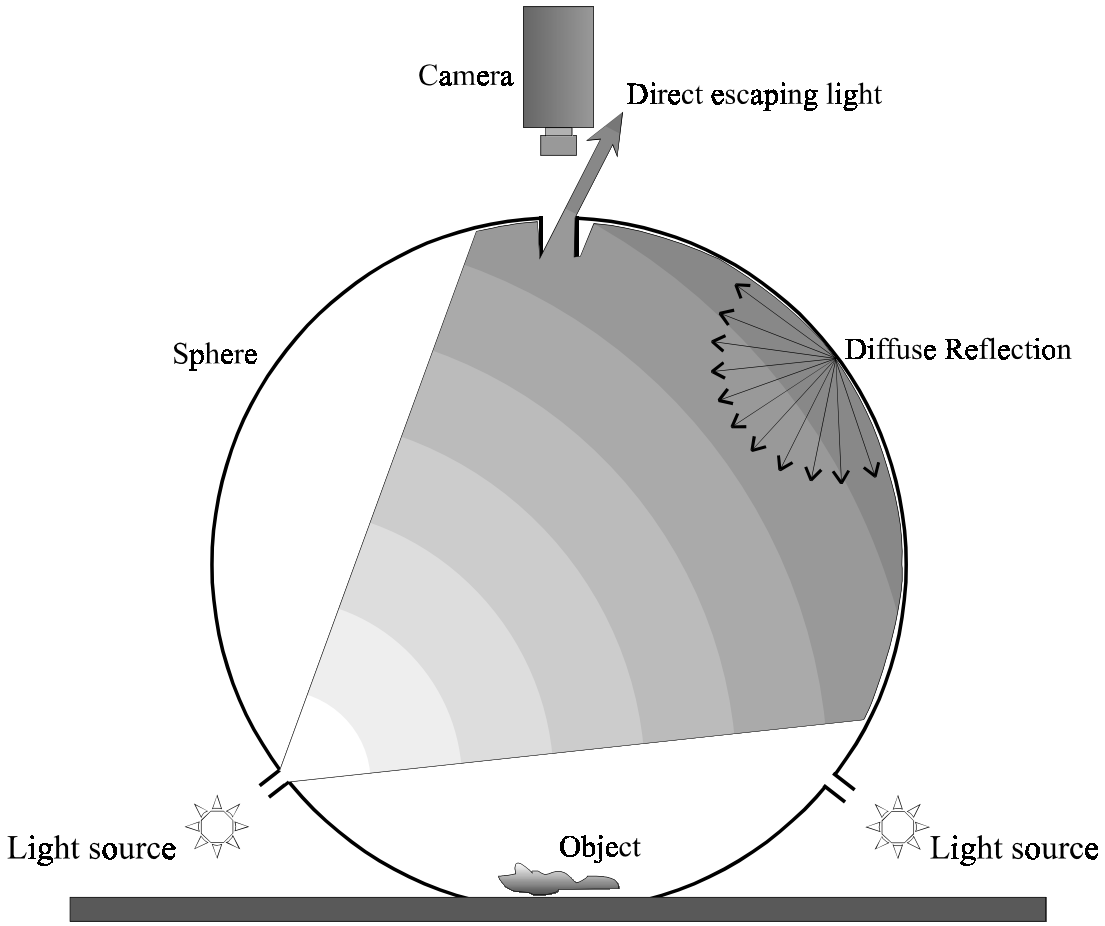


Figure 5.7 Integrated sphere with single ray.

The use of an integrated sphere on the same system as the Surfix device immediately gave better results since the extremely diffuse light inside the sphere removed all noticeable reflections with the samples used at that time. This allowed for a much better and sample glossiness independent calibration for uneven intensities in the images.

In order to get a better calibration for obtaining absolute colour (intensities) values it was decided to use some sort of high quality and common standard colour sheets. There are numerous different systems to choose from and for this project the Natural Colours System (NCS) from Scandinavian Colour Institute

AB Stockholm Sweden was considered the best. The NCS sheets are all painted, not printed, and have very small variations. There are 1750 available colour samples (sheets) with a naming convention that follows a IHS similar numbering. In this project only the nineteen grey sheets, named S0500-N, S1000-N, ... S9500-N (white to black) has been used. It is important to keep in mind that the NCS sheets have been produced to build on how the human eye sees colours and the grey sheets are therefore based on what an average human eye sees as a grey object and not e.g. an equal amount of all wavelengths.

A very interesting thing about the grey NCS sheets was that images of the black S9000-N and S9500-N were almost identical, see figure 5.10 and 5.11. A request to the Scandinavian Colour Institute gave the solution, the S9000-N and S9500-N almost only differ in the glossiness because they had problems getting something darker than S9000-N. The camera in the integrated sphere could not see the difference in glossiness because of the extreme diffuse light. The usable amount of grey NCS sheet for this purpose is therefore only 18.

The 360 mm integrated sphere with the 120 mm cap (140mm in the first sphere with a bottom plate and a rectangular 70x100 viewport) gave very uniform illumination. In order to check the uniformity a monochrome CCD camera with a long focal length lens (50 mm) was moved/rotated to change the captured area, but the images were almost identical. The non uniformity of the images of uniform surfaces like the grey NCS sheets must therefore mostly be caused by camera/lens errors like vignetting, colour shading and $\cos^4$ light fall-off. With an image of a NCS2000-N sheet adjusted so the mean intensity was 130.0 the intensities of the median filtered image (noise reduction) varied from 127.0 at one side to 131.5 at the other side of the 120 mm wide image.

Self-illumination

With the use of an integrated sphere the problem with self-illumination was increased compared to the Surfix illumination. This is due to the nature of the diffuse illumination in the sphere where the light is reflected all around in the sphere and the illumination on the sample has been reflected more than once. In

the Surfix the illumination was more direct from the light source (light ring) and the possibility to use a non glossy black surface inside the device could decrease the self-illumination even further.

For a fixed sample size the self-illumination is of course dependent on the size of the sphere, the larger the sphere the smaller the self-illumination. A large sphere requires more power of the light sources, a larger aperture or a longer exposure time. In general the power of the light in the sphere must be large enough for the specific exposure time and the use of a suitable small aperture in the lens to reduce aberrations. For a given sphere a smaller sample would of course decrease the self-illumination.

In order to check the amount of self-illumination a setup with an 360 mm integrated sphere, a JAI M50 monochrome CCD camera, a 23mm Scheider-Krautz lens and a NCS S0500-N sheet with a 7 mm hole in the center was used. For a fixed setup of the sphere and light source a number of grey NCS sheets was measured without any calibration for the mean value of the total image of the sheet and for three different places the small visible area of the same sheet below the S0500-N sheet with the 7 mm hole. The result of this is shown in table 5.1 where each row represents a grey NCS sheet. The mean value of the small area varies about 0-3 pixel values in different places in the image, as it can be seen from the table.

NCS	Full area	Small area 1	Small area 2	Small area 3	Mean small areas	Difference
500	200.81	169.38	166.87	168.5	168.25	32.56
1000	170.14	147.76	146.04	147.2	147	23.14
2000	128.97	115.89	114.07	114.4	114.79	14.18
3000	101.45	93.23	91.76	92.51	92.5	8.95
4000	76.02	71.71	71.22	71.5	71.48	4.54
5000	60.53	57.11	56.28	57.73	57.04	3.49
6000	46.12	45	43.75	45.17	44.64	1.48
7000	35.55	35.22	34.37	34.95	34.85	0.7
8000	25.19	25.48	25.5	24.73	25.24	-0.05
9000	17.91	18	18	18	18	-0.09

Table 5.1 Grey NCS sheet grabbed in full size and under S0500-N sheet with a 7 mm hole (small area).

The results from table 5.1 is plotted in figure 5.8 where the two interesting things can be seen. The increase in self-illumination is non-linear and the grey NCS-sheets are clearly non-linear relative to the pixel values.

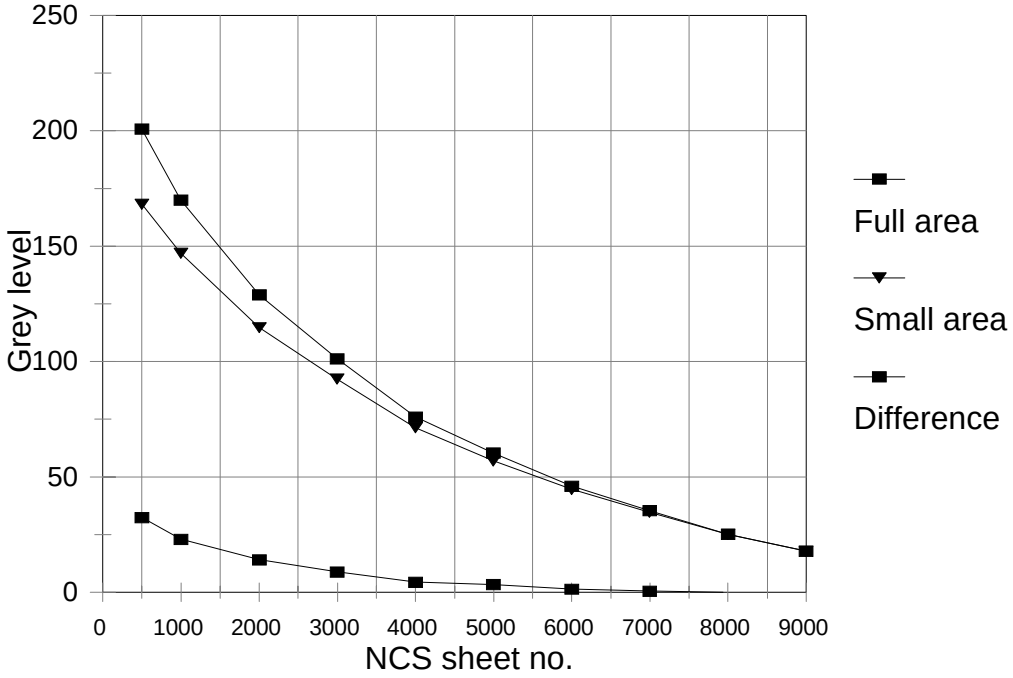


Figure 5.8 Plot of 10 grey NCS sheets. The small area is a 7 mm hole in a S0500-N sheet.

In order To check the non-linear relation between the grey NCS-sheets and the pixel values another experiment was needed. Since the integrated sphere has 4 entrance ports it is possible to reduce the full light (all four entrances) to ¾, ½ and ¼ power by shielding the entrance ports. The result of this experiment is in table 5.2.

Light Power	NCS S2000-N	NCS S8000-N
1/4	39.3479	10.6696
2/4	64.0983	14.1793
3/4	102.1629	19.4666
4/4	130.3084	23.3946

Table 5.2 Mean values of 200x200 window in the center of the image, with various of the four entrance ports shielded.

In figure 5.9 is plotted the result from table 5.2, when compared to figure 5.8 it is easily seen that the non-linear relation between the grey NCS sheets and the intensity values from the CCD-chip comes from the NCS-sheets. This means that the NCS system contains some non-linear relations to the number of photons, probably due to the same effect in the human eye. For a specific application it must therefore be considered if the measurements/calibration should use the linearity of the CCD chip (photons) or the NCS sheets.

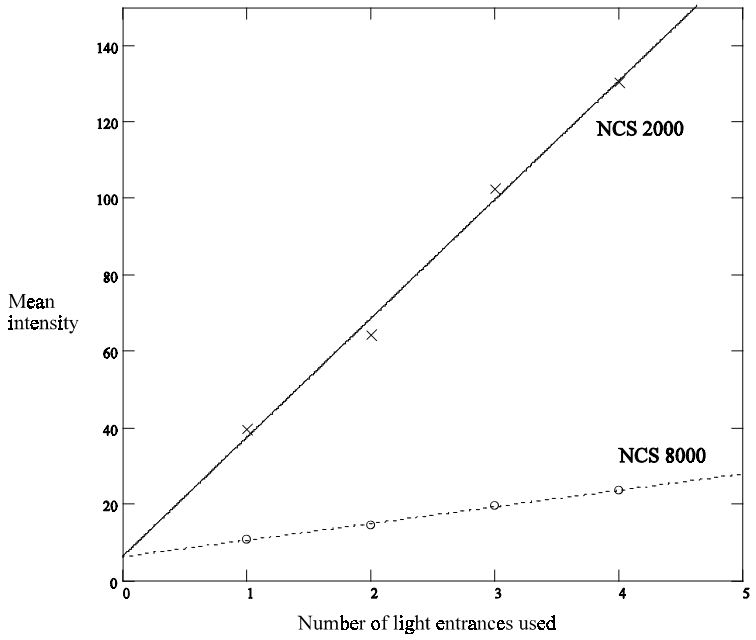


Figure 5.9 Mean values of images of NCS2000 and NCS8000 sheets with various number of light entrances turned on. The lines are the linear regressions for the four measurements.

As seen in figure 5.9 the two lines do not cross the Y-axis in zero, which shows that it is very important for the calibration for uneven illumination to make at least two measurements at different levels. The linear regression based on the four measurements gave a zero crossing of the S2000-N to 6.243 and the S8000-N to 6.062. In order to get the correct value the lens was shielded and an average darkness (black) image was captured, this gave an image with a mean value of 5.95619 with a standard deviation of 0.07699, which is very close to the extrapolated values from figure 5.9.

Depending on the settings of the camera and frame-grabber the pixel value of the darkness (black value) is not necessarily zero as shown above. The problem with this in a calibration for absolute colours and uneven intensities is that it is not granted that the darkness pixel value is negative and the method of grabbing a darkness image does not necessarily work. It was therefore decided to make the geometric calibration (aberrations etc) with the black disks on a white surface (see figure 5.3) and the calibration for colour shading, vignetting and $\cos^4$ light fall-off with multiple grey NCS sheets. The Surfix method of using the black disks and the white areas in between was too dependent on the printing device and the background material and was unsuitable for calculating absolute colours.

It was decided that the first calibration should be the calibration for uneven intensities to avoid the systematic calibration object displacements from the uneven intensities, see page 151. After the estimation of the parameters for the geometric calibration the order of calibration on subsequent measurement is of course less important, but there is a very small difference in the result if the geometric calibration is done before or after the calibration for the uneven intensities, but this is considered insignificant.

To keep the highest possible accuracy and to avoid truncation errors the Surfix software was rewritten to make all calculations and results in floating point, this means that the colour images was captured as floating point images (PAL 768x576 pixels, 4 bytes/pixel in three channels, 5.3 Mbyte per image !).

Calibration for uneven intensities

The calibration for uneven intensities was decided to be made as a calibration image generated from multiple images (more than two) of grey NCS-sheets. The calibration routines in the software asked for the correct NCS sheets and grabbed average images of the sheets. These sheets were converted into two calibration images by filtering the images with an initial 7x7 local median filter followed by two subsequent 25x25 local mean filters. This combination

removes outliers (dust, scratches etc in the surface) before the actual blurring (robust filtering). Subsequent images were then calibrated pixel by pixel using a linear stretch to fixed values between these calibration images (each grey NCS sheet was forced to have a specific value). This method was chosen since the unevenness of the images was too complex for modelling, see e.g. figure 3.15 and 3.16 in chapter 3.

In figure 5.10 is a subsequent colour measurement on the complete series of grey NCS sheets calibrated using the above describes method from the S9000-N forced to pixel value 45 and the S1000-N forced to 220. As seen from the plot of the mean values, the RGB values are very close near the two calibration values 45 and 220, but in between and outside they tend to spread out a little bit. This spread out is visible in the images as a colour shading similar appearance and are probable for the main part due to the missing correction for self-illumination. But despite these small spread-outs the result is much better than the Surfix method with use of the printed black disks and the white material.

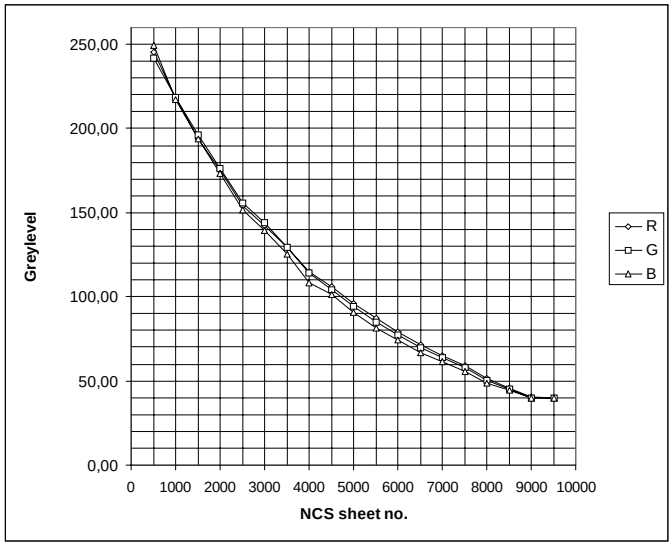


Figure 5.10 RGB values of the grey NCS sheets linear calibrated from the S1000-N and S9000-N sheets.

In order to get a better calibration using that method three sheets were used instead. In figure 5.11 is the plot similar to figure 5.10 except that the calibration is made from three NCS sheets. The S2000-N forced to pixel value 175, the S5000-N forced to pixel value 95 and the S8000-N forced to pixel value 50. As seen from the plot in figure 5.11 the result is much better with three sheets, except from the tails where a small spread out still occurs. The choice of NCS sheets to make the calibration from, should also be done from task to task to ensure a maximum accuracy inside the relevant intensities. If e.g. measuring on very bright textiles the relevant NCS sheets could be S1000-N, S3000-N and S5000-N.

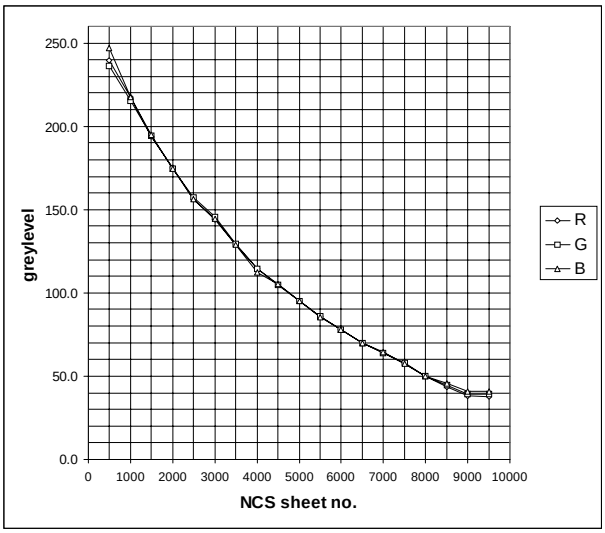


Figure 5.11 RGB values of the grey NCS sheets linear calibrated from the S2000-N, S5000-N and S9000-N sheets

The problem with the results in figure 5.10 and 5.11 is that the colour shading, vignetting and $\cos^4$ light fall-off is very difficult to separate from self-illumination in the sphere and the need for obtaining correct measurable colours in e.g. NCS values. For an accurate detection of the self-illumination it is required to have even intensities and in order to detect the uneven intensities accurately and get measurable colours the self-illumination must be known (as shown in the figures).

One solution is to have some sort of intensity calibration mark in each measurement, like a small piece of a NCS sheet inside the image plane. In this way the variations in self-illumination can be determined in each measurement. This method has been used in this project with measurement on fungus [Knudsen Jon 1998], where the background was chosen as a NCS 1000-N sheet. In each image (measurement) the channels where multiplied with a constant so that the background (the S1000-N) had the correct value. This ensured that the change in the spectral distribution of the light inside the sphere due to the coloured samples did not influence the measurements and of course that the overall intensity was correct and repeatable. This method of only calibrating from one calibration mark was much better, but not perfect if the images were analysed more intense, this was due to the lack of a zero value in a darkness image (black). For the best result from this method it is therefore required to use at least two different samples, if e.g. the background is a detectable mosaic of two different NCS sheets.

Since it is the size of the sample area (bottom cap) compared to the area of inner surface of the sphere that causes the self-illumination the use of larger sphere or a smaller sample area will also reduce the self-illumination significantly. The sphere used in this project is 360mm, giving an inside area of 407150mm^2, the cap is 120 mm in diameter giving an area of 11309mm^2, the sample surface is therefore about 2.8% of the illumination area in the sphere. If e.g. the diameter of the viewport is reduced from 120 mm to 60 mm (2827mm^2) the viewport area would be reduced by 75% and the sample surface would only be 0.7 % of the illumination area. If the sample area instead is remained at 120 mm and the size of the sphere is increased to 600 mm the inner surface of the sphere would be 1130973mm^2 and the sample area would be 1.0% of the illumination area.

It is also possible to mount a sensor inside the sphere and through some connection to the light source to control and adjust the light level to a constant value independent of the samples inside the sphere. This is though only realistic with the intensity level since it would be very difficult to accurately adjust the spectral distribution of the light. A sensor inside the sphere could also be connected to the computer and then be used as a parameter in the general calibration for colour shading, vignetting etc.

Another method is to estimate the actual self-illumination directly from the measurement (image). If the mean value of an grabbed image (after calibration) is very low in one channel it must mean that the self-illumination is very low (high degree of absorbed light) and opposite, if the mean value is very high it means a high self-illumination. The problem is to estimate the correct self-illumination value and to correct for it. The initial work on this problem has been done within this project, but due to time reasons this has not been completely finished.

For this initial self-illumination calibration there has been made three special NCS sheets with four 6 mm holes where there underneath each hole is glued a small sample of other NCS sheets. In figure 5.12 are images of these three special sheets calibrated for unevenness with the above described method. The calibration for self-illumination should be done after the calibration for unevenness because of the requirement for an even distribution in order to detect the values of the small samples correctly (this also makes an automatic detection algorithm much easier). These three sheets can also be used for the first calibration for unevenness (instead of the normal NCS sheets) by e.g. using morphology filters or large robust mean filters to separate the small sample holes out, this reduces the calibration time and the possibility of human errors.

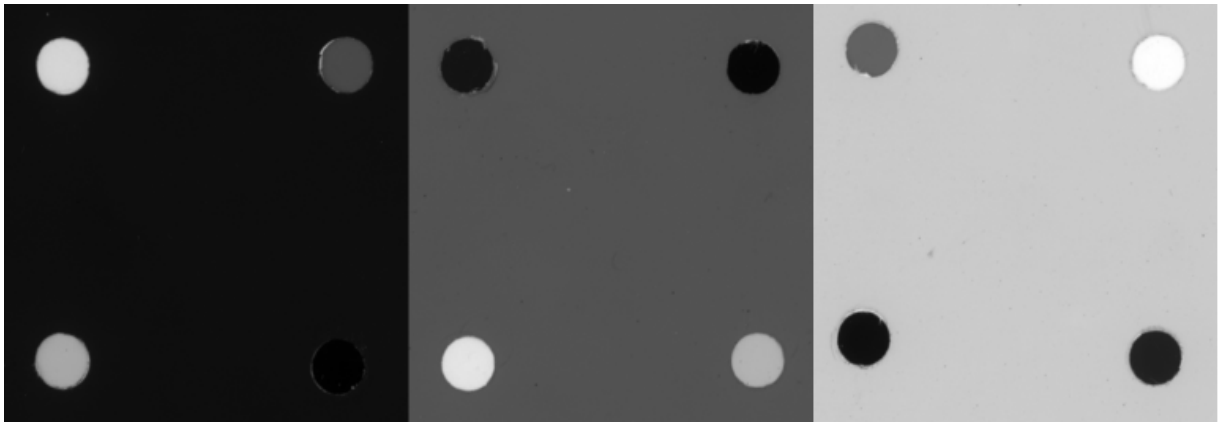


Figure 5.12 Three grey NCS sheets with four small samples of other NCS sheets. From left to right: S8000-N, S5000-N and S2000-N. The small samples are in all three S1000-N, S2000-N, S5000-N, S8000-N and S9000-N, if the sample and the sheet are equal the sample is omitted.

When having the calibrated images of the three sheets in figure 5.12, the mean intensity values of the small samples and the remaining foreground sheet are measured. In table 5.3 are these results from the images in figure 5.12, the values in italic and marked with an * are the mean value of the foreground NCS sheet. The values in the table are only the intensity values since the values in the colour channels are practically equal because of the initial calibration for unevenness.

Foreground	Small areas (through holes)				
	S1000-N	S2000-N	S5000-N	S8000-N	S9000-N
S2000-N	207.5	*173.6**	105	61.5	50.4
S5000-N	190.5	158.3	*94.7**	53.6	44.6
S8000-N	180.5	150.6	88.3	*49.9**	41.3

Table 5.3 Measurements made on NCS sheets with small samples of other sheets.

From table 5.3 it can be seen that the self-illumination with the sphere size used in this project is quite large and a calibration besides the unevenness is needed. To illustrate the measurements from table 5.3 each row is plotted in figure 5.13, as it can be seen the local intensity values are either increased or decreased depending on the global value of the image (the foreground NCS sheet).

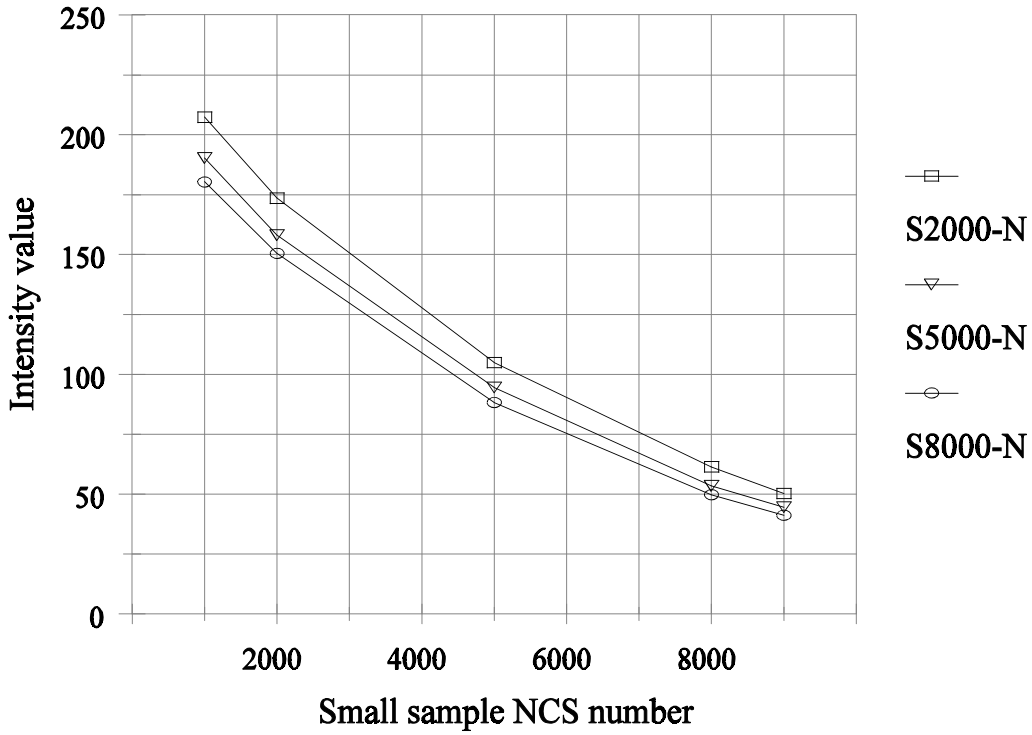


Figure 5.13 Plot of results from table 5.3 for each foreground sheet.

The non linear curves in figure 5.13 comes from the non linear relation between the CCD intensities and the grey NCS sheets. To illustrate the relations between the global and local values, in a calibration for self-illumination using an integrated sphere, the measurements from table 5.3 are replotted in figure 5.14 with the measured mean intensity value of the total image (global intensities) versus the local intensity values (from the small samples of NCS sheets), for each of the small samples of NCS sheet.

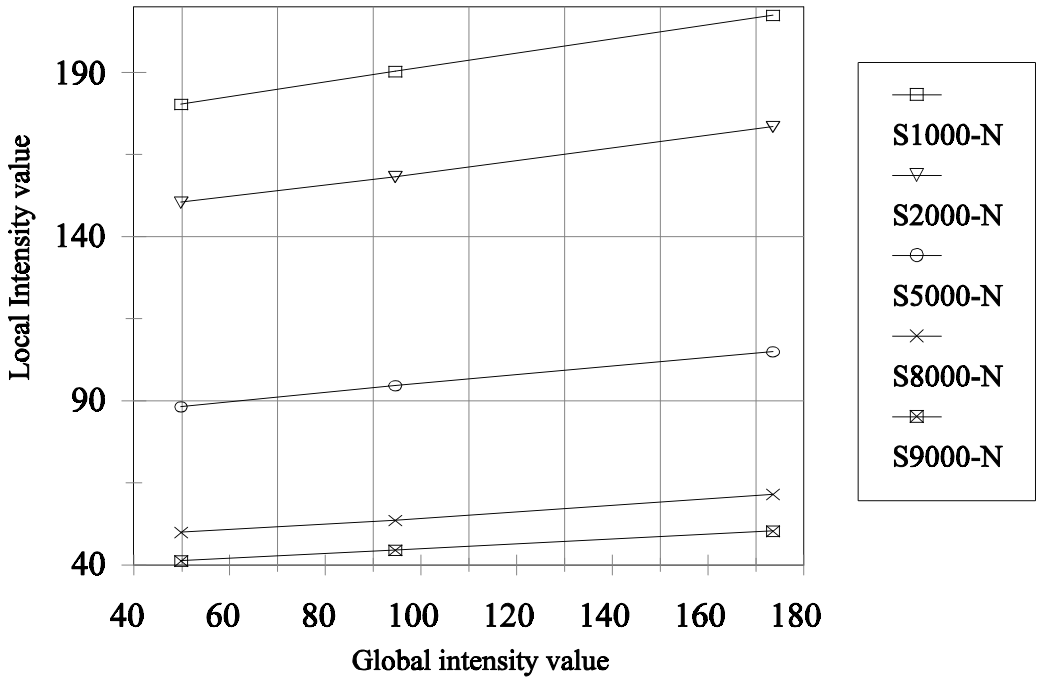


Figure 5.14 Plot of measurements from table 5.3 showing the linear relation between the local intensity values and the global intensity value.

This shows that given the global intensity level it is possible to do a linear transform of the local intensity values in order to reach some artificial local values that are independent of the global intensity level. It should also be mentioned that such a calibration should be done on each channel separately. Unfortunately there has not been time for further investigations of these aspect during this project and it is the author's opinion that this is one of the most important tasks of the future development of the devices for colour texture measurements.

Geometric calibration

Since the integrated sphere as it was designed in this project uses a camera with a parallel image and object plane and the depth of field is very low (close range 400mm between lens and object) the geometric calibration is done using warping, like with the Surfix device. The calibration is done based on a regular grid of black disks on a white background, see figure 5.3. The needed interpolation is done using a bilinear interpolation to avoid introduction of new pixel values above or below the original. All results are made in floating point to avoid truncation errors.

To avoid systematic positioning displacements caused by uneven intensities the above described method for calibrating for uneven intensities is done prior to the automatic positioning of calibration objects.

Reflections

A glossy plastic cup for growing fungus placed inside the sphere [Knudsen Jon 98] gave a minor reflection of the 21 mm exit port in the top of the sphere. This is a serious problem for accurate colour measurements on samples and one method to eliminate this is by using non-ortho images in the sphere making a system like shown in figure 5.15. Depending on the angles etc this could require some rearrangement of the light sources or by placing some extra shields inside the sphere. Using the warping method for the geometric calibration it is possible to warp/resample the images to be an estimate of parallel image plane and object plane. If needed the polynomial order of the warping should be increased to be able to make a better fit to both the lens distortions and the perspective transform of the image.

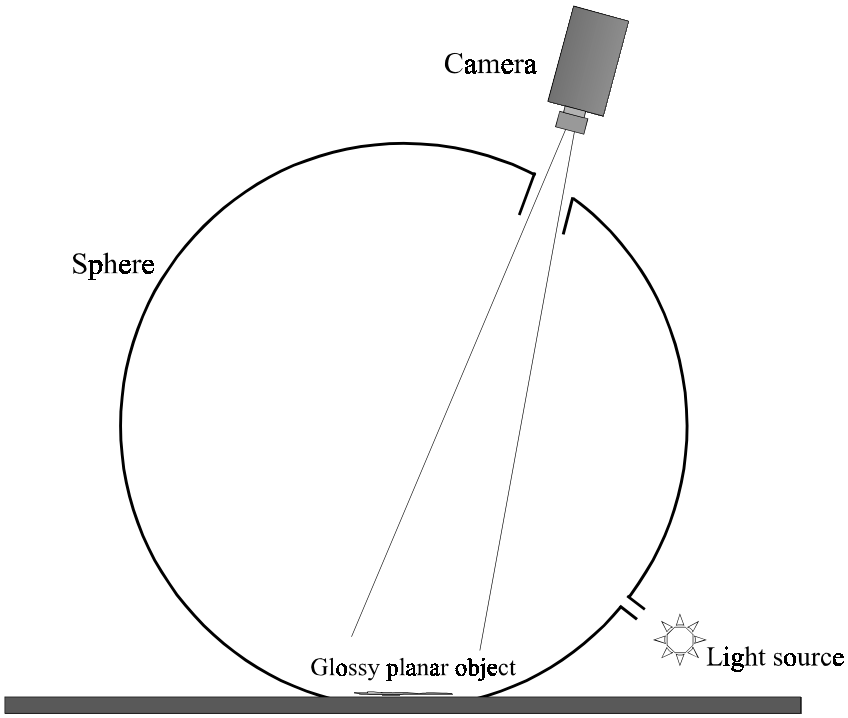


Figure 5.15 Integrated sphere for glossy and planar objects, that can avoid reflections.

The practical work with these highly calibrated and accurate systems has shown some small low frequency time varying intensities and colour distributions. These are probable coming from the temperature of the light source, since the opening of window in the room has been visible in measurements. This means that either some sort of online or between measurements calibration adjustment is needed to obtain very accurate and repeatable measurements.

VMX2000

To avoid foreign light (room light etc) to enter the sphere through the viewport in the top, the author of this project designed and created a big housing in glass fibre and PVC foam for the sphere, camera and linear up/down movement (see figure 5.6). The housing was painted black inside and was covering the upper part of the sphere. A large maintenance port was placed in the front for an easy

access to the camera and lens. The complete system with sphere/housing was called a VMX2000 device and is shown in figure 5.16. The up/down movement of the sphere was motor driven and completely controlled from the computer.

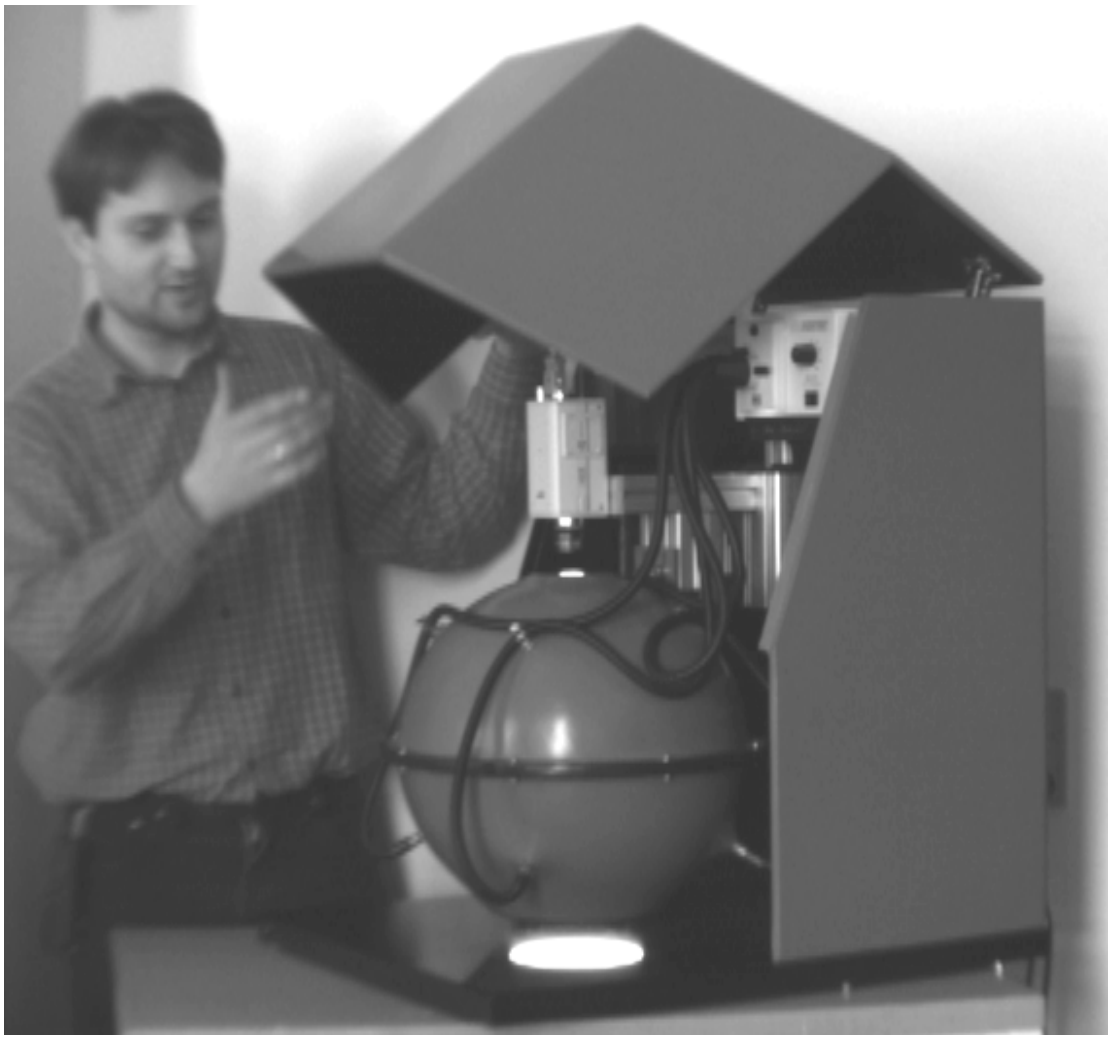


Figure 5.16 Jørgen Folm-Hansen demonstrating the VMX2000 prototype.

The VMX2000 device, the calibration methods and measurements on selected samples has been presented at *Visionday at the Technical University of Denmark* 6th of June 1998, as a poster at the DSAGM conference in August 1998, a firm seminar at Novo Nordisk A/S and for various industrial and academical visitors at the Departments during 1998.

5.1.3 The SurfView

In late 1998 the VMX2000 device was given to a Danish company called *DELTA Lys og Optik* (DELTA - Danish Electronics, Light & Acoustics) that already works with/design various devices for measuring light and colours, including integrated spheres but without the use of cameras. DELTA redesigned the VMX2000 device with new materials from their knowledge of illumination and must important to reduce the production costs. The carbon fibers in the sphere was replaced with bended aluminum and the glass fiber/PVC foam housing was replaced with sheet metal. To reduce the size of the device the camera is tilted 90° looking through a mirror and the light sources is mounted directly on the aluminum sphere with small shields below to remove direct light hitting the samples in the bottom, see figure 5.17.

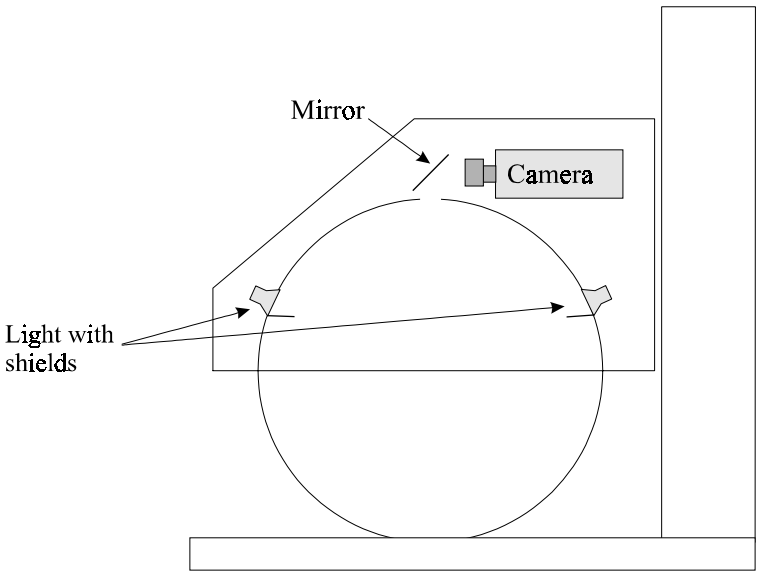


Figure 5.17 The principle of the SurfView device.

The software and calibration routines are still the same (at the time of writing this thesis) as with the VMX2000 device, but will be altered to a more robust and user-friendly software before the commercial distribution. In figure 5.18 is a drawing (no photo at the present time) of the SurfView device that looks almost exactly as the first prototype of the new SurfView device.

Figure 5.18 Drawing of the VMX2000 replacement called SurfView designed by DELTA lys & Optik in Denmark.

5.1.4 Measurements

The colour texture devices Surfix and VMX2000 has been used for measurements on various materials (textiles, fur, fungi, Rockwood etc) during this project. In this thesis only the two most important and some aspects about the colour texture measurements be mentioned. The use of an integrated sphere has also shown very good results with very textured surfaces like furs because of the very diffuse light.

Textiles

Measurements on textiles was one of the first to be done in this project, this work was done in cooperation with several people at Novo Nordisk A/S (among them Morten Rank and Tina Jacobsen).

When working with stonewash of the textile denim (e.g. used for blue jeans) that consists of the two different coloured threads blue (indigo) and white, it is unwanted that the blue treads have a rub off effect on the white treads. In figure 5.19 are shown both the front side (a) and the back side (b) of a small piece of denim. Novo Nordisk A/S among others have evaluated the rub off on the back side of the denim (called backstain and seen on the back side (b)) using a group of people to visually give a scoring (PSU - Panel Score Units).

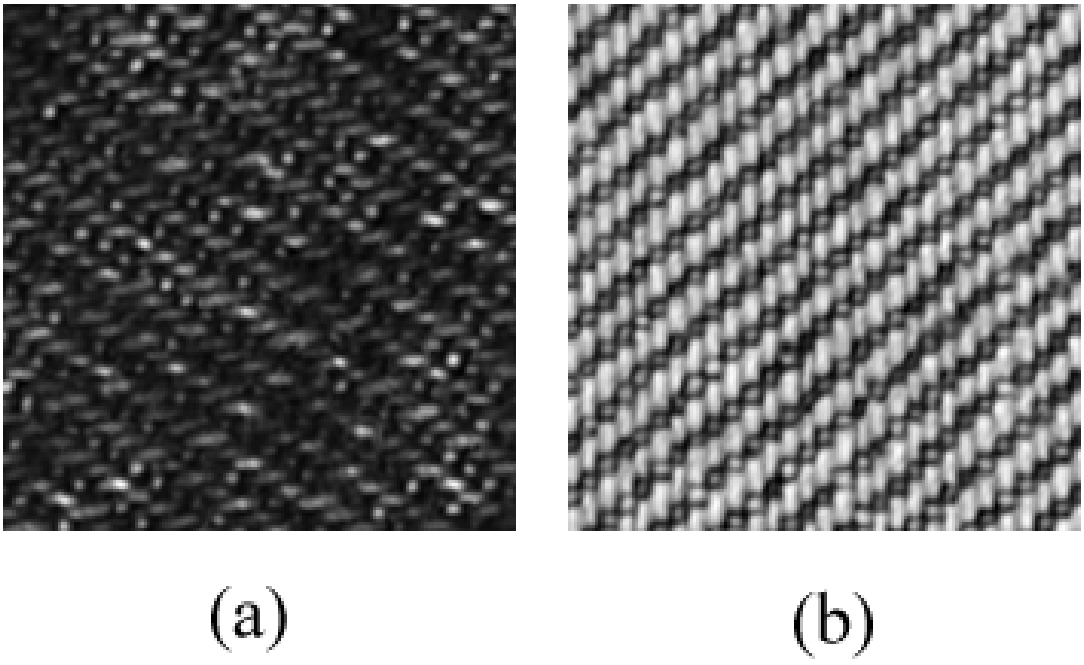


Figure 5.19 Small area of denim from VMX2000. (a) front side, abrasion. (b) back side, backstain.

In this project the VMX2000 has been used to obtain highly calibrated RGB images of the denim and from these images to measure the blueness of the automatic detected white treads. The white treads was detected simply by only looking at the most bright pixels (a fixed percentage) measured in the intensity (mean value of red, green and blue channel), a manual inspection of the selected areas has shown that this method is sufficient. In figure 5.20 is the plot of the VMX2000 backstain measurement, the PSU and the reflectometer at 420nm versus the enzyme dosage for a specific stonewash at Novo Nordisk A/S. The reflectometer can be used for the backstain measurement but since it measures on both the white and blue treads it is not a good measurement, but is included for the comparison [Carstensen99]. The correlation between the VMX2000 and the PSU was considered *excellent* by Morten Rank at Novo Nordisk A/S.

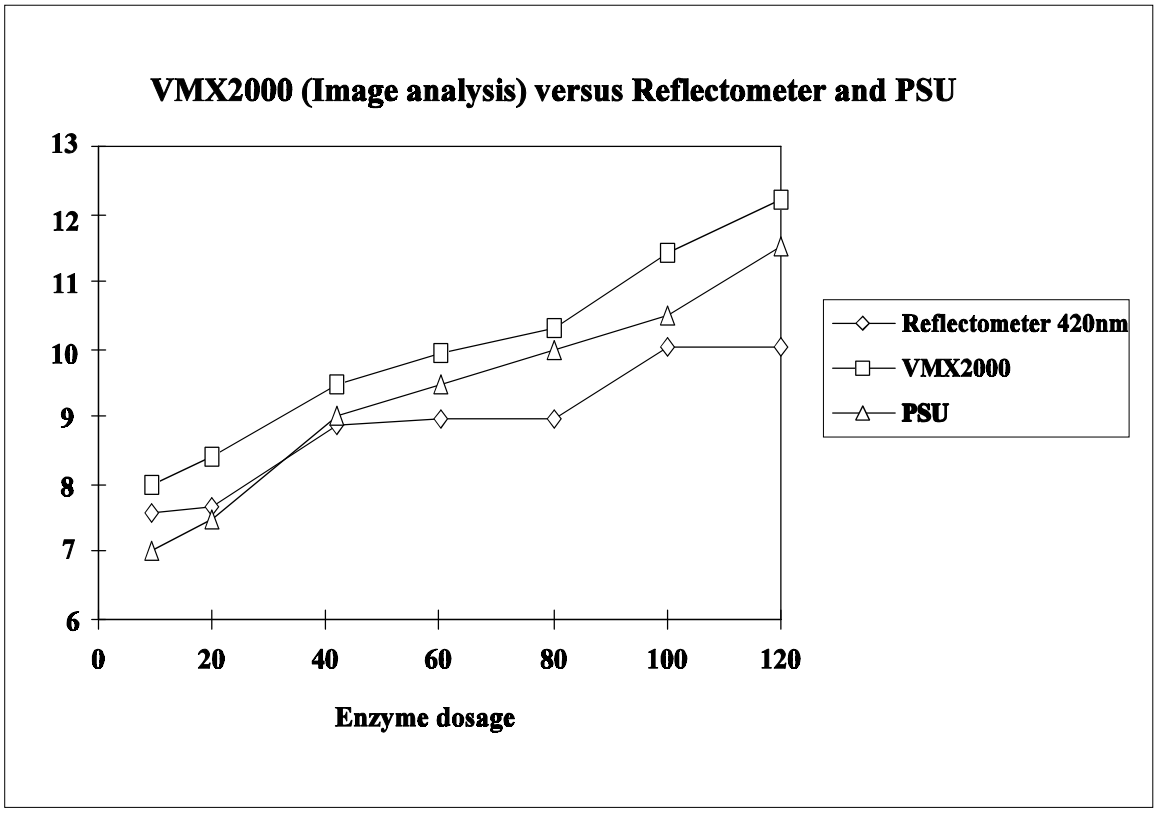


Figure 5.20 Measurements of the backstain using the reflectometer, VMX2000 and PSU. Notice the good correlation between the VMX2000 and the PSU.

Another problem with washing the blue denim is the amount of remaining blue colour (indigo) afterwards, this measurement is done on the front side of the denim (called abrasion) and is currently done using a PSU and/or a reflectometer at 420nm. The VMX2000 abrasion measurement is simply the average of the blue channel. In figure 5.21 is the plot of the VMX2000 abrasion measurement and the reflectometer at 420nm versus the incubation time for a specific low enzyme dosage wash at Novo Nordisk A/S. The VMX2000 abrasion measurement was considered very well by Morten Rank at Novo Nordisk A/S.

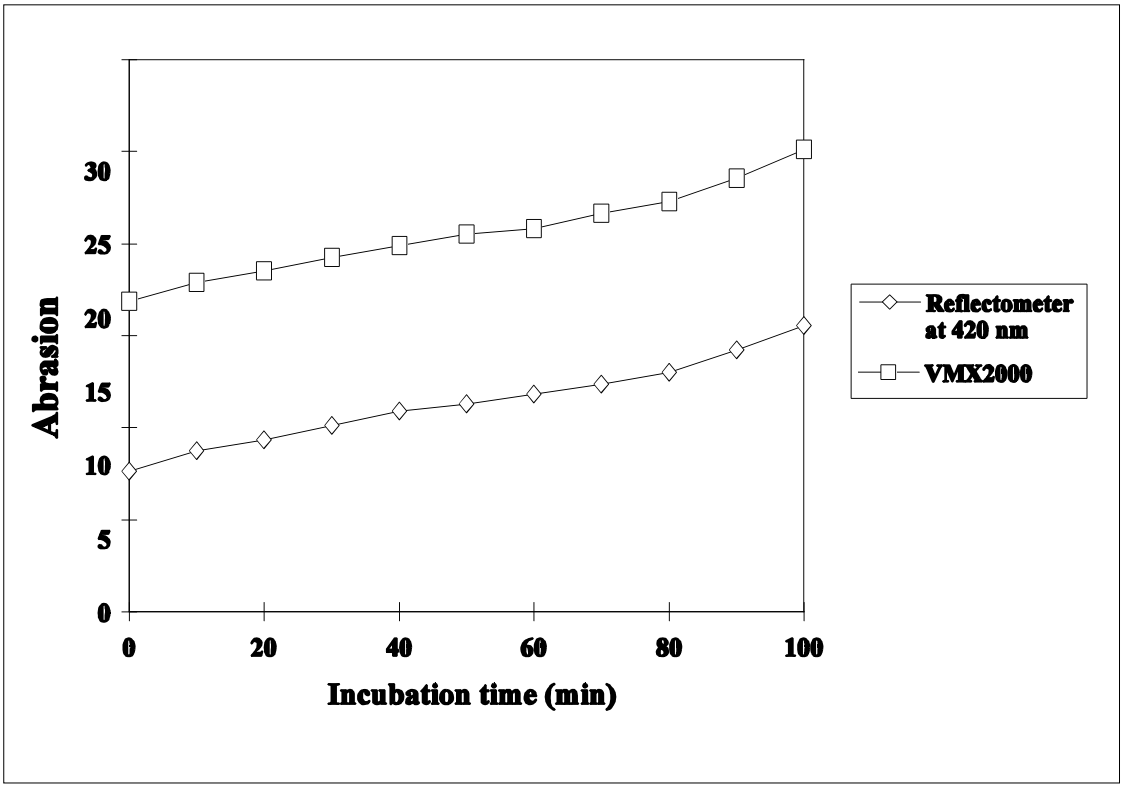


Figure 5.21 Measurement of abrasion on denim using a reflectometer and the VMX2000 versus the incubation time of a specific enzyme wash. Notice the very good correlation between the two devices.

The conclusion from Novo Nordisk was that the VMX2000 performed very well and that the repeatability was as low as 0.1% when making multiple measurements on the same place, 1.8% at different places on the same piece of denim and that the day to day variation was 2.5%. See [Carstensen99] for further details.

The Surfix and VMX2000 has also been used on other type of textile and types of measurement, where the general problem for e.g. Novo Nordisk A/S has been the lack of devices to do the measurements (currently only PSU). There has been made measurement on black sock before and after a number of machine washes with a specific washing enzyme to measure the amount of fluff (Tina Jacobsen at Novo Nordisk A/S), this has been done with success simply by using the mean value of the intensity image. Also washing powders ability to remove fat from textiles has been initial tested with some success, where the difficult part has been an automatic detection of the areas of the textiles with fat and those without.

It is the author's opinion that the use of a highly geometrically and chromatic calibrated colour camera device has some very high usability when measuring at textiles, especially because of the high accuracy, the high repeatability (calibrated against grey NCS sheets) and the very even illumination/intensities that makes it possible to measure correctly on different areas of the same textile.

Fungus

Another project that the VMX2000 has been successfully used for was a master thesis cooperation with Jon Knudsen at the Department of Biotechnology at the Technical University of Denmark [Knudsen Jon 1998]. This project was about measurement and classification of fungus from various sources like foods (bread) and building materials (wood). The main problem is the very small differences between the fungus and that it requires highly trained people to visually classify them today.

Here a series of plastic cup where fungus has been grown was measured in the VMX2000 device. In this case the most important was to make repeatable and accurate colour measurements of various manually selected areas of the fungus. Because of the big variation in colour and size of the fungus there was a serious need to calibrate for the self-illumination. This was done by placing a NCS S1000-N sheet below each plastic cup that also was visible in the images. By simply stretching the images to have a specific pixel RGB value in the visible NCS sheet, most of the self-illumination was calibrated for. There was unfortunately not time for a better calibration. In figure 5.22 are an example of one of these fungus, one of the interesting things are the colour of the dark (green) areas.

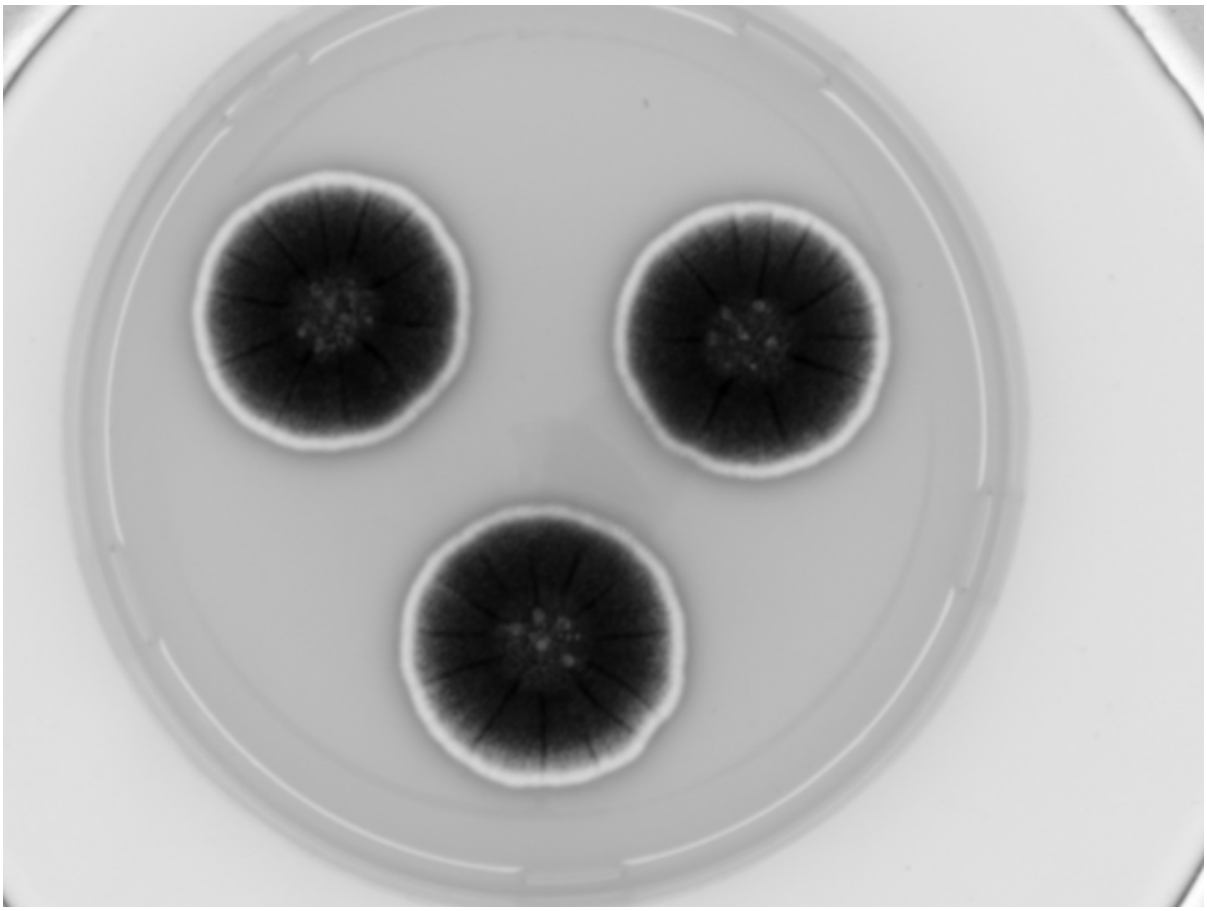


Figure 5.22 Image of fungus (from bread) captured with the VMX2000.

The conclusion from the thesis [Knudsen Jon 98] was that the VMX2000 was indeed very useful for such measurements and further research are in progress at the time of writing.

5.2 Laser Sheet Measurements

This section describes the work done on two smaller project in cooperation with the Department of Fluid Mechanics (AFM) and Dantec Measurement Technology A/S. Both projects are about the use of laser sheets but involves different types of calibrations.

5.2.1 Measuring Concentration Distribution

During 1998 a study on the concentration distribution of injected gas in a swirling bulk flow was done in cooperation with Simon Matlok at the Department of Fluid Mechanics. The contribution from this project to this study was the planning and software design of both the hardware and software to make the necessary image analysis. In figure 5.23 are a sketch of the setup placed at AFM, from [Matlok98], the laser sheet is placed in the mixing region between B and C, and the monochrome camera is placed at the end of the square tunnel looking at the laser sheet through the heating surface (C).

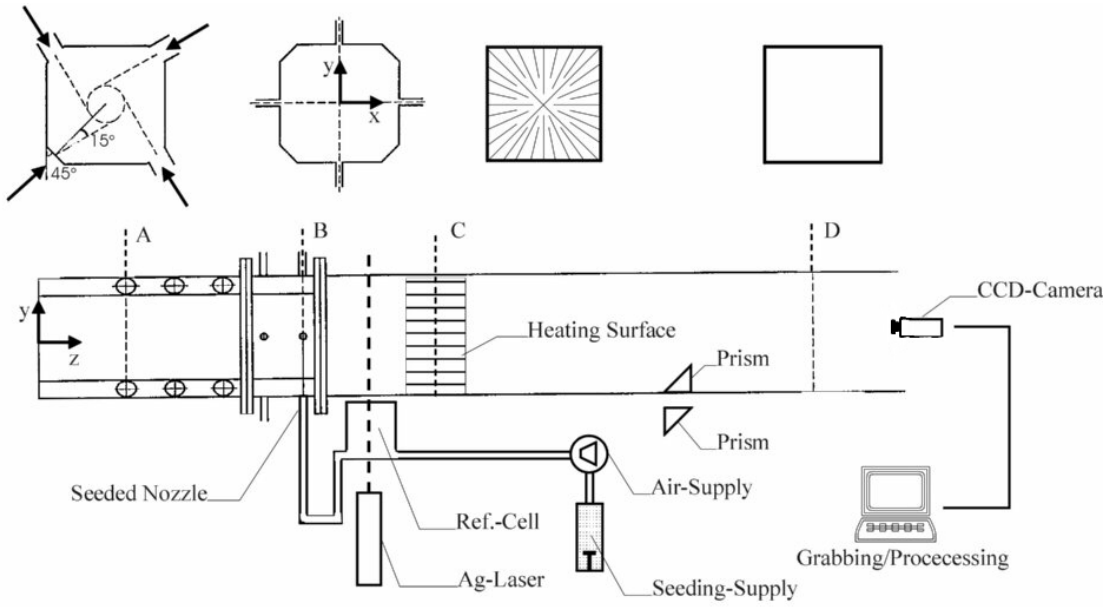


Figure 5.23 Sketch of the 1:60 scale model of a 204x204cm cross-section furnace.

The setup is a 1:60 scale model of real 415MW tangentially fired boiler, where the heating surface is modelled by the radial plates placed at C in figure 5.23. Since the heating surface is placed between the laser sheet and the camera the image calibrations is both the artificial removal of the heating surface from the images and the calibration for the uneven distribution of the laser sheet. In figure 5.24 are the first image to be captured when starting a measurement, this is the average image of about 100 frames where the gas was made almost evenly distributed. From this image both the unevenness of the laser light and the position of the heating surface are estimated.

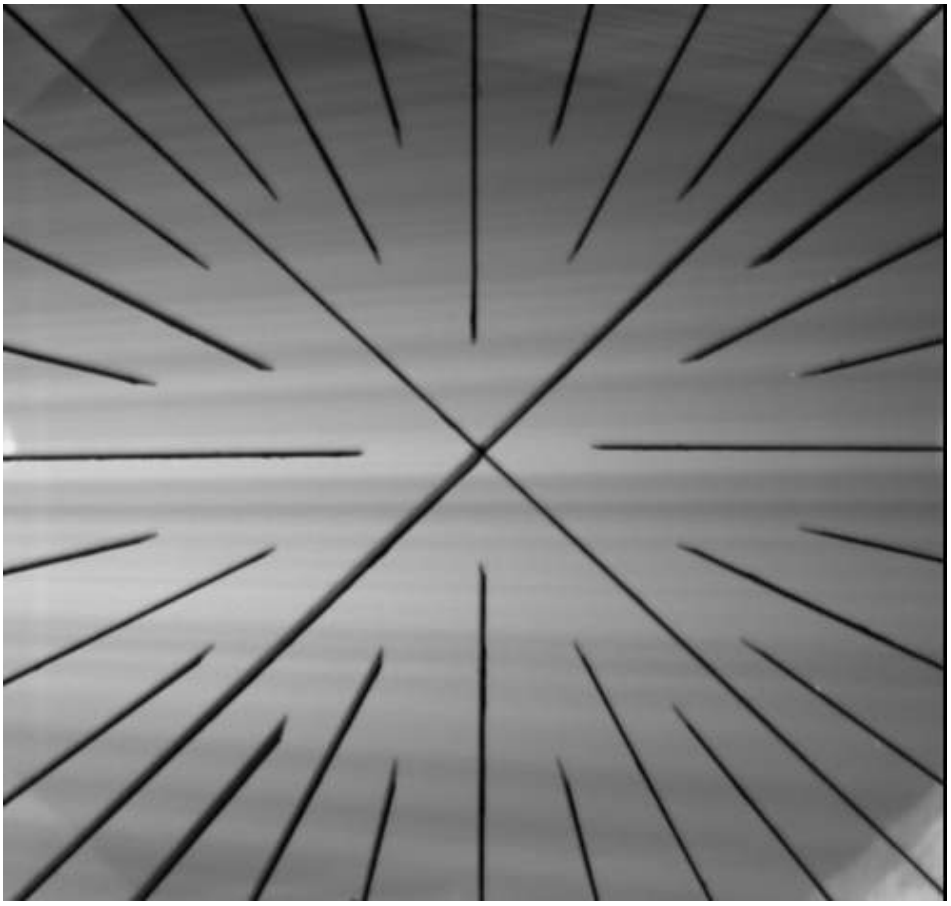


Figure 5.24 The laser sheet with an almost uniform gas distribution and the heating surface in front. Used for automatic detection of the heating surface and for estimation of the laser sheet intensity distribution.

The detection of the laser sheet cannot be done by a simple threshold since the laser light illumination is very uneven. So the detection of the heating surface and the unevenness of the laser sheet must be done simultaneous. First the image is heavily blurred (2 times a 25x25 local mean filter) to get a very rough estimate of the laser sheet. This is then subtracted from the original image to get a background corrected image that can be used to get the initial estimate of the heating surface. The initial heating surface is then detected as everything 10 % below and 40% above of the mean value of the image. This initial guess is then slightly eroded to make sure that all of the heating surface is selected.

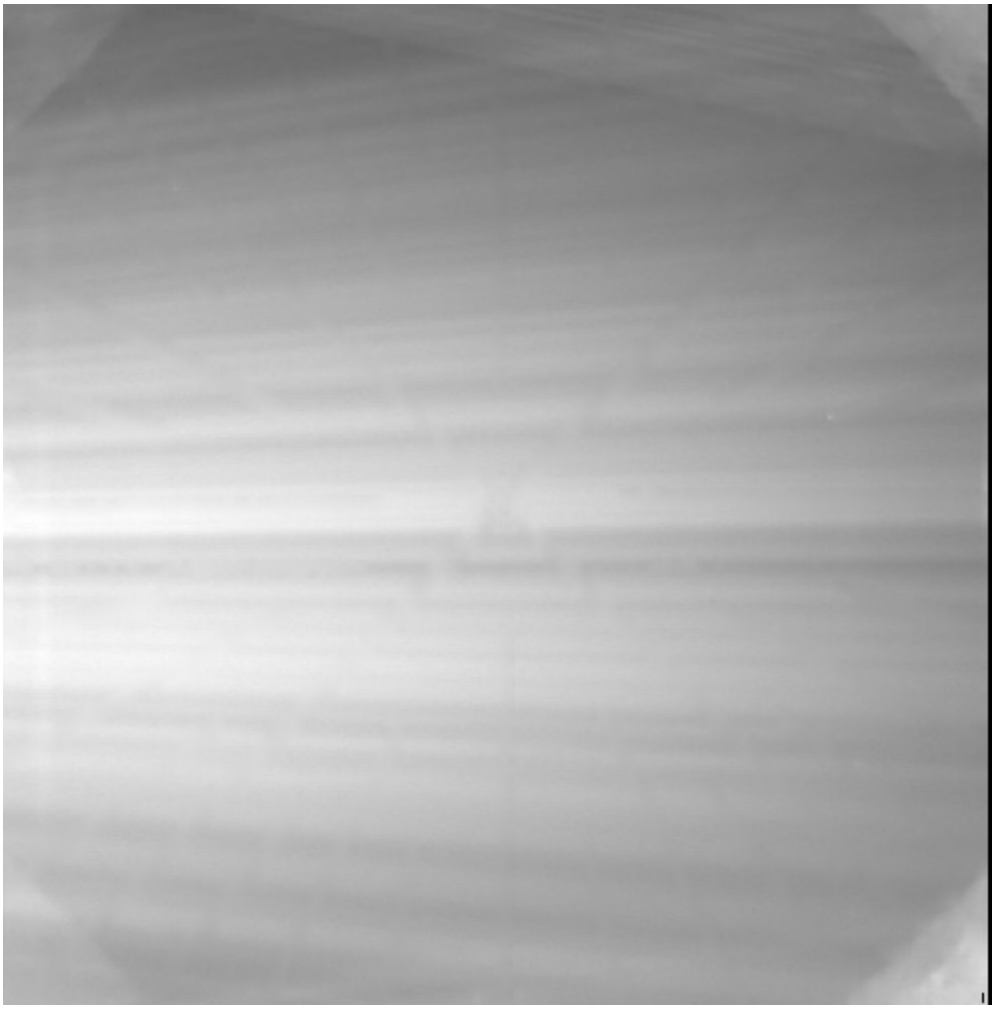


Figure 5.25 The estimate of the uneven distribution without the heating surface (interpolated away).

By making a simple interpolation over the pixels included in the initial detected heating surface on the original image, a new image where the heating surface has been removed is generated, see figure 5.25, this is used as the non-blurred laser light intensities image. This new image is then heavily blurred (2 times a 25x25 local mean filter) to get a blurred detection of the laser light intensities. By dividing the original image with the blurred detection of the laser light intensities, the heating surface can then be finally detected as the pixels where the value is below 0.88 or above 1.4. This method has been successfully used at AFM with only a few parameter adjustments in some cases.

After these automatic detections of the heating surface and the unevenness of the laser intensity, the real measurements on the gas mixing can be done. This was done by starting the system and then capture about 75 (PAL) frames at 25 frames/sec.

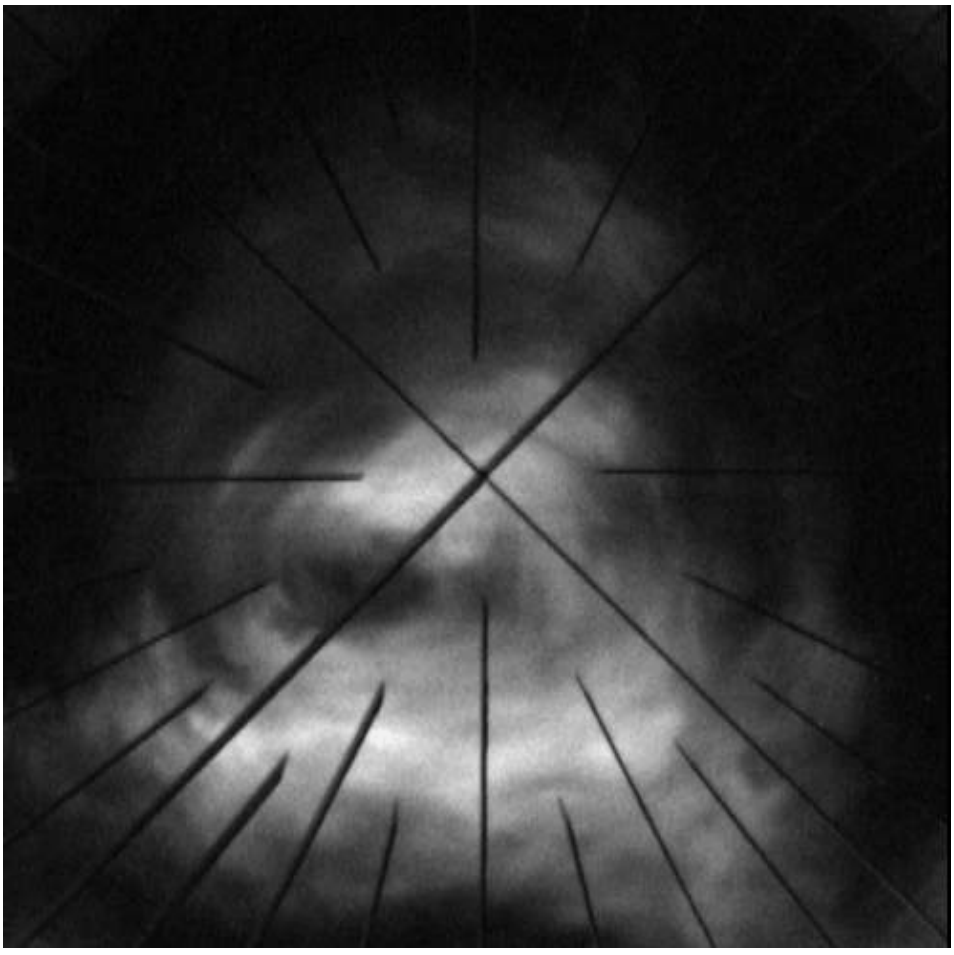


Figure 5.26 A single frame (interpolated field) from a sequence.

Because these frames has movements up to 6m/s the interlacing of the camera had to be removed by converting the 75 interlaced frames into a 150 fields sequence, that was converted to a full frames sequence by replacing the missing lines with the mean value of the pixels above and the below. In figure 5.26 are a single frame (interpolated field) from such a sequence. In order to make any measurements on these frames they had to be corrected for the unevenness of the laser light intensities and the heating surface had to be removed/ignored.

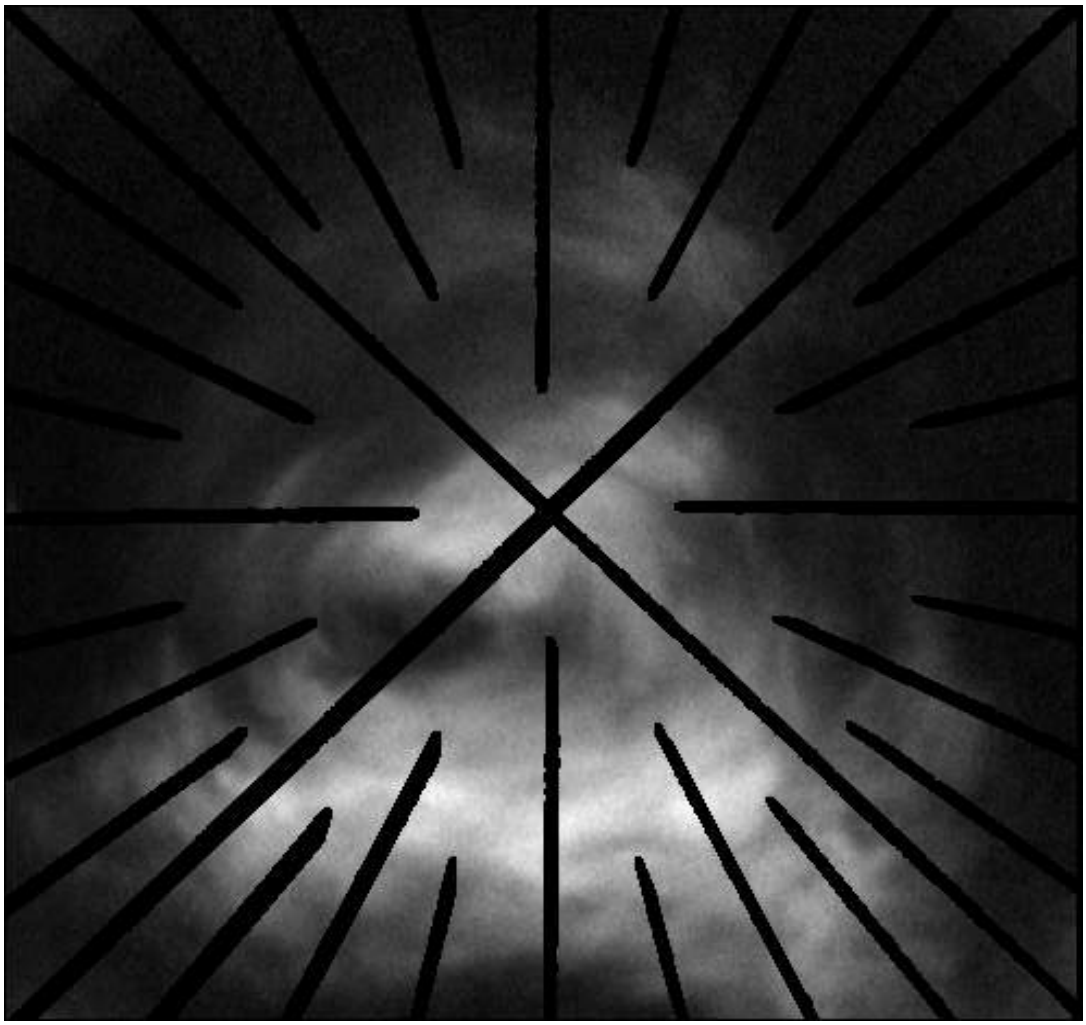


Figure 5.27 A frame from a sequence with the heating surface marked.

In figure 5.27 are the frame from figure 5.26 corrected for the unevenness of the laser sheet intensities and with the heating surface detected from figure 5.25 marked. The unevenness of the laser light was corrected by generating a calibration image from figure 5.25 that when multiplied together gave a uniform image with a value of 80. Each frame in a measurement series was then first multiplied with this calibration image. The heating surface was removed by an interpolation over the marked heating surface in each frame. In figure 5.28 are the frame from figure 5.24 corrected for the unevenness of the laser sheet and with the heating surface interpolated way.

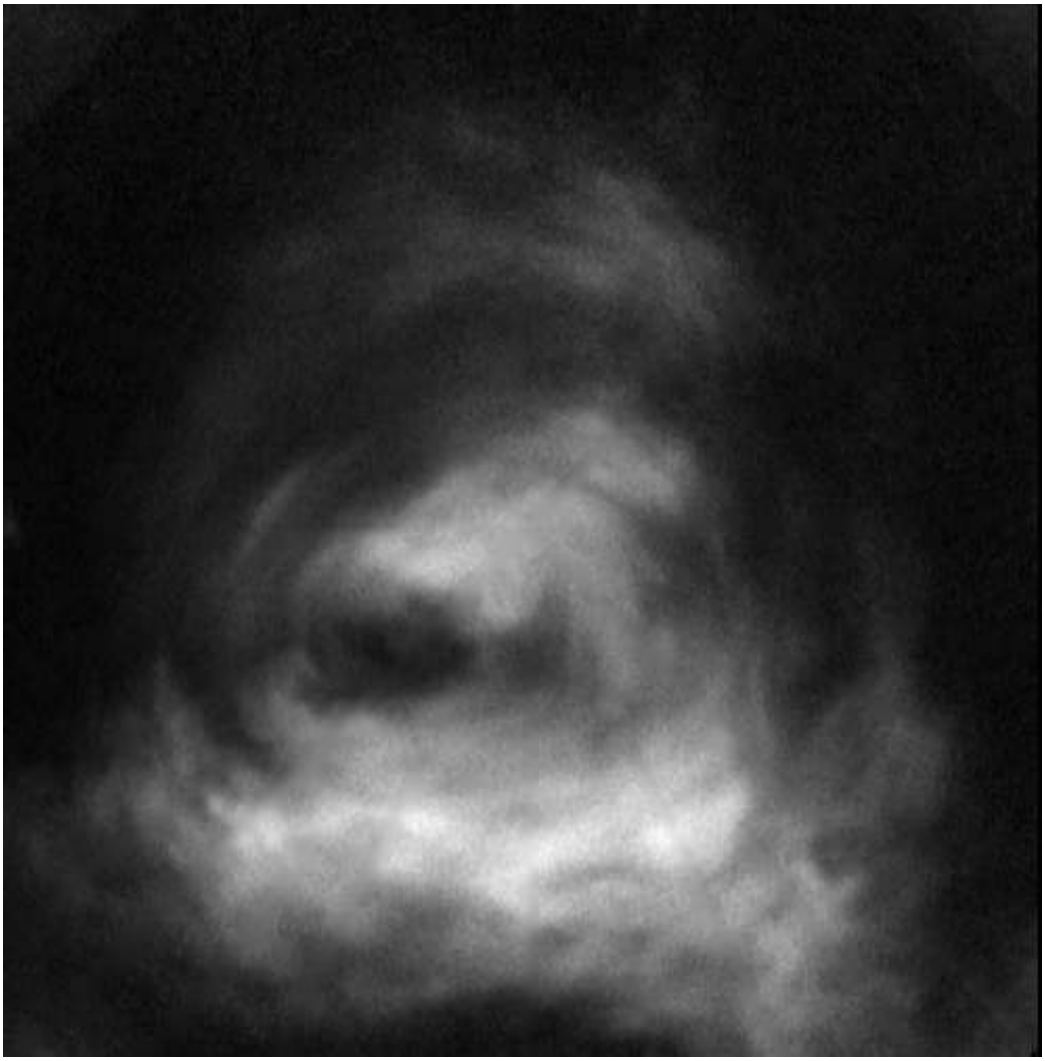


Figure 5.28 A frame from a sequence corrected for the unevenness of the laser sheet intensities and with the heating surface interpolated away.

The above described methods and the communication with the image capturing hardware was implemented in a Microsoft Windows 32 bit program with a graphical interface and a complete help system. By the use of this program a series of measurements on the setup in figure 5.23 was made by Simon Matlok at AFM. The result from these measurement and the use of image analysis are described in more details in [Matlok98], that was presented at the 8th international symposium on flow visualization in 1998. The conclusions on these measurements are that the calibrated pixels values from different experiments could be interpreted as normalized concentrations, to an accuracy within about 2-3%.

5.2.2 Particle Image Velocimetry in 3D

In the summer 1998 Finn Mengel from Dantec Measurement Technology A/S contacted the Image Analysis Group at IMM about a cooperation with a study on the possibility to make a 3D Particle Image Velocimetry (PIV) device.

To explain what a 3D PIV system is, it is best to explain the concept of the 2D PIV first. A PIV system is used for measuring a flow map by suspending seed particles in the fluid. The illumination is done by a laser sheet that is pulsed twice at a known interval (t) and an image is grabbed for each pulse. By dividing the two images into small areas called interrogation regions, the velocity vector map of the flow field are estimated. This is based on the displacement of groups of particles between the two images, the displacement is measured using a FFT based correlation technique. The result of a 2D PIV system is therefore only the movements in the laser sheet plane. The idea with 3D PIV is to make a real 3D velocity vector map (includes movements in and out of the laser plane). This can be done by using that the laser sheet has a given thickness (< 1mm) combined with 2 cameras in a stereo setup. In figure 5.29 are a sketch of the 3D PIV setup at Dantec M.T. Since the cameras are placed quite close to the laser sheet in a 45° angle (low depth of field) there is a fine control of the angle between the camera and camera lens so the Scheimpflug condition can be satisfied (the whole image plane in focus). For more information on 3D PIV see [Lawson97].

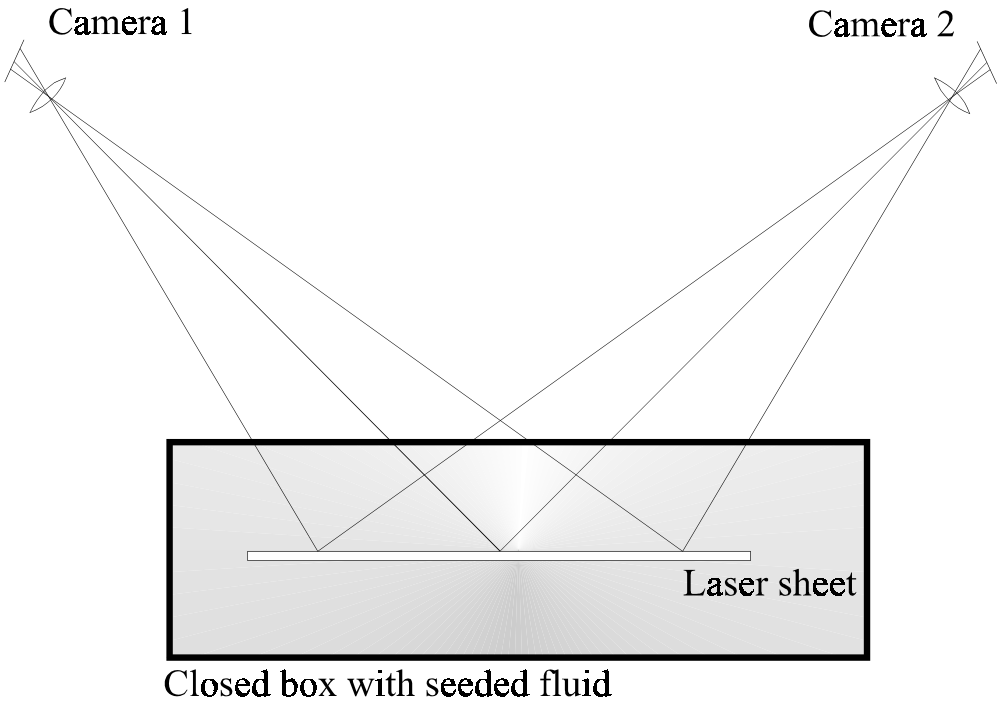


Figure 5.29 3D PIV setup, laser sheet inside box with seeded fluid and two cameras Scheimpflug tilted to give full image plane sharpness.

The contribution from this project to this study was helping with the needed calibration of the cameras both for the alignment of the images from each camera and for a DLT based positioning of the cameras.

The most important task was to get totally aligned images from the two cameras so that the PIV measurement from each camera could be accurately compared and the 3D vector could be estimated. For this purpose a special calibration target with a coded center object was designed, see figure in appendix C. In appendix C are shown an example of the automatic detection of the objects in the calibration target, because of the coding it was no problem to align the images from the two cameras. The planar calibration target was glued to a metal plate, placed approximately in the laser sheet. In order to get calibration points enough for a planar DLT (see chapter 2) the planar calibration plate holding the target could be moved forward and backward (in and out of the laser sheet) using a precise adjustment. But due to the lack of depth of field it was only possible to move it about ±3mm, the initial experiment with calculating the

DLT parameters was unfortunately not to good, but further research will be done in this area.

In the work done within this project, software for automatic detection of the calibration object, for calculating the parameters in the DLT and for warping/interpolating the two images into a common Cartesian coordinate system was made and handed over to Dantec Measurement Technology for further development towards a commercial 3D PIV system. Even though the positioning of the objects in the calibration target did not take the PPD into consideration the positioning of the objects was made within about 1 pixels.

The results of this smaller cooperation was a rough prototype of a 3D PIV system based on the warping into a common cartesian coordinate system that Dantec M. T. already has presented at one conference in Lisboa Portugal and various industrial partners. The system was able to automatically detect a coded calibration target from each camera at then to match subsequent images of seeded laser sheets, so that the 3D velocity vector could be estimated.

5.3 Summary

This chapter has described some of the implementations of the calibration methods described in the previous chapters. The main implementation has been on the development of a device for accurate colour texture measurement (VMX2000), that has been giving very good and previous undiscovered results in many areas of colour and monochrome texture measurements. Two smaller project involving calibration of cameras for flow measurements through laser sheets has also been described.

Chapter 6

Conclusion

It is an almost non-questionable fact that camera/image calibration is a very, perhaps the most, important part of many advanced image analysis tasks. It is also evident that many existing applications easily can do very precise geometrical measurements of various objects using the classic camera calibration methods from photogrammetry. The present author has only come across a sparse amount of literature on many of the subjects mentioned here. When using colour cameras for accurate colour measurement on various coloured samples there is practically no previous work done. The methods described are therefore a product of the author's expertise and experience in the area. Despite careful searches of the literature it cannot be excluded that similar or related results may have been published elsewhere. The use of an optically closed scene in form of an integrated sphere combined with a geometrical and chromatic calibration to capture accurate and repeatable colour images has been one of the big contributions from this thesis.

The studies made on calibrating colour images have also given a good insight in the calibration for even intensities using monochrome cameras. It is the author's opinion that there exists a gap between optics and image analysis, that, when bridged, opens for new and much more accurate measurements. Chapter 2 and partly 3 in this thesis are an attempt to build such a bridge.

The most noteworthy parts of the work presented in the thesis could be summarized as follows:

- A description, of the linear and non-linear camera calibration models combined with a description of the monochromatic lens aberrations, directed against the use in image analysis. The combination of the optics with the image analysis can bridge the gap between these two sciences, which, hopefully, can extend the area of cameras used as accurate measuring devices.

- A detailed description of the lens factors that influence the intensities in an image, based on the optical knowledge of vignetting and the fourth power of cosine light falloff. This can be a great help when working with calibration for uneven intensities.

- A significant amount of work done with the calibration of beam-splitter based cameras (3xCCD). The combination of vignetting, 4^{th} power of cosine light falloff and colour-shading have been described and various calibration methods have been proposed. This knowledge is very important when deciding on a calibration method when using e.g. 3xCCD cameras in accurate measurements.

- The initial work with optically closed scenes, like the cylindrically shaped first version of the Surfix device, has opened the eyes for the integrated sphere in digital image analysis for making accurate/repeatable colour/grey level measurement devices.

- The relation between using a warping and a non-linear camera calibration has been shown. It has also been shown that the use of a 3rd order polynomial warp to model the non-linear distortions is adequate when using parallel image plane and object plane. The use of warping can be necessary if the “depth of field” is too low to get non-planar calibration data.

- There has been given a detailed description of resampling and interpolation methods to be used for resampling an image, e.g. based on a calibration. This knowledge is useful when deciding on e.g. an interpolation method before further image analysis.

- The functionality of and suitable calibration methods for various colour cameras (single chip, beam-splitter and multiple shot cameras) have been presented. It is important to take the type of camera into consideration before deciding for a calibration method.

- Various methods for calibrating for uneven illumination of objects in both open and closed scenes have been presented. It has been shown that in many cases it is possible to calibrate for uneven intensities. With closed scenes the effect of self-illumination has been shown and initial attempts to calibrate for self-illumination have been done successfully. However, this requires more work.

- A detailed description of many various calibration targets and how to make automatic detections of these have been presented . It has been shown that, in most cases, it is possible to get accurate subpixel detections from these methods.

- It has been shown that uneven intensities in a blurred image can introduce a systematic displacement of calibration objects. It is therefore required to calibrate for uneven intensities before the geometrical calibration based on the correct detection of the objects.

- It has been shown that uneven blurring of an image, which is quite common in some lens/camera setups, can cause a systematic displacement of calibration objects.

- A useful model for evaluation of the perspective projection displacement (PPD) in a given setup has been presented.

The work done with the accurate colour texture measuring device VMX2000 has been presented. When implementing geometrical and chromatic calibrations based on suitable calibration targets in a device using an integrated sphere, it has been possible to make measurements on various samples that previously only could be done by a panel of humans making a visual judgement. The success of this implementation can also be expressed as the continuous development of the system.

Two smaller projects involving laser sheets have been presented, they have both shown the need for various camera calibrations in actual industrial applications. They have especially shown the need for automatic detection of calibration targets, since practical implementations cannot wait for a user to manually select the calibration objects.

Appendix A

Developed Software

This appendix gives a brief description of the main parts of the software designed and programmed by Jørgen Folm-Hansen during this Ph.D. project. This software can be considered as a part of this project's contribution to the society of Image Analysis and the future research and development at the section for Image Analysis at IMM. All software has been written in C/C++ and for some of the programs graphical user interfaces has been developed in unix/linux X11/motif and/or Microsoft windows 32bit. Please look at the WWW homepage of the section for Image Analysis (through www.imm.dtu.dk) for further information of the programs below.

HIPS

The standard software used for image analysis at IMM has during many years been the HIPS/HIPS-2 package from SharpImage Software New York. There has been made a lot of work on both the public domain version (hipspub) and the commercial version (HIPS-2) within this project.

A major contribution has been the porting of the complete HIPS-2 and HIPSPUB packages from UNIX to Microsoft win32 console programs and libraries. This work was based on the Microsoft Visual C++ 5.0 compiler. Due to this conversion all console based HIPS program now run in a MSDOS-prompt under Microsoft Windows 9x/NT.

- *wizard* for automatic ANSI-C source code generation of HIPSPUB programs, the code is generated on basis on user selectable options in a graphical user interface under X11/Motif called *xwiz*.

- *binfractile* makes an extreme fast local binary quantile in a HIPS sequence, can be used for both binary morphology and median filtering.

- *pcd2hips* convert from the Kodak PhotoCD to HIPS-2 image format (max. resolution is 768x512).

- *jpeg2hips* convert from the JPEG image format to HIPS-2. Based on the library by the Independent JPEG Group's software.

- *hips2tiff* convert a HIPS sequence into a tiff sequence, based on Jens Michael Carstensen’s first version of the same program and Sam Leffler’s TIFF library.

- *tiff2hips* convert a TIFF sequence into a HIPS sequence, based on Jens Michael Carstensen’s first version of the same program and Sam Leffler’s TIFF library.

- *tiffseq2hips* convert a numbered group of tiff images into a HIPS sequence. Based on Sam Leffler’s TIFF library.

- *rgb2cmap* converts a RGB HIPS image to a colour map HIPS image, based on a statistical selection in the various colour spaces. Very useful with 8 bit colour monitors.

- *mpeg2hips* convert a MPEG-1 movie stream to a HIPS sequence. Based on the MPEG library by the Portable Video Research Group at Stanford.

- *hips2dxf* convert a HIPS depth image to a 3D AutoCAD DXF file.

- *render* makes a very simple rendering of a HIPS depth map.

Special Software

The X11/unix program *xv 3.10a* has been modified for supporting almost all HIPS filetypes.

ITEMAN is a control program (Motif/X11) for an ITEM XYZ-table with a Motorola VME digital I/O board, only for Silicon Graphics ONYX Computers, this equipment is located in the Image Lab. at IMM.

AFMimg is a complete package for capturing, calibrating and measuring image sequences. It is made in a graphical and user-friendly user interface under Microsoft Windows 9X/NT and has been widely used in the cooperation with the Department of Fluid Mechanics, see chapter 5 and [Matlok98] for further details.

Winshow is the prototype software for the VMX2000 colour texture measurement device, see chapter 5. It can make automatic detection of calibration targets and use it for both geometrical and chromatic calibrations. It also controls the movement of the sphere and communicates with the frame-grabber.

Camera Calibration

Cali is a special camera calibration program for digital cameras. The implemented model calibrates for radial as well as tangential distortion, using an iterative Levenberg-Marquardt nonlinear optimization. Requires a calibration object with well known 3D points or multiple 2D points (planar objects) grabbed from various angles/heights. This program was made during a visit at the Department of Electrical Engineering, University of Oulu, Finland and is mainly based on a MATLAB program by Janne Heikkilä.

Simulation

IMAQA is an explorative simulation tool based on X11/Motif for simulation of 3D geological structures. Uses an explorative way of dealing with imported data (GSLIB and HIPS). This tool also gives the ability of interacting in a simulation during simulation in order to explore new simulation way and to incorporate intuitive data (experience of the user). It is primary used for geologist simulating a facies distribution in 3D. The simulation is based on a morphological Markov Random Field in 3D. This program was made as a part of a research project done in collaboration with Peter Frykman at GEUS.

Appendix B

Line-Jitter

This appendix will describe an experiment trying to detect the level of line-jittering with a camera and frame grabber used frequently in this project. The reason for trying to detect the line-jitter is various descriptions of horizontal pixel shifts depending of the intensity level of the grabbed image caused by errors in the synchronization between the analog camera signal and the frame grabber [Heikkilä97b]. When making a positioning of a group of calibration objects it is very critical that there are no systematic displacements, like a horizontal shift caused by the use of different illumination of the calibration target.

The experiment is made using an integrated sphere (from a prototype VMX videometer) with a diameter of 360 mm illuminated from four optic fibre cables connected to a Fostec DCRII (DLL) light source with an adjustable iris. The use of an integrated sphere gives almost complete uniform light on the calibration target and the iris on the light source makes it possible to adjust the illumination level without changing the temperature (spectral distribution) of the light significant. The voltage (temperature) setting on the Fostec light source was set at the step just before *"70"* in all the grabbed images and had about 15 min pre-warming.

The used camera was a JAI CV-M50 PAL industrial machine vision camera with a Schneider-Kreuznach 23 mm lens. The frame grabber was a Matrox Meteor/RGB in "CCIR via RGB" mode.

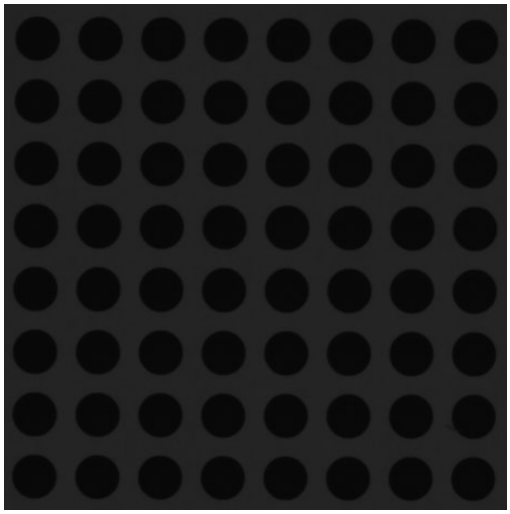


Figure B.1 (a) lowinten.hips

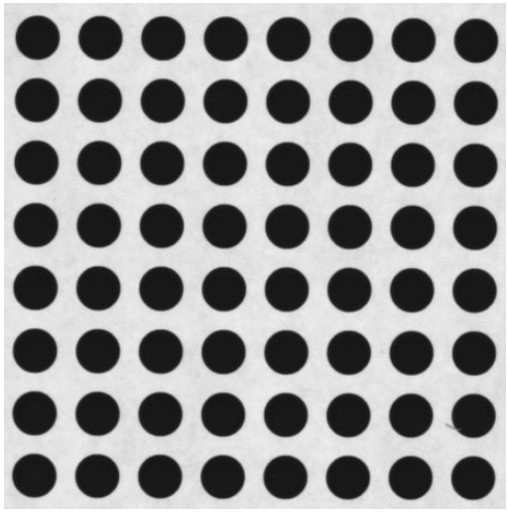


Figure B.1 (b) highinten.hips

The distance between camera and calibration target was about 380 mm, which when the wanted accuracy in the image plane is 0.01 pixel gives about 1.3 μm in world coordinates (100 mm / 768 pixels * 0.01 pixel), so the movement of the camera between two measurements must be smaller than 1.3 μm which in practice means that the setting not should be touched at all.

Two floating point images was grabbed, each the average of 25 normal 256 grey level PAL images, the first with the Fostec iris at “1.5" and the second with the iris at “6". The filenames and simple image characteristics are showed in table B.1.

Image	Iris on Fostec	Lowest pixel value	Highest pixel value
lowinten.hips	1.5	8	37
highinten.hips	6	22	239

Table B.1 Data on the two grabbed images of the calibration target.

In figure B.1 (a) and (b) are the two grabbed images, notice the very low intensities in *“lowinten.hips”*.

In figure B.2 (a) and (b) are the linearly stretched histograms for the two images, notice the small difference in appearance caused by the better signal/noise ratio using an average of 25 images.

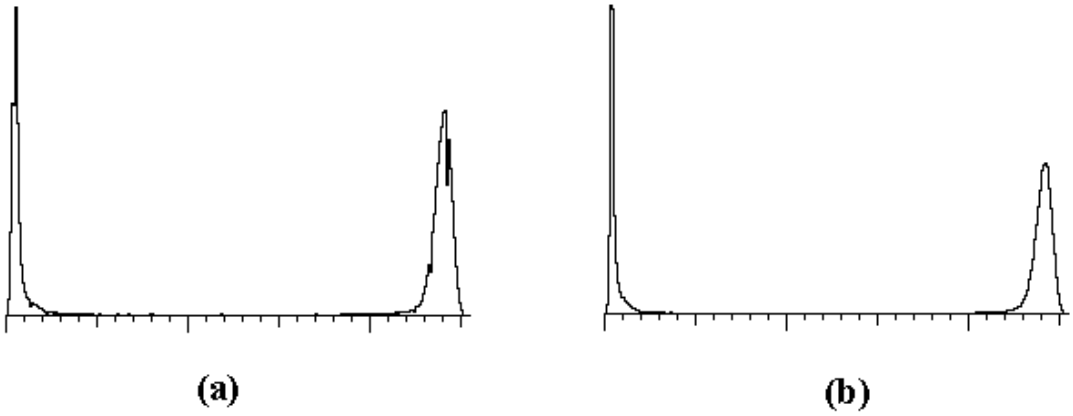


Figure B.2 Histogram for (a) low intensity image and (b) high intensity, both are linearly stretched to be between 0 and 255.

In order to detect any horizontal (or vertical) shift between the high intensity and low intensity image, the calibration objects (black disks) where detected automatic using the threshold method described in chapter 4. The result of the automatic detection is in table B.2.

As seen in table B.2 the horizontal displacement was -0.0073 pixels and the vertical was -0.0026 pixels which both should be compared to a std.dev. of about 0.013 pixels on the displacements. There is though a higher number (43/64) of negative values in the horizontal difference, indicating a very small line-jitter effect.

In table B.3 and B.4 are the values from table B.2 presented in a table where rows and columns correspond to the black disks in figure B.1. There is no obvious trend in table B.3 and B.4. This is confirmed in figure B.2 which is a

vector field plot of table B.3 and B.4 which look more like a random generated field than a systematic field. This experiment has been repeated with similar results.

The conclusion on this experiment must be that there is no significant line jitter based horizontal displacement between a high and a low intensity image.

num	lowinten.hips			highinten.hips			Displacements		
	Area	x	y	Area	x	y	Area	horiz. (x)	verti. (y)
1	1294	30.273	31.382	1294	30.273	31.382	0	0.0000	0.0000
2	1290	87.558	31.601	1289	87.544	31.594	1	0.0143	0.0065
3	1292	144.785	32.000	1294	144.802	32.000	-2	-0.0174	0.0000
4	1291	202.094	32.387	1290	202.084	32.375	1	0.0100	0.0121
5	1284	259.244	32.839	1286	259.261	32.852	-2	-0.0167	-0.0127
6	1292	316.331	33.256	1293	316.316	33.261	-1	0.0150	-0.0044
7	1286	373.348	33.708	1286	373.350	33.708	0	-0.0015	0.0000
8	1292	430.375	34.287	1293	430.390	34.291	-1	-0.0152	-0.0036
9	1288	29.648	88.561	1287	29.657	88.574	1	-0.0091	-0.0129
10	1294	87.124	88.977	1292	87.148	88.989	2	-0.0242	-0.0124
11	1286	144.436	89.378	1285	144.465	89.384	1	-0.0284	-0.0058
12	1286	201.610	89.632	1285	201.618	89.619	1	-0.0083	0.0135
13	1289	258.787	90.067	1289	258.787	90.067	0	0.0000	0.0000
14	1283	315.888	90.501	1283	315.914	90.512	0	-0.0257	-0.0109
15	1289	373.010	90.960	1290	373.019	90.974	-1	-0.0085	-0.0132
16	1279	430.156	91.419	1275	430.148	91.446	4	0.0074	-0.0272
17	1289	29.168	145.824	1289	29.189	145.803	0	-0.0217	0.0210
18	1292	86.568	146.265	1290	86.572	146.295	2	-0.0040	-0.0306
19	1288	143.860	146.554	1286	143.858	146.548	2	0.0018	0.0061
20	1290	201.204	146.940	1292	201.225	146.954	-2	-0.0206	-0.0133
21	1283	258.448	147.415	1284	258.455	147.401	-1	-0.0066	0.0143
22	1285	315.519	147.613	1285	315.519	147.613	0	0.0000	0.0000
23	1279	372.638	148.122	1282	372.672	148.151	-3	-0.0336	-0.0285
24	1281	429.686	148.564	1281	429.686	148.564	0	0.0000	0.0000
25	1303	28.664	203.200	1302	28.679	203.197	1	-0.0151	0.0037
26	1298	86.214	203.582	1294	86.211	203.569	4	0.0032	0.0136
27	1305	143.490	203.959	1306	143.500	203.971	-1	-0.0096	-0.0123
28	1293	200.695	204.364	1293	200.695	204.364	0	0.0000	0.0000
29	1290	258.000	204.671	1289	258.016	204.673	1	-0.0155	-0.0013
30	1296	315.186	205.048	1295	315.202	205.048	1	-0.0155	-0.0001
31	1290	372.355	205.483	1293	372.372	205.491	-3	-0.0170	-0.0082
32	1295	429.414	205.743	1295	429.423	205.748	0	-0.0093	-0.0046
33	1306	28.296	260.774	1304	28.310	260.802	2	-0.0142	-0.0280
34	1302	85.689	261.154	1301	85.674	261.159	1	0.0148	-0.0047
35	1300	143.078	261.540	1298	143.089	261.556	2	-0.0109	-0.0162
36	1301	200.334	261.881	1300	200.349	261.885	1	-0.0148	-0.0045
37	1295	257.543	262.302	1296	257.547	262.317	-1	-0.0042	-0.0152
38	1294	314.662	262.573	1293	314.663	262.558	1	-0.0005	0.0158
39	1297	371.847	262.994	1299	371.866	262.978	-2	-0.0188	0.0161
40	1295	429.030	263.231	1297	429.043	263.259	-2	-0.0131	-0.0282
41	1298	27.770	318.408	1297	27.779	318.395	1	-0.0091	0.0128
42	1299	85.341	318.816	1299	85.341	318.816	0	0.0000	0.0000
43	1298	142.633	319.223	1297	142.621	319.234	1	0.0119	-0.0102
44	1293	199.866	319.597	1293	199.866	319.597	0	0.0000	0.0000
45	1293	257.177	319.964	1293	257.177	319.964	0	0.0000	0.0000
46	1292	314.321	320.270	1290	314.338	320.266	2	-0.0168	0.0042
47	1287	371.479	320.514	1287	371.470	320.532	0	0.0085	-0.0186
48	1294	428.528	320.724	1293	428.537	320.737	1	-0.0089	-0.0129
49	1291	27.493	375.690	1290	27.494	375.674	1	-0.0004	0.0158
50	1291	84.962	376.184	1290	84.961	376.200	1	0.0008	-0.0156
51	1287	142.351	376.588	1287	142.363	376.587	0	-0.0117	0.0008
52	1288	199.583	377.045	1286	199.604	377.022	2	-0.0211	0.0232
53	1282	256.653	377.408	1280	256.681	377.406	2	-0.0283	0.0017
54	1280	313.843	377.638	1278	313.812	377.633	2	0.0308	0.0053
55	1288	370.983	377.944	1288	371.004	377.928	0	-0.0210	0.0163
56	1288	428.172	378.147	1291	428.177	378.126	-3	-0.0042	0.0204
57	1282	27.268	432.858	1280	27.292	432.866	2	-0.0239	-0.0084
58	1281	84.596	433.432	1281	84.596	433.432	0	0.0000	0.0000
59	1282	141.924	433.766	1281	141.956	433.775	1	-0.0319	-0.0092
60	1276	199.225	434.203	1276	199.225	434.203	0	0.0000	0.0000
61	1278	256.419	434.484	1278	256.419	434.484	0	0.0000	0.0000
62	1276	313.462	434.680	1277	313.469	434.695	-1	-0.0067	-0.0143
63	1271	370.622	434.936	1274	370.622	434.920	-3	-0.0008	0.0156
64	1270	427.630	435.221	1271	427.633	435.237	-1	-0.0027	-0.0155
Average	1,289.48	229.040	233.414	1,289.25	229.047	233.417	0.234	-0.0073	-0.0026
Std.dev.	7.75	132.133	132.558	7.59	132.133	132.557	1.540	0.0125	0.0130

Table B.2 Results from automatic positioning of high and low int images.

0.0000	-0.0091	-0.0217	-0.0151	-0.0142	-0.0091	-0.0004	-0.0239
0.0143	-0.0242	-0.0040	0.0032	0.0148	0.0000	0.0008	0.0000
-0.0174	-0.0284	0.0018	-0.0096	-0.0109	0.0119	-0.0117	-0.0319
0.0100	-0.0083	-0.0206	0.0000	-0.0148	0.0000	-0.0211	0.0000
-0.0167	0.0000	-0.0066	-0.0155	-0.0042	0.0000	-0.0283	0.0000
0.0150	-0.0257	0.0000	-0.0155	-0.0005	-0.0168	0.0308	-0.0067
-0.0015	-0.0085	-0.0336	-0.0170	-0.0188	0.0085	-0.0210	-0.0008
-0.0152	0.0074	0.0000	-0.0093	-0.0131	-0.0089	-0.0042	-0.0027

Table B.3 Horizontal difference between low and high intensity image of 8 x 8 black disks.

0.0000	-0.0129	0.0210	0.0037	-0.0280	0.0128	0.0158	-0.0084
0.0065	-0.0124	-0.0306	0.0136	-0.0047	0.0000	-0.0156	0.0000
0.0000	-0.0058	0.0061	-0.0123	-0.0162	-0.0102	0.0008	-0.0092
0.0121	0.0135	-0.0133	0.0000	-0.0045	0.0000	0.0232	0.0000
-0.0127	0.0000	0.0143	-0.0013	-0.0152	0.0000	0.0017	0.0000
-0.0044	-0.0109	0.0000	-0.0001	0.0158	0.0042	0.0053	-0.0143
0.0000	-0.0132	-0.0285	-0.0082	0.0161	-0.0186	0.0163	0.0156
-0.0036	-0.0272	0.0000	-0.0046	-0.0282	-0.0129	0.0204	-0.0155

Table B.4 Vertical difference between low and high intensity image of 8 x 8 black disks.

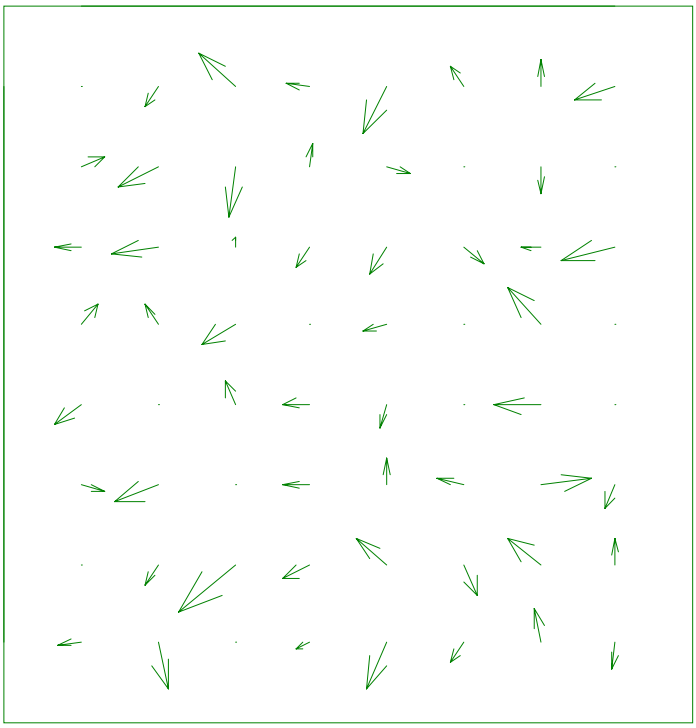


Figure B.1 Vector field of positioning displacement between low and high intensity images.

Appendix C

Warping Experiments

In this appendix is presented some useful experiments about warping as an approximation to the perspective transformation as it is given in (9.1) and how the quality depends on the order of the warping.

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \dfrac{a_1 x + a_2 y + a_3}{c_1 x + c_2 y + 1} \\ \dfrac{b_1 x + b_2 y + b_3}{c_1 x + c_2 y + 1} \end{bmatrix} \tag{9.1}$$

The experiments is based on an image (figure 9.1) grabbed at Dantec Measurement A/S using a Kodak ES-1 1k x 1k camera with a 60 mm lens.

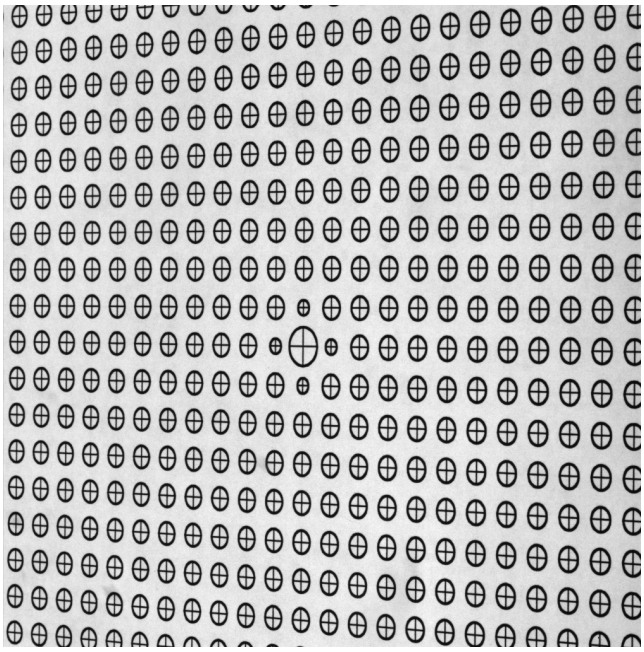


Figure 9.1 1008x1018 image of a calibration sheet.

The positioning of the calibration objects (circles with crosses) was fully automatically detected using an implementation based on the method described on page 138. The positioning was not corrected for PPD causing a systematic horizontal error of about 0.2 pixel in this case, but this does not influence the measurements made for comparing the warp model with the perspective model in (9.1) since post-calculated image coordinates is used.

The grabbed image contains 342 usable calibration objects, objects to close to the border is dismissed, the used objects are shown in figure 9.2. This high number of calibration objects allows for high order warp parameter estimations.

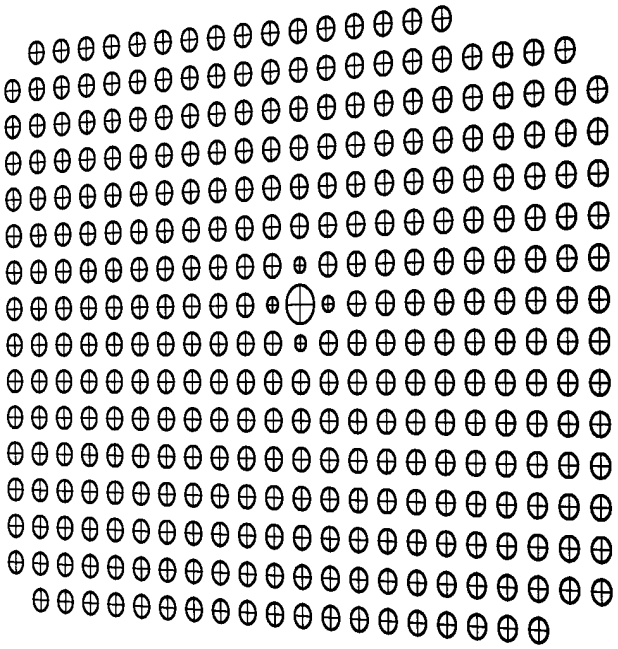


Figure 9.2 The calibration objects used in the estimation.

In order to support the statement in this thesis about the possibility to approximate a perspective transform with a polynomial based warp, the 3x3 DLT of the objects in figure 9.2 was estimated from coplanar world coordinates with origin in the large circle with the small neighbours and an interspace of

5mm, the result is in (9.2).

$$A^* = \begin{bmatrix} 8.168036160 & 0.102733030 & 473.531594097 \\ -0.791546731 & 12.368866447 & 475.094689671 \\ -0.001361188 & 0.000003354 & 1.000000000 \end{bmatrix} \quad (9.2)$$

The inverse of the estimated A^* matrix was then used to calculate new image coordinates from the world coordinates. The result of this operation was a coordinate series of 342 points that gives a perfect fit to the perspective transform in (9.1). In this way any defect in the positioning of the objects and all nonlinear distortion is removed and this coordinate series can be used for comparing warping to perspective transformation.

The coordinate series was then used for least square fits of 6 different warpings from 1st to 6th order. The warp coefficient for 1st to 4th order is presented below for the curious reader that e.g. want to see how the 0 order coefficient change with the order of the warp. The coefficients are presented in the same “layout” as they are calculated from (9.1).

N = 1 [a]
485.025222622998110 8.771931590980099
0.108637822787573
[b]
475.437536874936090 -0.088439236220123
12.311293412346435

N = 2 [a]
473.598338088434050 8.841031711064952 0.011920177687410
0.101373114936770 0.000112927571443
-0.000012268294371
[b]
475.020484078028010 -0.144390482323547 -0.000123088094922
12.387414902703833 0.016715508766558
-0.000005772731429

N = 3 [a]
473.516133571642110 8.812837251874061 0.012050884268201 0.000016193880124

```
  0.101161830262388    0.000108212779774    0.000000113386237
 -0.000001290283917    0.000000000692818
 -0.000000024469040
[b]
 475.093869615941740  -0.145025855198966  -0.000196788054552  -0.000000166940933
  12.367456478584247   0.016885994447829   0.000022725537394
 -0.000040784056415  -0.000000083483149
 -0.000000056105070
```

N = 4

```
[a]
473.53146611809757   8.812542703206496   0.011996216978930   0.000016425287667   0.000000021998247
  0.101145933742047   0.000108123003861   0.000000107244429   0.000000000110440
 -0.000000423164838  -0.000000002141410   0.000000000010708
 -0.000000001893531  -0.000000000038096
  0.000000000075696
[b]
 475.09476716515582  -0.144855450673869  -0.000197528064465  -0.000000267711363  -0.000000000218016
  12.367248845680185   0.016835194425350   0.000023017228090   0.000000030881835
 -0.000041367471291  -0.000000113750634  -0.000000000200183
 -0.000000003607767  -0.000000000007598
 -0.000000000093576
```

The result of this experiment is given in table C.1 that shows the mean and maximum value of the absolute errors in u (horizontal), v (vertical) and $dr=sqrt(du^2+dv^2)$ (distance) for the 1st to the 6th order warping.

Image coordinate differences		Order of warping					
		1	2	3	4	5	6
Mean	du	8.9815316	0.3356527	0.0124482	0.0004746	0.0000174	0.000000620
	dv	8.5904753	0.3219149	0.0120696	0.0004465	0.0000166	0.000000624
	dr	13.8109430	0.5182659	0.0193395	0.0007177	0.0000267	0.000000968
Max	du	22.7639569	0.9026006	0.0325656	0.0010955	0.0000364	0.000001350
	dv	29.8205552	1.4335696	0.0591883	0.0022032	0.0000768	0.000002660
	dr	37.4503131	1.6917938	0.0675558	0.0024451	0.0000850	0.000002886

Table C.1 Differences between image coordinates calculated from perspective transform of world coordinates and image coordinates warped (order 1 to 6) of world coordinates.

The conclusion on this must be that when approximating a perspective transformation using a polynomial based warping the order should be at least a full 3rd order warping. When a warping should model distortion and a perspective transformation with a high subpixel precision the use of a warping with an order larger than 3 should be considered.

An interesting aspect of the warping approximation of the perspective transformation is given in figure 9.3, where it is seen that the precession of the approximation gets about a factor 25 better for each increase in the order of the warping.

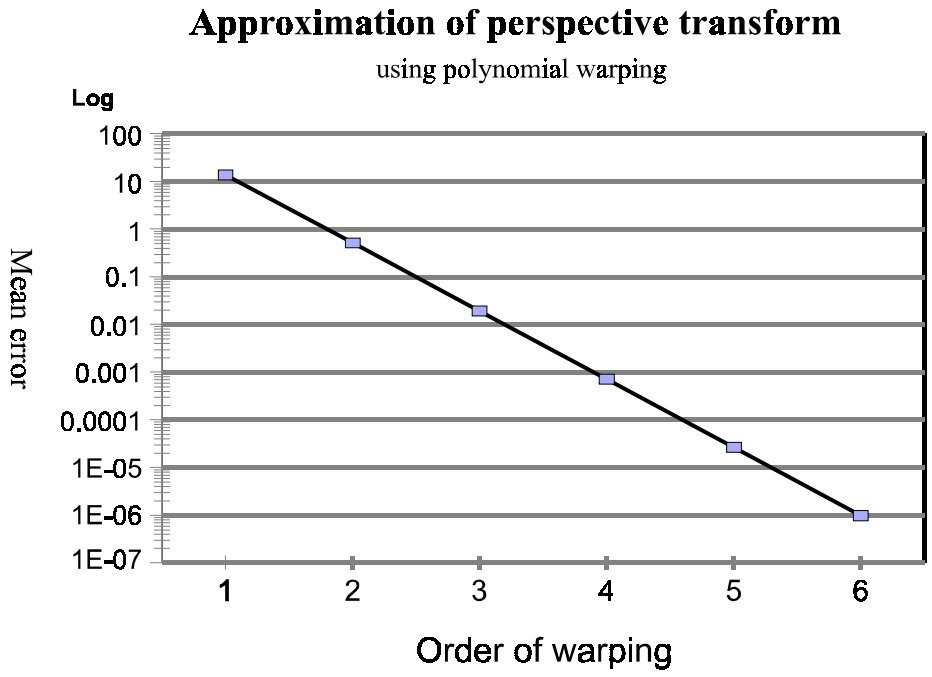


Figure 9.3 Logarithm Plot of mean error on approximation of perspective transformation using a polynomial warping against the order of the warping.

When using the estimated DLT model to transform from world coordinates to image coordinate and comparing them with the actual observed pixel coordinates the mean error is 0.479 pixels. The same error when comparing the warp with the observed image coordinates gives an error of 0.34 pixels with a 3^{rd} order, 0.24 pixels with a 4^{th} order and 0.23 pixels with a 5^{th} order. It is probably the measurement error on the calibration objects that is about 0.2 pixels, since the distortion is very low and should be sufficient modelled by a 5^{th} order polynomial.

Appendix D

Inverse Nonlinear Distortion Experiment

This experiment is made for control of the statement in [Weng92] that the nonlinear distortion model also is an approximation of its inverse. The experiment has been made entirely in MathCad7 by MathSoft which explains the notation.

The distortion model is made of a two coefficient radial-symmetrical distortion and a two coefficient decentering distortion. In the presented example there is only test for one image coordinate (1.0, 3.0) which with a 1/3" CCD chip is close to the border where the distortion is largest, the coefficients for the distortion has been manually chosen and should be representive.

The experiment, which included many point, showed that it is a good approximation as long as the distortion is relative small.

Radial symmetrical distortion is set to:

$$k_1 := 0.0001$$

$$k_2 := 0.00000000$$

Decentering distortion is set to:

$$p_1 := 0.0001$$

$$p_2 := 0.0002$$

Here we only present it for this point:

$$u := 1$$

$$v := 3$$

The result of applying the distortion to this point is:

$$\begin{bmatrix} uu \\ vv \end{bmatrix} := \begin{bmatrix} u\cdot\left[k_1\cdot\left(u^2+v^2\right)+k_2\cdot\left(u^2+v^2\right)^2\right]+p_1\cdot 2\cdot u\cdot v+p_2\cdot\left[\left(u^2+v^2\right)+2\cdot u^2\right] \\ v\cdot\left[k_1\cdot\left(u^2+v^2\right)+k_2\cdot\left(u^2+v^2\right)^2\right]+p_1\cdot\left[\left(u^2+v^2\right)+2\cdot v^2\right]+p_2\cdot 2\cdot u\cdot v \end{bmatrix} + \begin{bmatrix} u \\ v \end{bmatrix}$$

$$uu = 1.0040001$$

$$vv = 3.0070003$$

If inserting the two distorted points in the same model for back-projection we get:

$$\begin{bmatrix} uuu \\ vvv \end{bmatrix} := \begin{bmatrix} uu \\ vv \end{bmatrix} - \begin{bmatrix} uu \cdot \left[k_1 \cdot \left(uu^2 + vv^2 \right) + k_2 \cdot \left(uu^2 + vv^2 \right)^2 \right] + p_1 \cdot 2 \cdot uu \cdot vv + p_2 \cdot \left[\left(uu^2 + vv^2 \right) + 2 \cdot uu^2 \right] \\ vv \cdot \left[k_1 \cdot \left(uu^2 + vv^2 \right) + k_2 \cdot \left(uu^2 + vv^2 \right)^2 \right] + p_1 \cdot \left[\left(uu^2 + vv^2 \right) + 2 \cdot vv^2 \right] + p_2 \cdot 2 \cdot uu \cdot vv \end{bmatrix}$$

$uuu = 0.999973946162625$

$vvv = 2.999956912529397$

Where the error can be expressed as:

$$err_u := \frac{(u - uuu)}{u}$$

$$err_v := \frac{(v - vvv)}{v}$$

$err_u = 2.605383737497302 \cdot 10^{-5}$

$err_v = 1.436249020099941 \cdot 10^{-5}$

As seen the error on the back-projection is very small and the approximation seems to be quite good.

References

Andersen H.J. & Granum E. (1999). *Classifying Illumination Condition from two light Sources by Colour Histogram Assessment*. Article from AUC Denmark 16th march 1999. To be published in Journal of Optical Society of America A in 1999.

Blaszka T. & Deriche R. (1994). *Recovering and characterizing image features using an efficient model based approach*. Rapport de Recherche 2422, Institute National de Recherche en Informatique et en Automatique, November 1994.

Carstensen J.M. (1992). *Description and Simulation of Visual Texture*. Ph.D.-thesis DTU IMM 1992. ISSN 0107-525X.

Carstensen J.M. (1996). *An active lattice model in a bayesian framework*. Computer Vision and Image Understanding, 63(2), page 380-387.

Carstensen J.M., Folm-Hansen J. & Rank M. (1999). *Colorimetric image analysis of denim.* Internal IMM technical report. Expected to be published in journal during 1999.

Conradsen K. (1984). *En introduktion til Statistik*, Bind 2A and 2B, (IMSOR) Institute of Mathematical Modelling Technical University of Denmark, Lyngby, Denmark 1984.

Dalsgaard E., Skettrup T. and Owner-Petersen M. (1989). *Optik, optiske komponenter*. Polyteknisk Forlag, Lyngby, Denmark, 2. oplag 1992. ISBN 87-502-0686.

Dodgson Neil A. (1997). *Quadratic Interpolation for Image Resampling*. IEEE Transaction on Image Procesing. Volume 6. No. 9. September 1997.

Folm-Hansen J. & Carstensen J.M.(1997). *Kalibrering af farvekameraer med fokus på 3xCCD kameraer*. The Sixth Danish Conference on Pattern Recognition and Image Analysis (DSAGM) 1997, DIKU, Copenhagen, Denmark.

Fisker R, Poulsen H F, Schou J, Carstensen J M & Garbe S (1998). *Use of Image-Processing Tools for Texture Analysis of High-Energy X-ray Synchrotron Data*. Journal of Applied Crystallography, pp 647-653, ISSN 0021-8898, 1998.

Foley J. D., Dam A., Feiner S. K. & Hughes J. F. (1992). *Computer Graphics - Principles and Practice*. Addison-Wesley Publishing Company 1992.

Frydendal I. (1998). *Quality Inspection of Sugar Beets using Vision*. Ph.D.-thesis IMM 1998. ISSN 0909-3192.

Gonzalez R.C. & Wintz P. (1987). *Digital Image Processing*. Second Edition. Addison Wesley Publishing Company USA 1987. ISBN 0-201-11026-1.

Gramkow C. (1998). *Subpixel estimation of reference and calibration templates using parametric fits,* in Danish (Sub-pixel estimation af reference- og kalibreringsmærker v.h.a. parametriske funktioner.). Quarterly report 5 for ATV (EF660) 26th marts 1998, Odense Steel Shipyard A/S 01.70 Automation.

Haralick R.M. & Shapiro L.G. (1993). *Computer and Robot vision*. Volume 2. Addison-Wesley Publishing Company 1993. ISBN 0-201-56943-4 (v.2).

Hartelius K. (1996). *Analysis of Irregularly Distributed Points*. Ph.D.-thesis DTU IMM 1996-21. ISSN 0909-3192. Page 145-233.

Heikkilä J. & Silven O (1996a). *Accurate 3-D measurement using a single video camera*. International Journal of Pattern Recognition and Artificial Intelligence, Volume 10, No. 2., pp. 139-149.

Heikkilä J. & Silven O (1996b). *Calibration procedure for short focal length off-the-shelf CCD-cameras.* Proc. 13th International Conference for Pattern Recognition, Volume 1, August 25-30, Vienna, Austria, pp. 166-170.

Heikkilä J. & Silven O. (1997a). *A four-step camera calibration procedure with implicit image correction.* Proc. IEEE Conference on Computer Vision and Pattern Recognition, June 17-19, San Juan, Puerto Rico, 1:1106-1112.

Heikkilä J. (1997b). *Accurate Camera Calibration and feature based 3D reconstruction from monocular image sequences*. Ph.D.-thesis, Infotech Oulu and Department of Electrical Engineering, University of Oulu, Finland. October 1997..

Jacobi O. (1984). *Fotogrammetri Matematisk Grundlag*. Department of Planning Technical University of Denmark 1984.

Jacobsen D. (1997). *Photographic Lenses FAQ.* WWW posting from Hewlett-Packard Laboratories by David Jacobson, last modified 1997/03/11. Jacobsen@cello.hpl.hp.com.

Knudsen Jon (1998). *Digitalisering af svampebilleder og gruppering af skimmelsvampe ved hjælp af billedanalyse.* Master thesis from the Department of Biotechnology (in Danish) - Technical University of Denmark 1998.

Knudsen Ole (1998). Ph.D.-thesis from IMM and Odense Steel shipyard, ISSN 0909-3192.

Kraus Karl (1993). *Photogrammetry, Fundamentals and Standard Processes* (Volume 1). Ferd. Dümmlers Verlag, 5300 Bonn 1, Germany. 386 pp.

Lassen Niels Christian Krieger (1994). *Automated determination of crystal orientations from electron backscattering patterns*. Ph.D.-thesis from IMM 1994. 136 pp.

Lawson N.J. & Wu J. (1997). *Three-dimensional particle image velocimetry: experimental error analysis of a digital angular stereoscopic system*. Meas. Sci. Technol. 8, page 1455-1464, 1997.

Marszalec E. & Pietikäinen M. (1997). *Color Mesurements Based on a Color Camera*. SPIE, Volume 3101, 0277-786X/97.

Matlok S., Larsen P.S., Gjernes E. and Folm-Hansen Jørgen (1998). *Mixing studies in a 1:60 scale model of a cornerfired boiler with OFA*. 8th. International Symposium on Flow Visualization. Sept.1-4, 1998, Sorrento, Italy.

Melen T. (1994). *Geometrical modelling and calibration of video cameras for underwater navigation*, Ph.D.-thesis 94-103-W, Norges Tekniske Høgskole, Trondheim, Norway 1994.

Niblack W. (1985). *An Introduction to Digital Image Processing*. Strandberg Publishing Company, Birkerød, Denmark 1985. ISBN 87-872-0055-4.

Pertl F. A. J., Craven R. P. M., Smith J. E. & Duhn C. (1997). *Color measurement of non-homogenous colored flat surfaces using a PC-based color camera vision system.* SPIE Vol. 3205, 0277-786X/97.

Pratt W. K. (1991). *Digital Image Processing*. John Wiley & Sons inc. 2nd edition, USA 1991. ISBN 0-471-85766-1.

Press W.H, Flannery B.P., Teukolsky S.A. and Vetterling W.T. (1986). *Numerical Recipes, The Art of Scientific Computing*. Cambridge University Press, ISBN 0521308119, USA, 1986.

Russ J. C. (1990). *Computer-Assisted Microscopy - The Measurement and analysis of Images*. Plenum Press New York. ISBN 0-306-43410-5. Third Printing March 1992.

Sonka M., Hlavac V. & Boyle R. (1993). *Image Processing, Analysis and Machine Vision*. Chapman & Hall Computing London UK 1993, ISBN 0-412-45570-6.

Sternberg J. (1997). *Scanning and modelling of 3D objects*. Master Thesis 1997-12, IMM, Technical University of Denmark ISSN 0909-6256, Lyngby, Denmark 1997.

Tsai R. Y. (1986). *An Efficient and Accurate Camera Calibration Technique for 3D Machine Vision*. CH2290-5/86/0000/0364$01.00 © 1986 IEEE.

Walther A. (1995). *The Ray and Wave Theory of Lenses*. Cambridge University Press 1995.

Weng J., Cohen P. and Herniou M. (1992). *Camera Calibration with Distortion Models and Accuracy Evaluation*. IEEE transactions on Pattern Analysis and Machine Intelligence, Volume 14, No. 10, October 1992, page 965-980.

Ph. D. theses from IMM

1. **Larsen, Rasmus**. (1994). *Estimation of visual motion in image sequences.* xiv + 143 pp.

2. **Rygard, Jens Moberg**. (1994). *Design and optimization of flexible manufacturing systems.* xiii + 232 pp.

3. **Lassen, Niels Christian Krieger**. (1994). *Automated determination of crystal orientations from electron backscattering patterns.* xv + 136 pp.

4. **Melgaard, Henrik**. (1994). *Identification of physical models.* xvii + 246 pp.

5. **Wang, Chunyan.** (1994). *Stochastic differential equations and a biological system.* xxii + 153 pp.

6. **Nielsen, Allan Aasbjerg**. (1994). *Analysis of regularly and irregularly sampled spatial, multivariate, and multi-temporal data.* xxiv + 213 pp.

7. **Ersbøll, Annette Kjær**. (1994). *On the spatial and temporal correlations in experimentation with agricultural applications.* xviii + 345 pp.

8. **Møller, Dorte.** (1994). *Methods for analysis and design of heterogeneous telecommunication networks.* Volume 1-2, xxxviii + 282 pp., 283-569 pp.

9. **Jensen, Jens Christian**. (1995). *Teoretiske og eksperimentelle dynamiske undersøgelser af jernbanekøretøjer.* viii + 174 pp.

10. **Kuhlmann, Lionel.** (1995). *On automatic visual inspection of reflective surfaces.* Volume 1, xviii + 220 pp., (Volume 2, vi + 54 pp., fortrolig).

11. **Lazarides, Nikolaos.** (1995). *Nonlinearity in superconductivity and Josephson Junctions.* iv + 154 pp.

12. **Rostgaard, Morten.** (1995). *Modelling, estimation and control of fast sampled dynamical systems.* xiv + 348 pp.

13. **Schultz, Nette.** (1995). *Segmentation and classification of biological objects.* xiv + 194 pp.

14. **Jørgensen, Michael Finn.** (1995). *Nonlinear Hamiltonian systems.* xiv + 120 pp.

15. **Balle, Susanne M.** (1995). *Distributed-memory matrix computations.* iii + 101 pp.

16. **Kohl, Niklas.** (1995). *Exact methods for time constrained routing and related scheduling problems.* xviii + 234 pp.

17. **Rogon, Thomas.** (1995). *Porous media: Analysis, reconstruction and percolation.* xiv + 165 pp.

18. **Andersen, Allan Theodor.** (1995). *Modelling of packet traffic with matrix analytic methods.* xvi + 242 pp.

19. **Hesthaven, Jan.** (1995). *Numerical studies of unsteady coherent structures and transport in two-dimensional flows.* Risø-R-835(EN) 203 pp.

20. **Slivsgaard, Eva Charlotte.** (l995). *On the interaction between wheels and rails in railway dynamics.* viii + 196 pp.

21. **Hartelius, Karsten.** (1996). *Analysis of irregularly distributed points.* xvi + 260 pp.

22. **Hansen, Anca Daniela.** (1996). *Predictive control and identification - Applications to steering dynamics.* xviii + 307 pp.

23. **Sadegh, Payman.** (1996). *Experiment design and optimization in complex systems.* xiv + 162 pp.

24. **Skands, Ulrik.** (1996). *Quantitative methods for the analysis of electron microscope images.* xvi + 198 pp.

25. **Bro-Nielsen, Morten.** (1996). *Medical image registration and surgery simulation.* xxvii + 274 pp.

26. **Bendtsen, Claus.** (1996). *Parallel numerical algorithms for the solution of systems of ordinary differential equations.* viii + 79 pp.

27. **Lauritsen, Morten Bach.** (1997). *Delta-domain predictive control and identification for control.* xxii + 292 pp.

28. **Bischoff, Svend.** (1997). *Modelling colliding-pulse mode-locked semiconductor lasers.* xxii + 217 pp.

29. **Arnbjerg-Nielsen, Karsten.** (1997). *Statistical analysis of urban hydrology with special emphasis on rainfall modelling.* Institut for Miljø teknik, DTU. xiv + 161 pp.

30. **Jacobsen, Judith L.** (1997). *Dynamic modelling of processes in rivers affected by precipitation runoff.* xix + 213 pp.

31. **Sommer, Helle Mølgaard.** (1997). *Variability in microbiological degradation experiments - Analysis and case study.* xiv + 211 pp.

32. **Ma, Xin.** (1997). *Adaptive extremum control and wind turbine control.* xix + 293 pp.

33. **Rasmussen, Kim Ørskov.** (1997). *Nonlinear and stochastic dynamics of coherent structures.* x + 215 pp.

34. **Hansen, Lars Henrik.** (1997). *Stochastic modelling of central heating systems.* xxii + 301 pp.

35. **Jørgensen, Claus.** (1997). *Driftsoptimering på kraftvarmesystemer.* 290 pp.

36. **Stauning, Ole.** (1997). *Automatic validation of numerical solutions.* viii + 116 pp.

37. **Pedersen, Morten With.** (1997). *Optimization of recurrent neural networks for time series modeling.* x + 322 pp.

38. **Thorsen, Rune.** (1997). *Restoration of hand function in tetraplegics using myoelectrically controlled functional electrical stimulation of the controlling muscle.* x + 154 pp. + Appendix.

39. **Rosholm, Anders.** (1997). Statistical methods for segmentation and classification of images. xvi + 183 pp.

40. **Petersen, Kim Tilgaard.** (1997). *Estimation of speech quality in telecommunication systems.* x + 259 pp.

41. **Jensen, Carsten Nordstrøm.** (1997). *Nonlinear systems with discrete and continuous elements.* 195 pp.

42. **Hansen, Peter S.K.** (1997). *Signal subspace methods for speech enhancement.* x + 226 pp.

43. **Nielsen, Ole Møller.** (1998). *Wavelets in scientific computing.* xiv + 232 pp.

44. **Kjems, Ulrik.** (1998). *Bayesian signal processing and interpretation of brain scans.* iv + 129 pp.

45. **Hansen, Michael Pilegaard.** (1998). *Metaheuristics for multiple objective combinatorial optimization.* x + 163 pp.

46. **Riis, Søren Kamaric.** (1998). *Hidden markov models and neural networks for speech recognition.* x + 223 pp.

47. **Mørch, Niels Jacob Sand.** (1998). *A multivariate approach to functional neuro modeling.* xvi + 147 pp.

48. **Frydendal, Ib.** (1998.) *Quality inspection of sugar beets using vision.* iv + 97 pp. + app.

49. **Lundin, Lars Kristian.** (1998). *Parallel computation of rotating flows.* viii + 106 pp.

50. **Borges, Pedro.** (1998). *Multicriteria planning and optimization. - Heuristic approaches.* xiv + 219 pp.

51. **Nielsen, Jakob Birkedal.** (1998). *New developments in the theory of wheel/rail contact mechanics.* xviii + 223 pp.

52. **Fog, Torben.** (1998). *Condition monitoring and fault diagnosis in marine diesel engines.* xii + 178 pp.

53. **Knudsen, Ole.** (1998). *Industrial vision.* xii + 129 pp.

54. **Andersen, Jens Strodl.** (1998). *Statistical analysis of biotests. - Applied to complex polluted samples.* xx + 207 pp.

55. **Philipsen, Peter Alshede.** (1998). *Reconstruction and restoration of PET images.* vi + 132 pp.

56. **Thygesen, Uffe Høgsbro.** (1998). *Robust performance and dissipation of stochastic control systems.* 185 pp.

57. **Hintz-Madsen, Mads.** (1998). *A probabilistic framework for classification of dermatoscopic images.* xi + 153 pp.

58. **Schramm-Nielsen, Karina.** (1998). *Environmental reference materials methods and case studies.* xxvi + 261 pp.

59. **Skyggebjerg, Ole.** (1999). *Acquisition and analysis of complex dynamic intra- and intercellular signaling events.* 83 pp.

60. **Jensen, Kåre Jean.** (1999). *Signal processing for distribution network monitoring.* xv + 199 pp.

61. **Folm-Hansen, Jørgen.** (1999). *On chromatic and geometrical calibration.* xiv + 238 pp.