# Multi-camera system for close-up supervision in test-fixtures using openCV and optics

Author: Henrik Liljefors
November 7, 2023
Supervisors: Andreas Ehn and Johan Mauritsson
Supervisor at mikrodust: Ola Thernström
Examiner: Lars Rippe
LRAP-595

# Abstract

The test-fixture is a device used for testing of electronic circuits in production, that is produced by the Lund based company mikrodust AB. In the test-fixture, cameras are both used for supervision of products-in-testing and in the products themselves. This poses a challenge due to the compact size of the test-fixture and the number of mechanical elements.

This thesis aims to investigate how clear and readable images can be captured in the test-fixture, while also not be to hard to implement from a mechanical stand-point. Solutions using both optical components and image processing algorithms are investigated. Moreover, also to explore the ideal lens combination and its optimal placement based on specific camera attributes, including field of view, minimum focus distance, and physical size.

The results of this thesis show that this problem could be solved with the use of optical components while the use of image processing algorithms were slightly redundant. This work also provides information on what solution is most appropriate given a camera product's characteristics.

# Populärvetenskaplig sammanfattning

Företaget mikrodust AB producerar så kallade test-fixturer. Test-fixturen är en apparat som används för att testa elektriska kretsar och produkter i produktion. Genom att strömsätta produkterna som testas i fixturen så kan det verifieras att produkter som ska till försäljning fungerar som avsett.

I test-fixturen används kameror på två olika sätt. Dels används de för övervakning av produkterna som testas i test-fixturen för att till exempel kontrollera att en LCD-skärm fungerar som den ska. Det andra sättet är när en kamera sitter på någon av produkterna som testas inuti test-fixturen, då måste man verifiera att kamerorna fungerar. Kamerorna som testas i fixturen verifieras genom att de tar en bild på ett referensmönster som sitter i test-fixturen, bilden jämnförs sedan med en referensbild. Användningen av kameror kan vara problematisk inuti test-fixturen på grunda av dess kompakta storlek, det är inte ovanliga att man i fixturen enbart har cirka 60 mm i fri höjd. Många kamerorna som testas i fixturen har ett fokusavstånd på några meter, alltså det avstånd som krävs för att en kamera ska ta bilder som är i fokus. När man försöker använda kameror med ett långt fokusavstånd på ett kort avstånd så blir resultatet oläsligt, bilderna kommer att vara suddiga på grunda av att kameran inte är i fokus. Detta är ett problem hos mikrodust eftersom suddiga bilder inte kan jämnföras mot en referensbilden och kan därmed inte avgöra om kameran fungerar som den ska.

För att lösa detta problem har man hos mikrodust tidigare använt sig av positiva linser, genom att placera en positiv lins mellan kameran och referensmönstret så kan man korta ner fokusavståndet för kameran. Detta har emellertid enbart varit framgångsrikt i specifika situationer, och det har inte

funnits något tydligt tillvägagångsätt för att avgöra vilka linser som bör användas för varje specifika scenario och hur de bör positioneras. För att lösa situationerna som inte kan lösas med enbart positiva linser så introduceras båda negativa linser och speglar.

Förutom att utreda vilka optiska komponenter som ska användas vid specifika scenarion så utreddes även vad som kunde åstadkommas med bildalgoritmer. Algoritmer för att kompensera för suddighet och distortion i bilderna implementerades.

Resultatet visade vilka optiska komponenter som skulle användas och hur de skulle placeras berode på vilken typ av kamera som testades. Det blev även tydligt att bildbehandlings algoritmer inte var nödvändiga för att lösa problemen som uppstod.

# Acknowledgements

I would like to thank my supervisors at LTH, Johan and Andreas. Thank you for your support and patience. I would also like to thank my colleagues at Mikrodust, especially my supervisor Ola and Johan for answering all my programming questions.

# Contents

# Chapter 1

# Introduction

## 1.1 Background

### 1.1.1 Mikrodust test-fixture

A test-fixture is a device that is used to test electronic circuits in manufacturing. The test fixture can have a fixed camera located on top of the fixture, which is used for supervision during the testing of for instance LEDs and displays in a product. Certain products that are tested in the fixture can also contain a variety of different cameras, which needs to be verified that they work properly. That is done by capturing an image of a reference object located in the top of the test-fixture.

### 1.1.2 Aims

The aim of this project is to develop and evaluate a generalized method for a two way close-range camera supervision in the Mikrodust test fixture. Using both a system of optical components and the open-source software openCameraVision.
The first case contains a fixed camera of a certain model, and will be referred to as case I in this thesis. This camera is located in the top of the test fixture, and it is used for surveillance of the circuits being tested.
For the first case, acquiring a clear image is done in three steps:

- Given the size of the product being tested, a lens with a field of view matching the product size is chosen.

- Blur and image distortion is compensated for using openCameraVision.
- The processed image is compared to a reference image to detect any faulty circuits.

The second case that are looked into here, which will be referred to as caseII in this thesis, is when a camera is located on any of the electronic circuits tested in the test-fixture. These tests will be carried out on various camera models and the camera models themselves needs to be verified so that they are working properly. The process is done in two steps:

- Given camera parameters and the distance between the camera and the reference image, the optimal combination of optical components are determined.
- The image that is acquired in the test is compared to a reference image to identify cameras that does not work properly.

### 1.1.3 Challenges

The biggest challenge in the development for the optical systems is to fit it into the compact test-fixture. Most cameras that are being tested have a focusing distance far beyond the height of the test-fixture, and taking clear images without high levels of optical aberrations is extremely challenging. Other difficulties that are faced in this thesis is the amount of changing variables of each case. Since every camera model that is being tested in the test-fixture all have different properties, both cases carries a different set of challenges.

From an image processing perspective, the biggest challenge is time complexity. Every test in the fixture should be as fast as possible, and therefore the computer vision algorithms can not be too computationally heavy.

### 1.1.4 Thesis structure

In chapter two, the theory behind the project is introduced and explained. First, an overview on the theory of the optical systems and how these systems cause blur and distortion in the images are explained, along with and description of how digital cameras acquire and store images. The theory of

computer vision algorithms used to compensate for blur and distortion is also explained.

In chapter three the material that is used in this thesis is presented.

Chapter four describes the methodology that are used in the project is presented and how they are linked to the theory presented from chapter one. The results obtained are presented in chapter five, as well as a discussion on the results.

# Chapter 2

# Theory

## 2.1 Digital cameras

### 2.1.1 Color coding

**Red-blue-green coding** or RGB colour coding, is when every pixel in an image is represented with a green, blue and red value. RGB coding is used to produce regular colour images. **Gray scale** coding is when each pixel is represented with a value of intensity of light. For an image with 8-bit pixel value, 255 represents completely white, while 0 is a completely black pixel. A RGB image can be converted to gray scale by taking the mean of every pixel's RGB-values[11]:

$$I(x,y) = \frac{R(x,y) + B(x,y) + G(x,y)}{3} \tag{2.1}$$

### 2.1.2 Camera concepts

This section will explain some camera concepts that are frequently used in this thesis, which can be seen in figure2.1

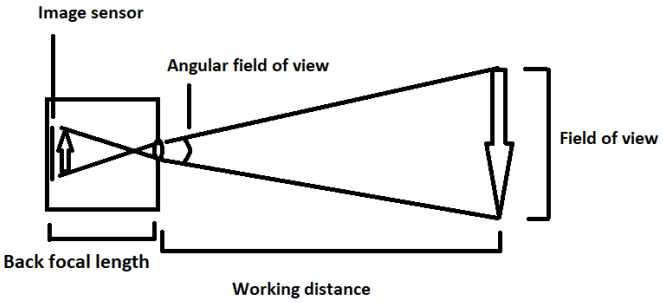



Figure 2.1: Image formation for a simple camera

### Angular field of view

The field of view for a camera is the maximum area a camera can image, see figure 2.1. This is the area which cover the entirety of the camera's image sensor. The angular field of view is the field of view expressed in degrees. It is dependent on the focal length of the lens used and the size of image sensor. It can be mathematically described as[11]:

$$AFOV = 2\tan^{-1}(\frac{h}{2F}) \tag{2.2}$$

Where $h$ is the size of the sensor and $F$ is the focal length of the lens used.

### Working distance

The working distance is the distance between an object and the front element of a camera's lens.

### Aperture

The aperture of a camera refers to the hole where light enters. The size of the aperture controls the amount of light that enters the camera. It is

numerically expressed as the f-stop, which is the focal length divided by the diameter of the opening of the aperture[11].

#### Back focal length

The back focal length of a camera is defined as the distance between the last surface of the last optical element and the cameras image sensor.

#### Minimum focus distance

The minimum focus distance is the smallest working distance were a camera can capture a all-in-focus image.

## 2.2 Optical Components

### 2.2.1 Positive and negative lenses

Lenses are optical components that focuses or disperse light. How light waves changes direction when they pass from one medium to another can be described with Snells law of refraction:

$$n_1 \sin(\theta_1) = n_2 \sin(\theta_2) \tag{2.3}$$

Where $n_x$ are the mediums refractive index and $\theta_1$ is the angle of incidence ray and $\theta_2$ is the angle of the refracted ray[9].

Lenses that focuses incoming light rays are called positive lenses. The incoming light will refract according to Snell's law and converge towards the positive lens' focal on the opposite side of the origin of the light waves. The positive lens will create an inverted real image if the object is located further away than its focal length, meaning the image it produces will be able to be projected. Lenses that disperse incoming light rays are called negative lenses. A negative lens has a negative focal length, and incoming light rays will be diverged from the lens focal point, which is located on the left side of the lens.

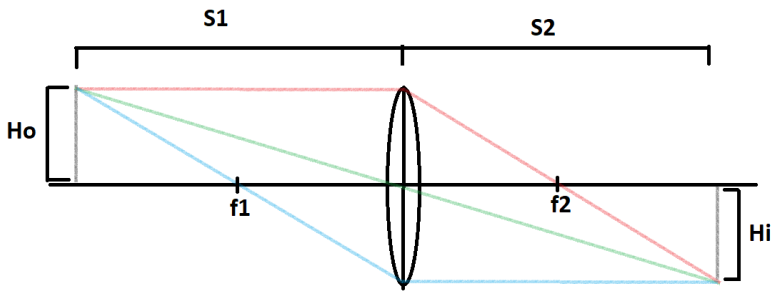



Figure 2.2: Image formation for a positive lens

The image formation, assuming that the lenses are thin, for both types of lenses can be described with the thin lens formula:

$$\frac{1}{s_1} + \frac{1}{s_2} = \frac{1}{f} \tag{2.4}$$

Here $s_1$ is the distance between the object and the lens, $s_2$ is the distance between the formed image and the lens and $f$ is the lens focal length.

The magnification of a lens can be described with the formula:

$$m = \frac{s_2}{s_1} = \frac{h_i}{h_o} \tag{2.5}$$

Where $h_i$ is the height of image and $h_o$ is the height of the object. A positive lens will in most cases create a real and inverted image, which means that the magnification will in most cases be larger than 1. While a negative lens will create a virtual and non-inverted image, which means that the magnification for a negative lens will be in most cases less than 1.

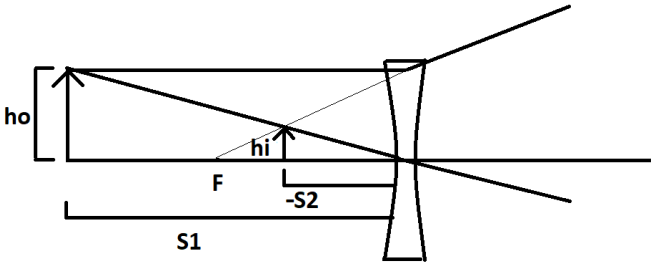



Figure 2.3: Image formation for a negative lens

### 2.2.2 Reversed Galilean telescope

A reversed Galilean telescope uses a negative lens as an objective and a positive lens as an eyepiece lens. The reversed Galilean telescope is what is known as a visual field extender, and will to the negative lens as the first element create a reduced image. Since it creates a reduced image the field of view is increased. The lenses are placed so the both the lenses front focal point coincide, see figure2.4, which means that the total system magnification can be described as:

$$m = \frac{f_{negative}}{f_{positive}} \tag{2.6}$$

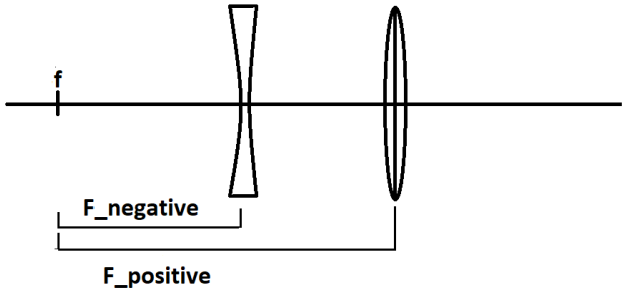



Figure 2.4: Lenses placed in a reversed Galilean formation

### 2.2.3 Optical mirrors

#### Plane mirrors

Unlike lenses when rays coincides with mirrors the rays are reflected rather than refracted. Snell's law provides the law of reflection, that states that the angle of incidence is equal to the angle of reflection. Both the reflected rays and the refracted rays angles are measured form the normal of the surface. The plane mirror creates a virtual image behind the mirror at the same distance as the object's distance from the mirror, and has a constant magnification of 1.[3]

## 2.3 Image processing

### 2.3.1 Out-of-focus Defocus

Blurring in images can be caused by a few different factors such as: defocus, motion and insufficient depth of field. Since stationary cameras and objects are used and every part of the imaging object is at the same distance from the camera, it can be assumed that all image blurring in this case will be the result of defocusing.
In a perfect image system each imaging point will result in a small point on the camera's image sensor. However, if the focus plane is not located on the camera's image sensor the point will instead be a circle, called the circle of confusion[6]. The diameter of the circle of confusion can be described as:

$$\sigma = \frac{k}{A} \frac{f^2 |u - u_f|}{u(u_f - f)} \tag{2.7}$$

Where A is the f-number, $u_f$ is the working distance when the camera is in focus ,$u$ is the working distance out-of-focus and $f$ is the focal length of the lens. The blur kernel that is used in the image deblurring algorithms is retrieved by the result from equation 2.5.

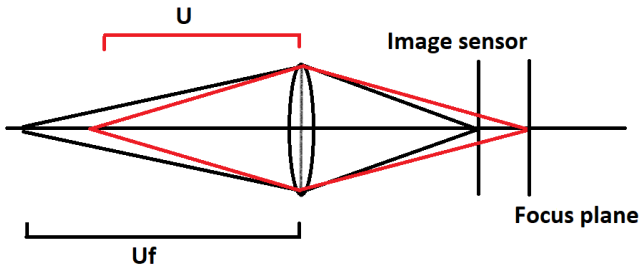



Figure 2.5: Defocus of a imaging system. Where the black rays show a point that is in focus, and the red rays are of a point that is out-of-focus.

### 2.3.2 Point spread function

The point spread function can be seen as an optical system's impulse response. It describes how an imaging system responds to a point source or a point object[12]. In a system with a coherent light source the point spread function of a point source can be described as the Gaussian[10]:

$$h(x, y) = \frac{1}{2\pi\sigma^2} \exp(-(x^2 + y^2)/2\sigma^2) \tag{2.8}$$

Where x and y are pixel coordinates and $\sigma$ diameter of the circle of confusion[6]. Since the aperture is round it is assumed that the point spread function is the shape of a disk.

### 2.3.3 Image deblurring

Blurring in an image can be described as the discrete linear system:

$$b = h * i + n \tag{2.9}$$

Where * denotes the convolution operator, $b$ is the blurred image, $i$ is the clear latent image, $h$ is the point spread function, also called the blur kernel, and $n$ is the adaptive noise. The processes of estimating the latent image $i$ is called image deblurring. Image deblurring can be split into two categories, blind and non-blind deblurring. In blind deblurring the blur kernel h is unknown, while in non-blind deblurring the blur kernel is known. The most common ways to recover the clear latent image is with the help of either neural networks, Wiener deconvolution or iterative algorithms[13].

#### Wiener Deconvolution

According to the convolution theorem, which states that the convolution of two signals in spacial domain is equal to the point-wise multiplication of their Fourier transforms. From this the equation 2.9 can be re-written as[2]:

$$B = H \cdot I + N \tag{2.10}$$

Where $B$ is the spectrum of the blurred image, $I$ is the spectrum of the true image, $H$ is the blur kernel, which is often represented by the frequency response of the Point spread function and $N$ is the spectrum of additive

noise[8].The most straight forward method to recover the latent image, is by inverse filtering.

$$i = \mathcal{F}^{-1}(\frac{B}{H}) \tag{2.11}$$

Inverse filtering can be efficient, but it is problematic when the values of the blur kernel in the Fourier domain are small. This is often the case for most point spread functions in imaging and optics. Divisions by zero or values close to zero will amplify the additive noise.
The solution to this issue, is to use a Wiener filter. The Wiener filter assumes a probabilistic distribution of the blur and noise, and solves the deblurring problem by minimizing the least square error. The wiener filter is expressed as[4]:

$$H_w = \frac{H}{|H^2| + \frac{1}{SNR}} \tag{2.12}$$

Where $H$ is the spectrum of the blur kernel and $SNR$ is the signal-to-noise ration. To restore the latent image, the inverse convolution also known as the deconvolution is used. The deconvolution in the frequency domain turns to a simple division. By applying the Wiener filter the latent image can be expressed as:

$$i = \mathcal{F}^{-1}(B \cdot H_W) = \mathcal{F}^{-1}(B\frac{H}{|H^2| + \frac{1}{SNR}}) \tag{2.13}$$

### Gradient decent

Since the signal-to-noise ratio can be hard to approximate, a different approach to the image deblurring problem is to solve the linear system in the time domain. Solving the linear system is done by minimizing the data error to the equation below with the respect to $i$ [1]:

$$\tilde{x} = argmin||hi - b||^2 \tag{2.14}$$

One method for solving the above problem is the gradient decent algorithm. The gradient decent is an iterative method, which estimate the sharp image by calculating the gradient of the pixel values. That can be described as: [5]:

$$x_k = x_{k-1} - t\nabla f(x_k) \tag{2.15}$$

Where $x_k$ is the current iteration of the deblurred image, $x_{k-1}$ is the previous iteration of the deblurred image, $\nabla f(x_k)$ is the gradient and $t$ is the learning

rate. The gradient can be described as:

$$\nabla f(x) = h' * (I * h - b) \tag{2.16}$$

Where * denotes the convolution, $h$ is the blur kernel and $b$ is the blurred image [5]. This process is done until a acceptable result is obtained.

### 2.3.4 The pinhole model

The pinhole model is the camera model that OpenCV uses. OpenCV is an open source computer vision software. The pinhole model describes the mathematical relationship between coordinates of points in real 3D space and its projection on a 2D image plane[7]. To create a model, a coordinate system is introduced called the camera coordinate system, which has its origin, denoted C, in the camera center. The coordinate system is located so the z-axis is pointing forward, the x-axis is pointing to the right and y is downwards. To generate a 2D: $x = (x, y, 1)$ point form a 3D world point $X = (X, Y, Z)$, a viewing ray is formed between $X$ and the $C$. The line will intersect the image plane located at z = 1, which is referred to as the principal point and is the center of the image. The viewing ray can be parameterized with the expression: $C + s(X - C) = sX$. To find the intersection between the line and the image plane, the third coordinate needs to fulfil $sZ = 1$, which means that $s = 1/Z$. The projection then becomes:

$$x = \begin{pmatrix} X/Z \\ Y/Z \\ 1 \end{pmatrix} \tag{2.17}$$

Since the camera is subject to movement in the 3D world, there needs to be a strategy for model that. A new coordinate system is introduced called the global coordinate system to account for that. The rotation matrix $R$ is a 3x3 matrix. It describes the cameras orientation in the 3D world, which means it describes how much the camera coordinate system is rotated in relation to the global coordinate system. The translation vector $t$ is a 3x1 vector. The translation vector represents the cameras position in the global coordinate system, and how much it moves with respect to the global coordinate system. Both $R$ and $t$ are often represented as one 4x4 matrix as:

$$\begin{pmatrix} R & t \\ 0 & 1 \end{pmatrix} \tag{2.18}$$

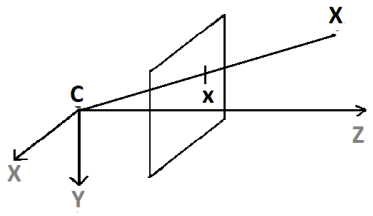



Figure 2.6: Pinhole camera, the mathematical model

$K$ is known as the camera's intrinsic matrix. The intrinsic matrix contains information about the camera's inner parameters and how 3D point in the camera's coordinate system is transformed into 2D image points. The intrinsic matrix can be expressed as:

$$K = \begin{pmatrix} \gamma f & 0 & x_0 \\ 0 & f & y_0 \\ 0 & 0 & 1 \end{pmatrix} \tag{2.19}$$

Where $f$ is the focal length expressed in pixels, $x$ and $y$ is the principal point, and $\gamma$ is the aspect ratio.

The entire pin hole models projection can now be described as:

$$sp = K[R|t]P_w \tag{2.20}$$

Where $p$ is the 2D pixels on the image plane and $P_w$ is the real world coordinates, is $K$ is the camera intrinsic matrix, $R$ is the rotation matrix and $t$ is the translation matrix[7].

### 2.3.5 Image distortion

The pinhole model described above has the property that straight lines are preserved when projected from 3D to the 2D image plane. However, the

pinhole model does not work for fisheye or wide eye lenses. Fisheye lenses have a large field of view but it comes at the expense of introducing radial distortion. Radial distiortion makes straight lines appear curved. Radial distiortion is modeled by moving points away from the principal points and increases further away from the image center points are. The relationship between the undistorted and distorted points is described as[:]

$$x_u \sim \begin{pmatrix} d(r_d)x_d \\ d(r_d)y_d \\ 1 \end{pmatrix} = \begin{pmatrix} d(r_d) & 0 & 0 \\ 0 & d(r_d) & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_d \\ y_d \\ 1 \end{pmatrix} \tag{2.21}$$

Where $r_d = \sqrt{x_d^2 + y_d^2}$ is the distance form the principal point and $d(r_d)$ is the control function. The distortion function controls how far points should be moved. There are different ways to model the control function, but openCV models it as:

$$d(r_d) = 1 + k_1^2 r_d + k_2^4 r_d + k_4^6 r_d \tag{2.22}$$

If the camera's intrinsic parameters as well as the distortion coefficients are known, the radial distortion can be removed calculating[7]:

$$x_u \sim \begin{pmatrix} d(r_d)x_d \\ d(r_d)y_d \\ 1 \end{pmatrix} = \begin{pmatrix} d(r_d) & 0 & 0 \\ 0 & d(r_d) & 0 \\ 0 & 0 & 1 \end{pmatrix} K^{-1} \begin{pmatrix} x_d \\ y_d \\ 1 \end{pmatrix} \tag{2.23}$$

# Chapter 3

# Material

## 3.1 Test camera

As previously mentioned in the case with the fixed camera (case I), i.e. the case when a camera is used for surveillance of products in testing, a single model of camera is used. In this case it is the Arducam OV-5642 mini. The OV5642 has a resolution of 5 megapixeles. The OV-5642 uses both SPI and I2C as communication protocols. Where I2C was used for communication between integrated circuits on the camera printed circuit board. SPI was used for communication between the master chip and the control chip developed by mikrodust.

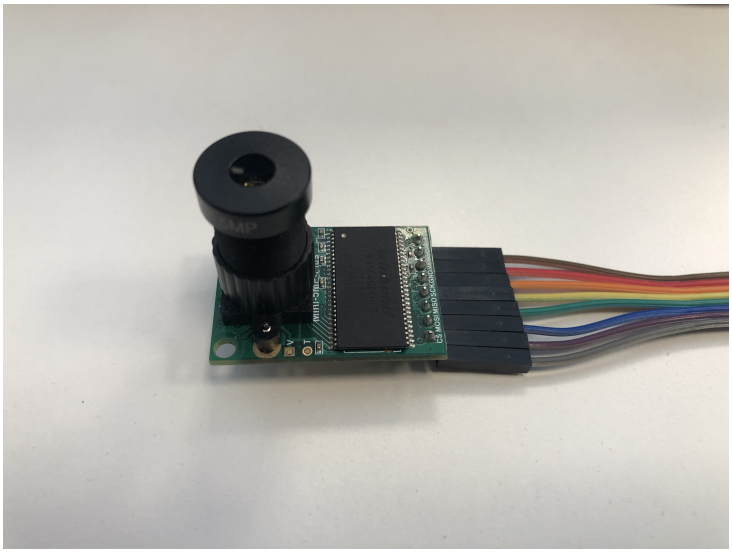


Figure 3.1: OV-5642

## 3.2 M12 lenses

In order to change the field of view and the minimum focusing distance for the OV-5642, a kit with ten different m12 lenses were used. The m12 lenses' angular field of view was in the range between 10-200 degrees, and their focal lengths varied between 16 mm - 0.8 mm.

Figure 3.2: Some of the m12 lenses used

## 3.3 External lenses

For caseII a set of external lenses was used. The lenses that were investigated in this work were ordered before the work started, with the exception of three negative lenses that was purchased during the thesis. Every lens had a diameter of 25 mm and had focal lengths ranging between -75 to 400 mm.

## 3.4 Experimental setup

Two different setups were used for the experiments in this thesis. The first setup was designed by mikrodust, but had previously not been used. The first setup was used to quickly identify the placement of certain lenses to achieve sufficient focus.

The second setup was designed in Autocad in order to simulate the conditions inside the Mikrodust test-fixture. The experimental setup had the same height as the top lid of the test-fixture, where the camera or the reference image is located.

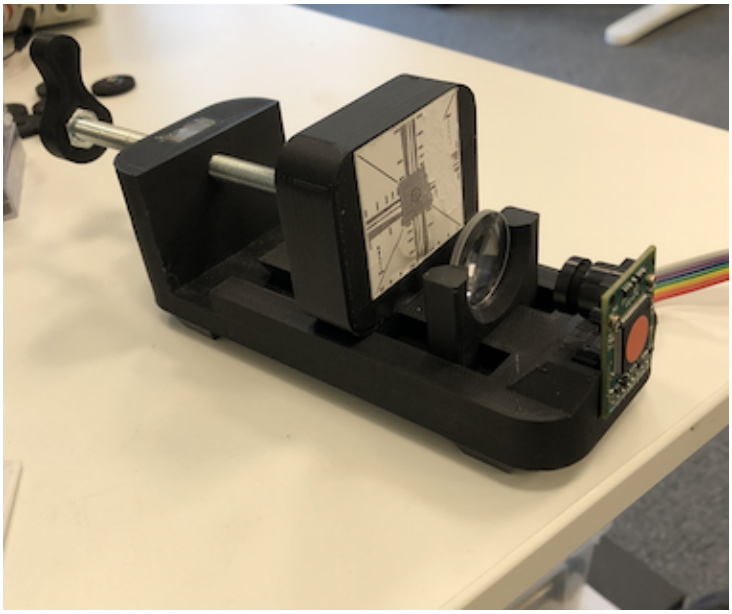


Figure 3.3: The first test setup. In the figure you can see the camera to the right and the reference image to the left. The distances L1 and L2 is adjustable.

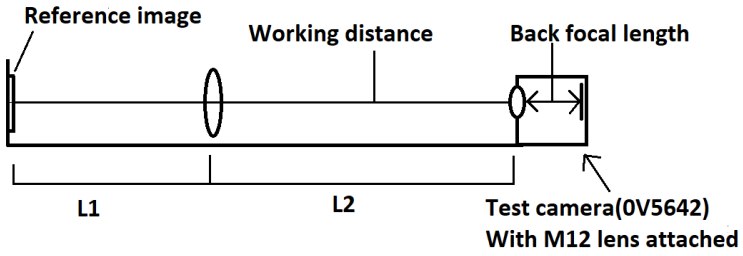
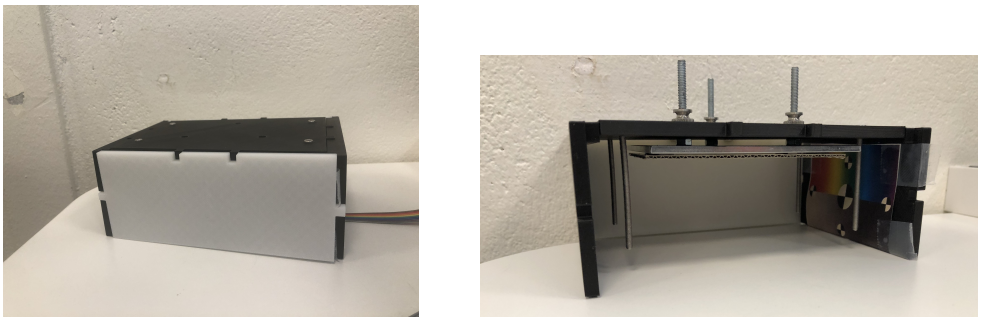



Figure 3.4: A schematic of the first test setup. The distances between the camera and the lenses are quickly verified this way.



(a) (b)

Figure 3.5: The second test setup. Image (a) shows how the setup looks when it is closed. Image (b) how it looks inside of the test setup.

# Chapter 4

# Method

## 4.1 Implementation of test camera

Previously the camera mikrodust used in the test-fixtures was the OV-2640, which has now been discontinued. The first step was therefore to determine a suitable replacement. The OV-5642 was selected, since it uses the same communication protocols. However, the OV-5642 uses 16-bit register address, instead of an 8-bit. So the the old user interface that mikrodust previously used had to be re-written to accommodate a 16-bit register address. Mikrodust uses a text-based graphical user interface to retrieve the images. The interface works by reading in a bit-stream of the image in JPEG-format, the JPEG-format contains a start and end header in the bit-stream. By looking for the start and end header in the bit-stream, every byte containing the JPEG can be saved and the bits before the start header and the bits after the end header was discarded.

## 4.2 Case I

Mikrodust were interested in knowing what m12 lens that should be used given the size of the product that is tested and the distance from the lens to the product. It was apparent from testing that finding focus was not an issue. By adjusting the back focal length, see figure 3.4, every m12 lens was able to find focus with a working distance that was equivalent to the dimensions of the test-fixture. However, some amount of blur was still present. Hence, the

only parameters that will differ for the lenses that were tested is the field of view that is used to image the entire object.

A python script was constructed to identify which one of the m12-lenses should be used, the script carried out simple geometric calculations to determine how big the angular field of view was needed. However, since the numbers given by the producers of the m12 lenses are just approximations, to get a better understanding of what m12 lens needed to be used the first test set up was used. The test worked by simply acquiring images and measuring how large area it covered. The camera performance is not of importance, only that a clear non distorted image can be captured. Images taken from the first case was used to investigate how well the image processing algorithms preformed.

### 4.2.1 Implementation of deblurring algorithms

The two different deblurring algorithms was implemented in Matlab, to investigate how much blur could be present, while still being able to recover an image that is clear enough for mikrodust purposes.

First, the algorithms had to be tested with simulated blur. A clear greyscale image was obtained, where the every point in the image was at the same distance from the camera aperture. The algorithms were verified by applying blur with a known blur kernel, and then attempted to recover the original image. The blur kernel was produced by using the command fspecial in matlab which produces a Gaussian kernel of known size and variance. A blurred image was created by calculating the convolution of the clear image and the blur kernel. The execution time was measured for both algorithms. The different parameters of each algorithm was varied until the best possible result was achieved. Once the algorithms was functioning as expected they were implemented in python, and the point spread function was estimated. By finding at what working distance the object was in focus, and then measuring the working distance of the out-of-focus images. These values could then be inserted in to equation 2.8, to construct the kernel.

### 4.2.2 Distortion compensation

To compensate for the distortion caused by the fish-eye lenses, a classic camera calibration method was used. A known image was acquired with structures that can be identified, in this case a chessboard pattern. Since the pattern is known the coordinates of the 3D real world coordinates are known. In order to find the 2D image points the openCV function findChessBoard-Corners was used, it returns the coordinates of the 2D points. When both the 3D and the 2D points are known the intrinsic camera matrix, the extrinsic camera matrix and the distortion coefficients can be determined. The parameters was saved and are later used to undistort images taken inside the test setup.

## 4.3 Case II

For a given camera product, a few factors needed to be taken into account:

1. How large is the product? Since the test-fixture's top-box has a set height of 57 mm, the working distance is given by $57 - height_{camera}$.

2. How large is the cameras' angular field of view? How much does the image need to be magnified or reduced for the camera to be able to capture the entire image. For the purpose of this thesis, the reference image in figure 4.1 was used.

3. What is the cameras minimum focusing distance? At what distance was the camera able to capture a clear image?

4. The last factor that needed to be taken into account is the mechanical aspect, how are the optical components placed so it does not interfere with other structures in top-box?

Two different possible solutions were examined. Both solution had strengths and weaknesses, the limitations were also examined. The same procedure was used for each solution.

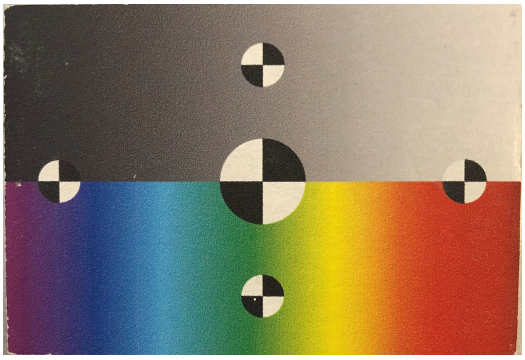


Figure 4.1: Reference image

### 4.3.1 Solution one

The first possible solution was to use external lenses placed between the camera and the reference image. This method has previously been used at mikrodust. They have placed a positive lens between the camera and the reference image when the camera could not achieve a sharp image of the reference image.

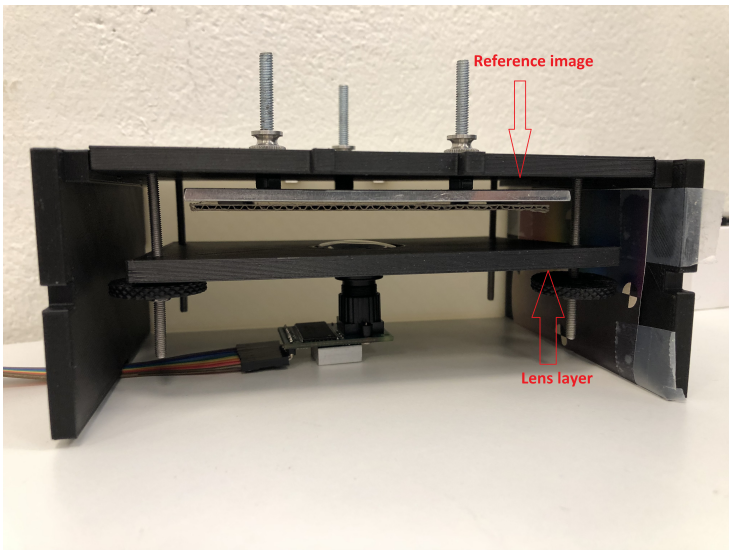



Figure 4.2: Case II, solution one. The reference image is located on the top side of the test setup

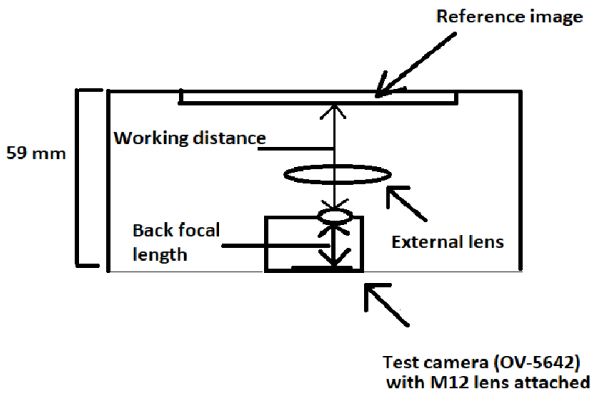



Figure 4.3: Schematic image for solution one for case II

To simulate the conditions with a camera that has a focus distance longer than the height of the test-fixture top-box. A m12 lens was attached to the OV-5642 with the back focal length adjusted so the camera had a focus distance of approximately a 1 meter to infinity. This was done for every m12 lens that was tested. To take a clear image containing a majority of the reference image, different configurations of lenses was used. The placement of the lens or lenses was approximated using the thin lens equation and by the non-linear mappings of the lenses. For example, if a camera had a focusing distance of 1 m to infinity and needed a demaginfication to capture the entire reference image, the process of determining the placement of the lenses were done in two steps. First the location of the virtual image produced by the negative lens was approximated by:

$$\frac{1}{s_1} - \frac{1}{s_2} = -\frac{1}{f_1} \implies \frac{1}{s_2} = \frac{1}{s_1} + \frac{1}{f_1} \implies s_2 = \frac{f_1 \cdot s_1}{s_1 + f_1} \tag{4.1}$$

The virtual image of the negative lens then becomes the object for the positive lens. By placing the positive lens so it is a focal lengths distance of the object, i.e. $s_1 = f_2$, the location of the image becomes:

$$\frac{1}{s_1} + \frac{1}{s_2} = \frac{1}{f_2} \implies \frac{1}{s_2} = \frac{1}{s_1} - \frac{1}{f_2} \implies s_2 \approx \infty \tag{4.2}$$

### 4.3.2 Solution two

The second solution was to introduce a plane mirror into the system to increase the total distance between the camera and the reference image. Since the top-box is wide but not that high, a mirror allowed the reference image to be placed on the side of the top-box instead of on the inside of the top lid. Similar to the first solution the back focal length was adjusted so the

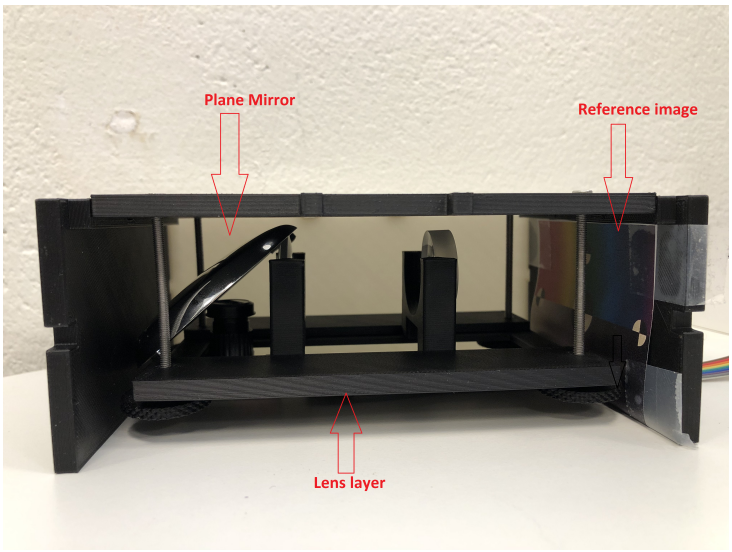



Figure 4.4: Case II, solution two. The reference image is located to the right of the test setup.

camera had a minimum focus distance longer then the working distance. The position of the external lenses used in this solution was approximated in a similar fashion as in the first solution.

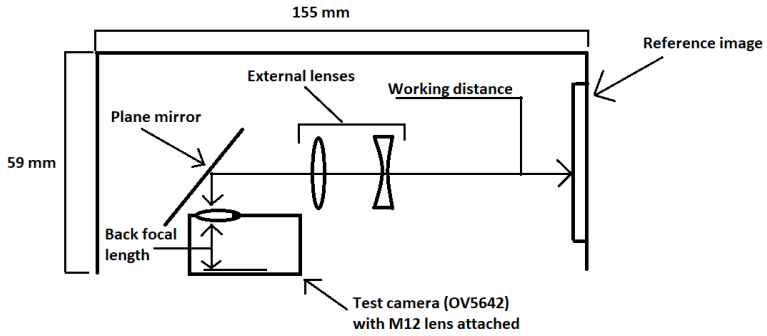



Figure 4.5: Schematic image for solution two for case II.

# Chapter 5

# Discussion and Results

## 5.1 Case I

From testing it became apparent that it was not an issue finding focus, and external lenses would not be needed. Focus could be found at every distance by changing the back focal length. However, some amount of noise and blur was still present. This is likely a problem with the camera itself, since it is a rather cheap product. The biggest problem that arose was when the object was large, in that case a fisheye lens needed to be used. Fisheye lenses by design cause heavy distortion. Distortion can be problematic when for instance a LED-screen needs to be tested. Due to the fact that this cases really did not need to be investigated any further the image processing algorithms were tested. The results for the image deblurring algorithms are presented below, as well as animage pipe line that both deblurrs and undistorts an image.

### 5.1.1 Deblurring

#### Simulated blur

All tests with simulated blur used figure 5.1 as a ground truth. The image was blurred with Gaussian kernels with different width. Blurring is done by calculating the 2D convolution of the original clear image and the blur kernel. For the gradient decent 100 iterations and a learning rate of 0.5 were used. For the Wiener deconvolution a signal-to-noise ratio of 1/0.01 was used. Both algorithms preformed well on the simulated blur, but the images that were

reconstructed with Wiener deconvolution experienced some ringing effects. The ringing effects is likely due to sharp colour gradients in the images. The gradient decent algorithm was slightly slower to execute despite few iterations used.

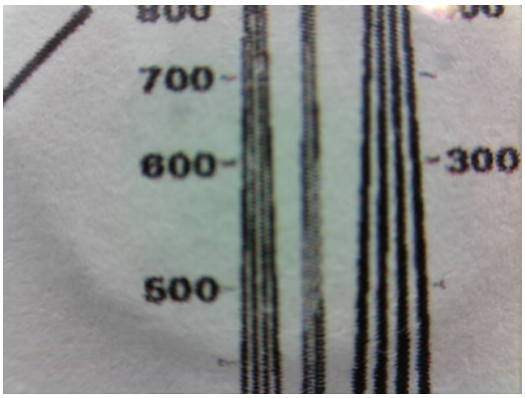
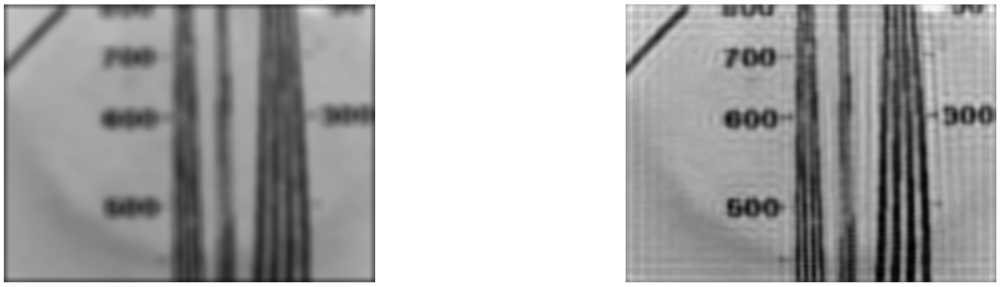
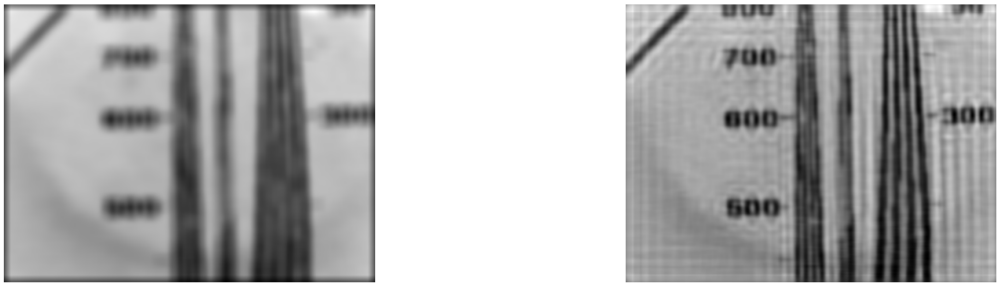



Figure 5.1: The original image.



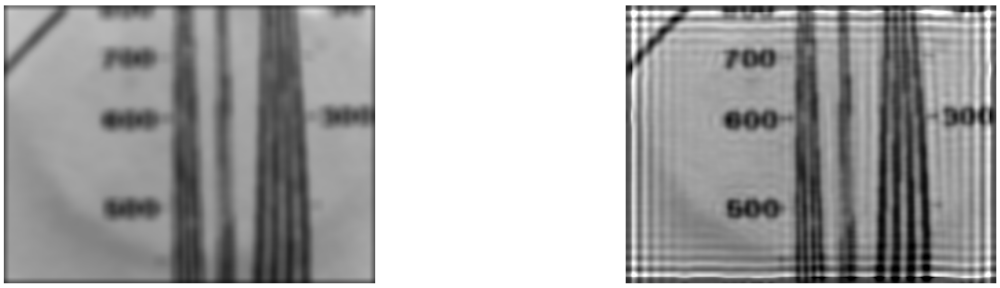

(a) Blurred with a 21x21 kernel (b) Reconstructed image

Figure 5.2: The image to the left shows a the original image that has been blurred with a 21x21 blur kernel. The right image shows the image once it has been reconstructed using gradient decent.



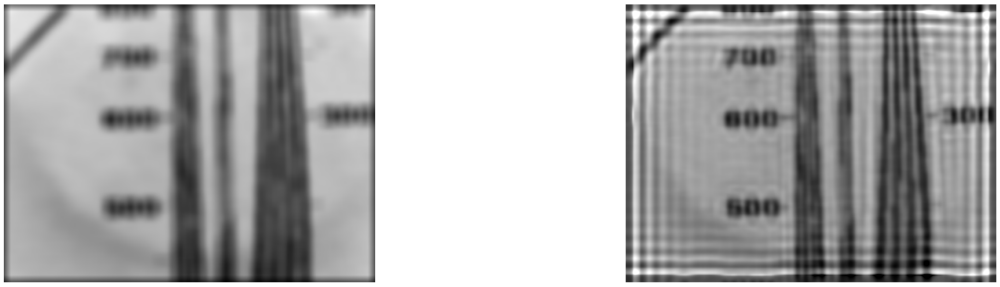

(a) blurred with a 25x25 kernel (b) Reconstructed image

Figure 5.3: The image to the left shows the original image that has been blurred with a 25x25 blur kernel. The right image shows the image once it has been reconstructed using gradient decent.




(a) Blurred with a 21x21 kernel (b) Reconstructed image

Figure 5.4: The image to the left shows a the original image that has been blurred with a 21x21 blur kernel. The right image shows the image once it has been reconstructed using Wiener deconvolution.



(a) Blurred with a 25x25 kernel

(b) Reconstructed image

Figure 5.5: The image to the left shows a the original image that has been blurred with a 25x25 blur kernel. The right image shows the image once it has been reconstructed using Wiener deconvolution.

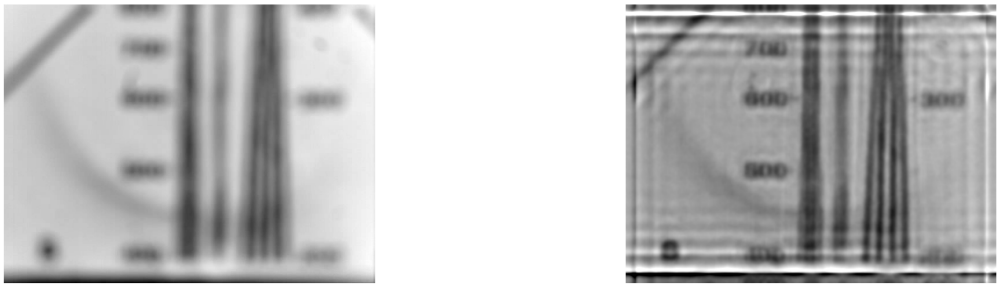
### 5.1.2 Real blur

The figures below show naturally occurring blur due to the camera being defocused. The circle of confusion was estimated using the method described in section 4.2.1. The kernel size was estimated to have a size of 41x41 pixels. For the real blur the deblurring algorithms did not perform that well. Both methods suffer form ringing, which is in this case likely due to that the image does not have enough high frequencies. The ringing is the black lines that appear in the recovered images. The reason the ringing is more severe for real deblurring is likely due to the fact that the kernel being not being estimated correct. The algorithms preformed well with a kernel known, so it can be assumed that the method used to estimate the blur kernel is flawed.

(a) Blurred image

(b) Reconstructed with estimated kernel

Figure 5.6: This figure shows a naturally blurred image to the left, and the reconstructed image with an estimated kernel and gradient decent.



(a) Blurred with a 25x25 kernel (b) Reconstructed image

Figure 5.7: Wiener deconvolution

Figure 5.8: This figure shows a naturally blurred image to the left, and the reconstructed image with an estimated kernel and Wiener deconvolution.

| Algorithm | kernel size | Execution time |
|---|---|---|
| Wiener deconvolution | 21x21 | 0.062505 |
| Wiener deconvolution | 25x25 | 0.072634 |
| Wiener deconvolution | 40x40 | 0.055012 |
| Gradient decent | 21x21 | 4.936729 |
| Gradient decent | 25x25 | 4.956476 |
| Gradient decent | 25x25 | 9.107195 |

Table 5.1: Exectution times for both alogrithms

### 5.1.3 Deblurr and undistorted

The original image is captured using a m12 lens with a 180° angular field of view and a focal length of 1.71 mm. To estimate the intrinsic, extrinsic camera matrices as well as the distortion coefficients, the procedure described in the methods in 4.2.2 is used. The image is first deblurred using gradient decent and then the image is undistorted. For the figure bellow no external lenses were used and the working distance was approximately 30 mm.


Figure 5.9: Original image.


Figure 5.10: Image that has both been deblurred and undistorted

## 5.2 Case II

When testing a product in the test-fixture, images taken by products in testing are compared to image taken by a product that is proven to work as expected. How the images are compared varies, but usually a few things needs to be included. Every acquired image needs to have a clear center of focus, i.e there must be some part of the image that is in focus. The black and white spheres, see figure 4.1, needs to be located on a straight line and the camera should be able to handle the entire frequency spectrum. This means the solutions should produce an image that contains almost the entire reference image, with the exception of the corners, and have a clear center of focus. As mentioned in the method section, two different approaches were investigate, and the results are presented below.

### 5.2.1 Solution 1

Given the lenses that mikrodust already had, this solution was appropriate to use if the height of the camera product was small, had a short minimum focus distance and a larger field of view, see figure 5.11. However, if the camera product was large and had a low angular field of view, only one lens could be used and the entire image could not be captured, see figure 5.12. Unfortunately, for most cases two lenses needed to be used, since just using a positive lens would not allow for the camera to capture the entire reference image. The use of two lenses meant that quite a bit of working distance is needed. Even some cases where it was possible to place to lenses, if the distance between those lenses was to small the image would suffer from distortion, see figure 5.13. By using the non-linear mappings of the lenses' the smallest working distance that gave a reasonable result was 42 mm. However, this could probably be reduced using external lenses with a short focal length. From a mechanical stand point solution one is preferred since it is only vertical was the only direction that ensured free field of view from any mechanical elements.

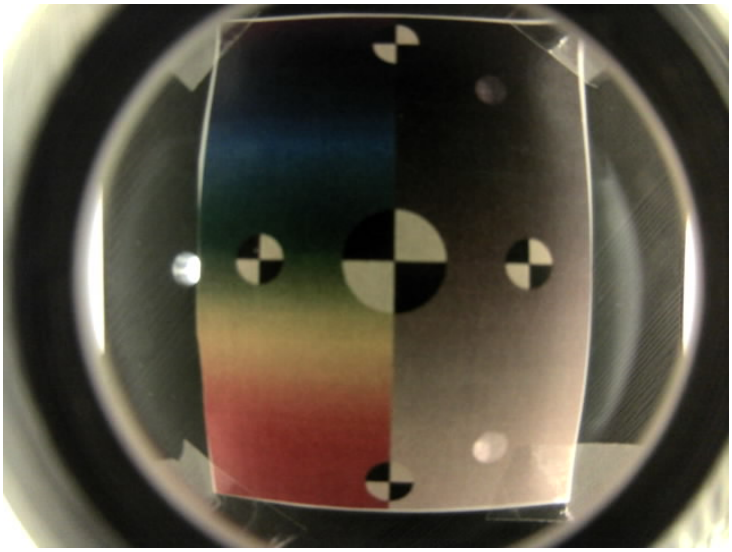


Figure 5.11: A m12 lens with a 140° angular field of view, and a minimum focus distance of approximately 30 cm. One external lens was used with the focal length of 150 mm. Here the entire reference image is visible and is the focus is sufficient for mikrodust's use.

| figure | Working distance | L1 | L2 | L3 | focal length |
|---|---|---|---|---|---|
| 5.11 | 50 mm | 45 mm | 5 mm | - | 1.56 mm |
| 5.12 | 35 mm | 30 mm | 5 mm | - | 8 mm |
| 5.13 | 43 mm | 25 mm | 5 mm | 12.5 mm | 3.6 mm |

Table 5.2: Table of values for the three tests. L1 is the distance between the reference image and the external lens, L2 is the distance between the external lens and the m12 lens, L3 is the distance between the two external lenses if two lenses were used and focal length is the focal length of the M12 lens.

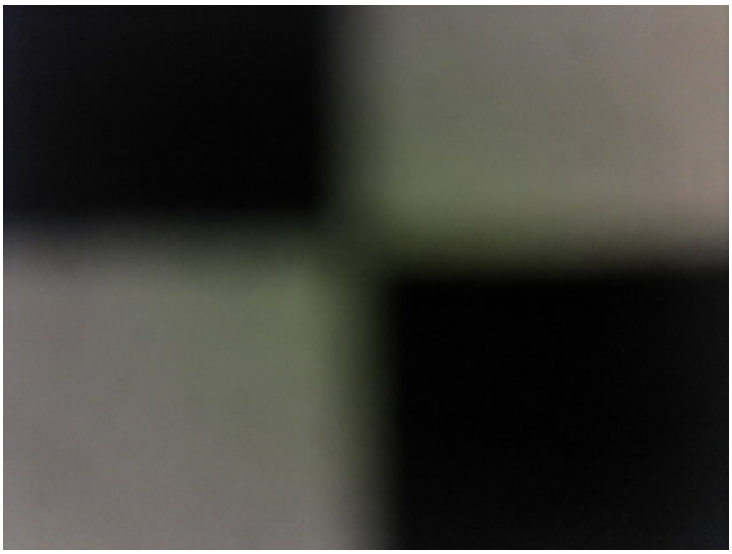


Figure 5.12: A m12 lens with a 26° angular field of view, and a minimum focus distance of approximately 1m. One external lens was used with the focal length of 30mm. Here only a only a fraction of the reference image is captured, therefore this is not a usable image.

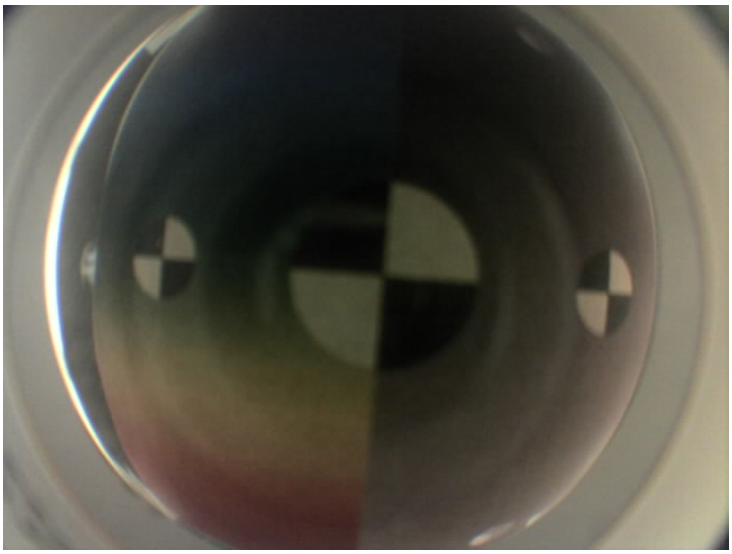


Figure 5.13: A m12 lens with a 67° angular field of view, and a minimum focus distance of approximately 1m. Two external lenses was used, with the focal lengths of -25 mm and 30 mm. The entire reference image is not in the captured image therefore this image is also not usable.

### 5.2.2 Solution 2

The first solution, though mechanically preferred, had optical limitations due to size limitations in the vertical. When the height of a product was to large to place two lenses between the camera aperture and the reference image, the second solution is preferred. By placing a plane mirror directing the reference image 90°, the working distance is increased, which allowed cameras with a longer minimum focusing distance to preform better. The number of lenses used was dependent on the angular field of view. If the camera product used a fish eye lens, the image needed to be magnified, which meant that only a positive lenses was used, the result is seen in figure 5.14. Otherwise two lenses placed in a reversed Galilean formation was used, which made it simple to dictate how much the image was reduced, see 5.15. By placing the lenses so their focal points coincided, the demagnification could be calculated using equation 2.6. However, this solution did not preform as well for camera's with a larger field of view than 120°. This was due to the fact that even with a magnified image the region of interest was small in comparison to the field of view.

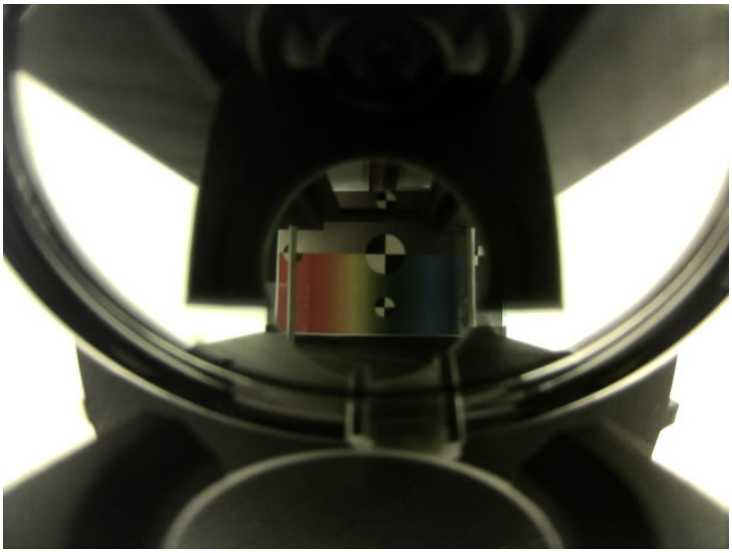


Figure 5.14: A m12 lens with a 140° angular field of view and a minimum focus distance of approximately 30 cm. One external lens was used with the focal length of 200 mm. In this figure the reference image is only a small part of the captured image.

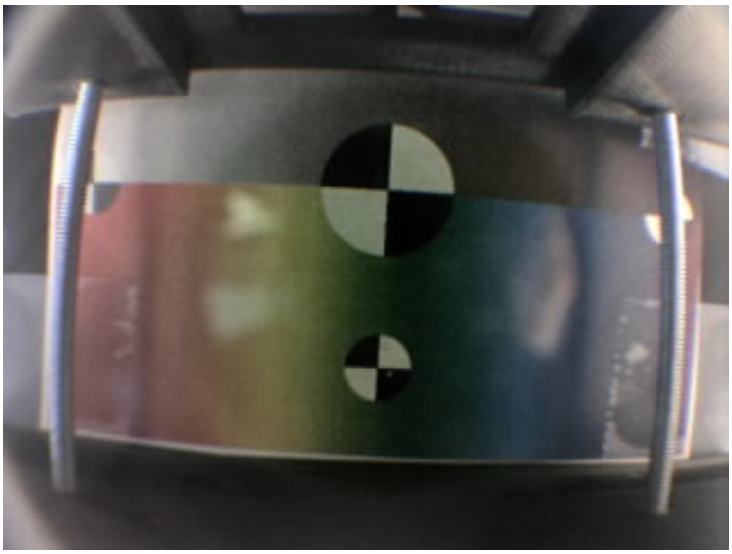


Figure 5.15: A m12 lens with a 26° angular field of view and a minimum focus distance of approximately 1m. Two external lenses used in a reverse galiliean formation. The negative lens had a focal length of -25 mm and the positive lens had a focal length of 60 mm. The entire reference image is present in the captured image, as well as a clear center of focus. Which means that it is usable for mikrodust.

| figure | Working distance | L1 | L2 | L3 | focal length |
|---|---|---|---|---|---|
| 5.14 | 140 mm | 135 mm | 5 mm | - | 1.56 mm |
| 5.15 | 65 mm | 25 mm | 5 mm | 47.5 mm | 8 mm |
| 5.16 | 92.5 mm | 75 mm | 5 mm | 37.5 mm | 3.6 mm |

Table 5.3: Table of values for the three tests. L1 is the distance between the reference image and the external lens, L2 is the distance between the external lens and the m12 lens, L3 is the distance between the two external lenses if two lenses were used and focal length is the focal length of the M12 lens.

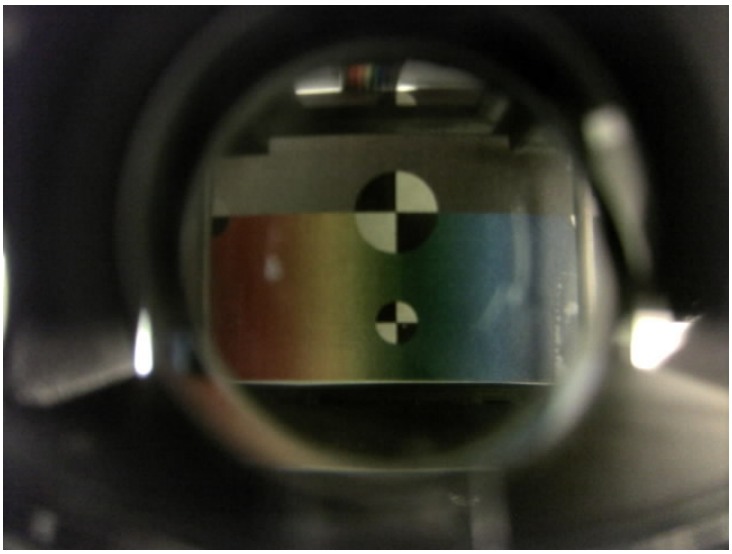


Figure 5.16: A m12 lens with a 67° angular field of view and a minimum focus distance of approximately 1m. Two external lenses used in a reverse galiliean formation. The negative lens had a focal length of -75 mm and the positive lens had a focal length of 50 mm. The entire reference image is present in the captured image, as well as a clear center of focus. Which means that it is usable for mikrodust.

# Chapter 6

# Conclusion

In this thesis a two different cases for camera supervision in the mikrodust test-fixture was investigated. For the case when the camera was used for surveillance of products in testing it was discovered that there was no need for any further development. It became apparent that cheap m12 lenses with varied angular field of view solves any issues mikrodust has, by simple calculations a appropriate m12 lens could be estimated. Instead image processing algorithms for removing blur and distortion was investigated. Two different common deblurring algorithms were used. By both image quality and execution time Wiener deconvolution performed the best. However, none of the algorithms performed well on real blur, which indicates that the way the blur kernel was estimated was inaccurate to some extent.

For the case when the camera was in the product being tested two different solutions was investigated. The best solution was to place a plane mirror to increase the working distance, which was limited by the height of the top box. The size of the image and the focus distance could then be decided by an external lens or a pair of lenses. By being able to capture a clear image with a camera that had a focusing distance of infinity this solution should work for most camera products tested in the test-fixture. The only exception when the solution preforms slightly worse is if the camera product uses fish eye lens. The performance of fisheye lenses could probably be improved, if a smaller mirror is used.

# Bibliography


[1] Amir Beck and Marc Teboulle. "A Fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse Problems". In: *SIAM J. IMAGING SCIENCES* 2.1 (2009), pp. 183–202.

[2] Jonathan M. Blackledge. "Chapter 2 - 2D Fourier Theory". In: *Digital Image Processing*. Ed. by Jonathan M. Blackledge. Woodhead Publishing Series in Electronic and Optical Materials. Woodhead Publishing, 2005, pp. 30–49. ISBN: 978-1-898563-49-5. DOI: https://doi.org/10.1533/9780857099464.1.30. URL: https://www.sciencedirect.com/science/article/pii/B9781898563495500021.

[3] Eustace L. Dereniak and Teresa D. Dereniak. *Geometrical and Trigonometric Optics*. Cambridge University Press, 2008. DOI: 10.1017/CBO9780511755637.

[4] G. Dougherty and Z. Kawaf. "The point spread function revisited: image restoration using 2-D deconvolution". In: *Radiography* 7.4 (2001), pp. 255–262. ISSN: 1078-8174. DOI: https://doi.org/10.1053/radi.2001.0341. URL: https://www.sciencedirect.com/science/article/pii/S1078817401903414.

[5] Dong Gong et al. "Learning Deep Gradient Descent Optimization for Image Deconvolution". In: *IEEE Transactions on Neural Networks and Learning Systems* 31.12 (2020), pp. 5468–5482. DOI: 10.1109/TNNLS.2020.2968289.

[6] Renzhi He et al. *Point Spread Function Estimation of Defocus*. 2022. arXiv: 2203.02953 [cs.CV].

[7] Carl Olsson. "Lecture notes in Computer Vision". In: *LTH* (2022).

[8] OpenCV. *Out of focus deblur filter*. URL: https://docs.opencv.org/4.x/de/d3c/tutorial_out_of_focus_deblur_filter.html.

[9] Frank L. Pedrotti, Leno M. Pedrotti, and Leno S. Pedrotti. *Introduction to Optics*. 3rd ed. Cambridge University Press, 2017. DOI: 10.1017/9781108552493.

[10] Said Pertuz, Domenec Puig, and Miguel Angel Garcia. "Analysis of focus measure operators for shape-from-focus". In: *Pattern Recognition* 46.5 (2013), pp. 1415–1432. ISSN: 0031-3203. DOI: https://doi.org/10.1016/j.patcog.2012.11.011. URL: https://www.sciencedirect.com/science/article/pii/S0031320312004736.

[11] Stephen Kyle Thomas Luhmann Stuart Robson and Jan Boehm. *Close-Range Photogrammetry and 3D Imaging, 3rd Edition*. De Gruyter, 2020.

[12] Wikipedia. *Point spread function*. URL: https://en.wikipedia.org/wiki/Point_spread_function.

[13] Sudha Yadav, Charu Jain, and Aarti Chugh. "Evaluation of image deblurring techniques". In: *International Journal of Computer Applications* 139.12 (2016), pp. 32–36.

# Appendices

# Appendix A

# More images - Case II solution 2

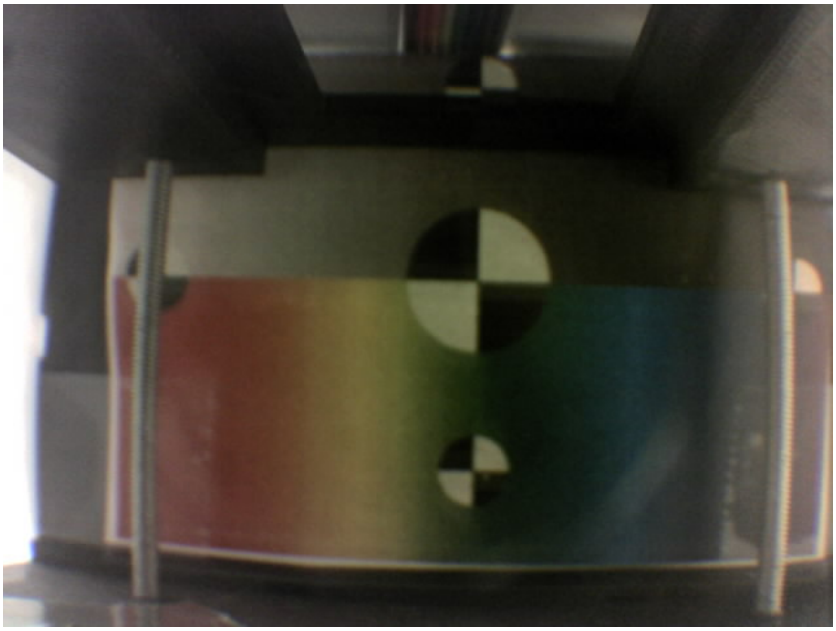


Figure A.1: Solution 2: A m12 lens with a 26° angular field of view and 8mm focal length. Two lenses used in a reverse Galilean formation with focal lengths of -25 mm and 50 mm. The distance between the reference image and the first lens was 25 mm, the distance between the two lenses was 37.5 mm and the distance between the second lens and the camera was 1 mm.

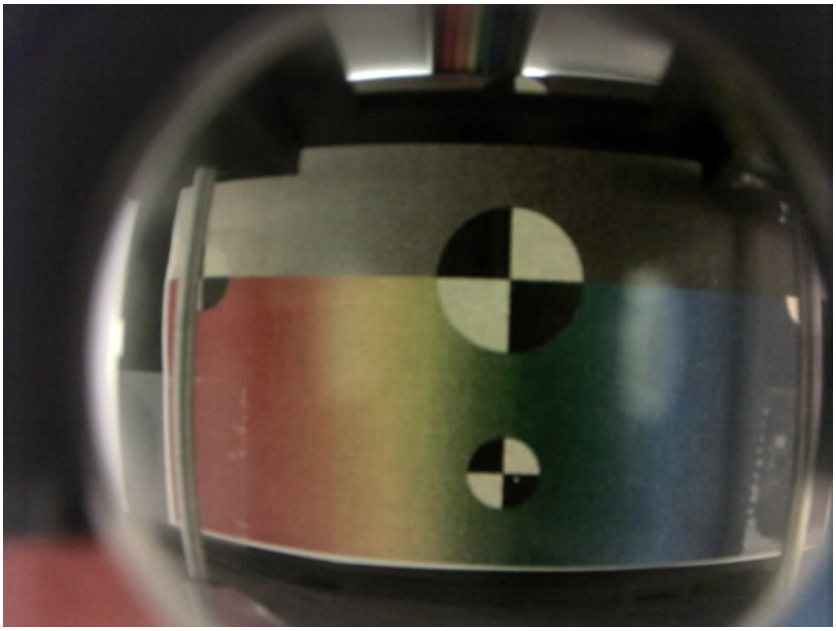


Figure A.2: Solution 2: A m12 lens with a 33° angular field of view and 6 mm focal length. Two lenses used in a reverse Galilean formation with focal lengths of -25 mm and 50 mm. The distance between the reference image and the first lens was 25 mm, the distance between the two lenses was 37.5 mm and the distance between the second lens and the camera was 1 mm.

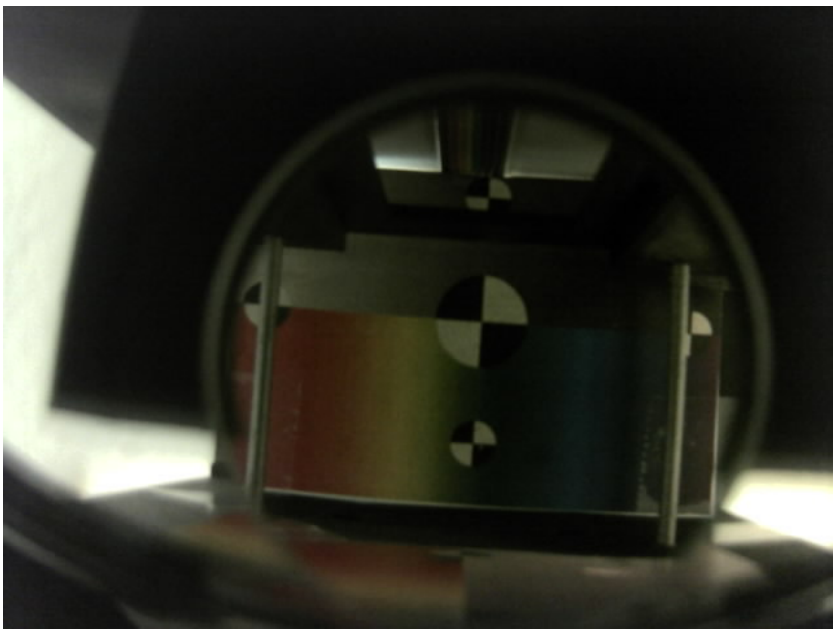


Figure A.3: Solution 2: A m12 lens with a 100° angular field of view and 1.8 mm focal length. A positive lens was used with a focal length of 100 mm was used. The distance between the lens and the reference image was 150 mm and the distance between the lens and the camera was 20 mm.

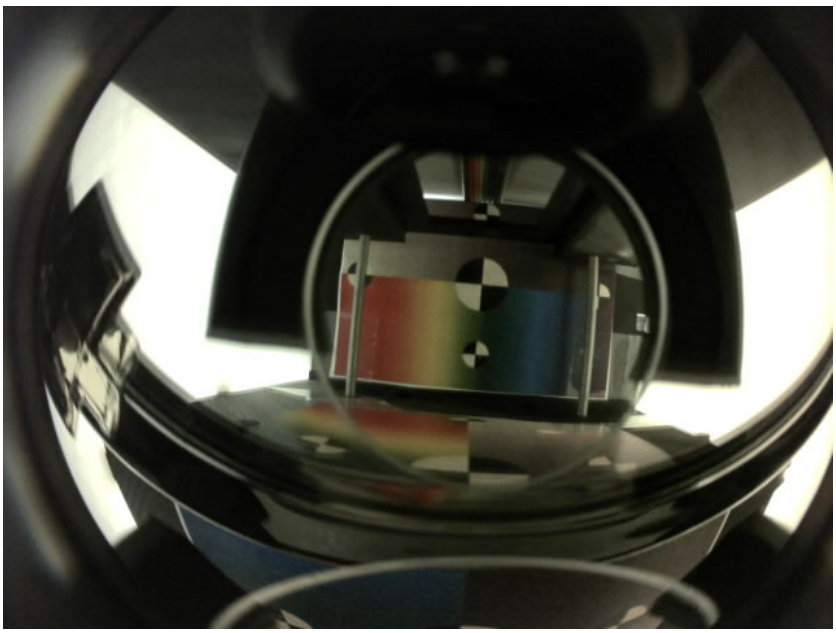


Figure A.4: Solution 2: A m12 lens with a 160° angular field of view and 1.71 mm focal length. A positive lens was used with a focal length of 400 mm was used. The distance between the lens and the reference image was 150 mm and the distance between the lens and the camera was 20 mm.