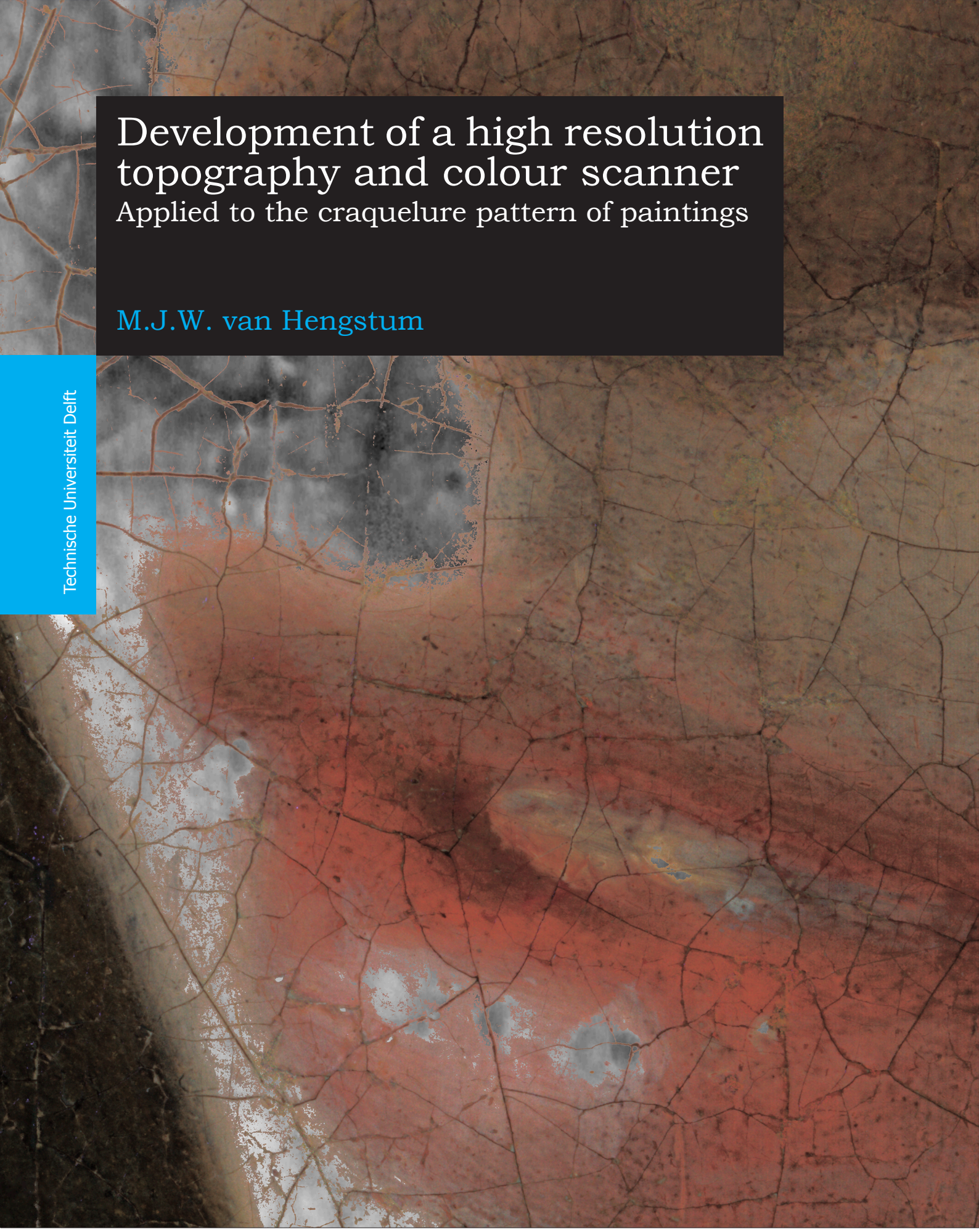


# Development of a high resolution topography and colour scanner

## Applied to the craquelure pattern of paintings

M.J.W. van Hengstum

Technische Universiteit Delft

TU Delft Delft University of Technology

Challenge the future

# Development of a high resolution topography and colour scanner

## Applied to the craquelure pattern of paintings

by

**M.J.W. van Hengstum**

in partial fulfillment of the requirements for the degree of

**Master of Science**
in Mechanical Engineering

at the Delft University of Technology
to be defended publicly on Tuesday July 10, 2018 at 15:00 PM.

Supervisors: Prof. dr. ir. J.M.P. Geraedts
Dr. Y. Song
Ir. T.T.W. Essers
Prof. dr. J. Dik
Dr. D. Dodou



TU Delft Delft University of Technology

# Preface

This Master Thesis is carried out as part of the Master thesis graduation requirements of Bio-Mechanical Engineering at the Technical University of Delft. This thesis consists of a literature study and the graduation, which took about 4 and 6 months, from September 2017 to the beginning of July 2018 to complete. The project was initiated by Prof.Dr. Joris Dik in September 2017 with the idea to develop a 3D high resolution scanner that could capture the colour and topography of a painting's surface.

The subject of this Master thesis was formulated in collaboration with Prof.Dr.Ir. Jo Geraedts from the faculty Industrial Design Engineering at the Technical University of Delft. The faculty was able to provide essential hardware, from a previous project with Océ, a canon company, to develop the 3D high resolution scanner. Moreover, I would like to thank Prof.Dr.Ir. Jo Geraedts, Dr.ir Y. Song and Ir. T.T.W. Essers for their weekly input and valuable feedback.

At the beginning of this thesis it became known that the Mauritshuis would initiate a thorough research project on their famous painting "The girl with the pearl earring" in March 2018. The Mauritshuis made it possible to use the scanning of the painting as a case study. I would like to thank the Mauritshuis for giving us the opportunity to scan "The girl with the pearl earring" and Océ a canon company for providing the required hardware for the high resolution scanner.

*M.J.W. van Hengstum*
*Delft*

# Abstract


## Paper

We present a high resolution topography and colour scanner applied to the craquelure pattern of oil paintings. Craquelure is a result of a highly complex process of ageing and environmental influences and is therefore often used in research as a possibility to visualise and monitor the quality of a painting. Historians and restorers are highly interested in using digitisation techniques to analyse craquelure patterns in oil paintings to determine the painting's quality. The scanner that we propose simultaneously captures the colour and topography information of the oil painting with fringe encoded stereo imaging. The scanner is capable of capturing a painting with a spatial resolution of 7 µm and a depth precision of 34 µm. Moreover, the scanner is able to capture a painting of 1750 $cm^2$ within 2 hours with a tile overlap of 25%. In addition, we were given the opportunity to scan 'The girl with the pearl earring' to capture the craquelure pattern. Preliminary analysis of the results show that the scanner can digitise the craquelure in three-dimension, contributing to the exact measurement research on craquelure patterns.

## Literature study

The craquelure pattern is one of the main factors in determining the state of a painting. The state-of-the-art devices either lack the ultra-high resolution or manoeuvrability to capture the craquelure pattern of the entire painting. Few devices can accurately capture near-planar surface in 3D ultra-high resolution, e.g., 2500 dpi, for the entire surface of paintings.

The literature study aims to analyse the current colour and topography capturing techniques that could accurately capture the craquelure pattern. In addition, the associating issues of the technique are examined. The study begins with a holistic view of techniques that can be used in capturing colour and topographical information of the painting. The first step of the literature study concludes that the Fringe encoded Stereo Imaging holds the most potential for painting scanning. However, the current resolution is insufficient to capture craquelure. Further literature research indicated that fringe encoded stereo imaging has multiple camera lens solutions to increase the resolution, from which extension tubes is the most suited solution, from both theoretical and practical points of view. Increasing resolution is always associated with a smaller Depth of Field, therefore the position and orientation of the cameras have to be accurately positioned regarding the painting on the easel. Thus, another literature study on techniques to accurately estimate the positioning of the device relative to the painting is conducted. Sparse stereo image matching is the most promising technique both in simplicity and accuracy.

This literature concludes that it is feasible to accurately scan the craquelure pattern of the entire painting with enhanced fringe encoded stereo imaging and accurate positioning.

## Instruction manual FAST

The FAST scanner focuses on ultra-high resolution colour and topography of the paintings, to digitally visualise the craquelure pattern of painting. The instruction manual and some of the software were based or copied from the Fine Art Scanner, which focuses at scanning the visual appearances (colour, texture and gloss) of a painting and reproducing it with 3D printing. The FAST scanner does not scan gloss, but there might be some software titles in the manual that still contain gloss, which is not part of the actual software.

The manual explains in detail the work flow of the FAST scanning process. In seven chapter the entire scanning for capturing a painting is explained. This manual includes an introduction to the different processes, system requirements, the assembly of the scanner, acquiring data and image processing.

# Contents

# I

# Paper

# Development of a high resolution topography and colour scanner, applied to the craquelure pattern of paintings

M.J.W. van Hengstum, J.M.P. Geraedts, T.T.W. Essers, D. Dodou, Y. Song, and J. Dik

**Abstract**—We present a high resolution topography and colour scanner applied to the craquelure pattern of oil paintings. Craquelure is a results of a complex process of ageing and environmental influences and is therefore often used in research as a possibility to visualise and monitor the quality of a painting. Historians and restorers are highly interested in using digitisation techniques to analyse craquelure patterns in oil paintings to determine the painting's quality. The scanner that we propose simultaneously captures the colour and topography information of the oil painting with fringe encoded stereo imaging. The scanner is capable of capturing a painting with a spatial resolution of 7 $\mu$m and a depth precision of 34 $\mu$m. Moreover, the scanner is able to capture a painting of 1750 cm$^2$ within 2 hours with a tile overlap of 25%. In addition, we were given the opportunity to scan 'The girl with the pearl earring' to capture the craquelure pattern. Preliminary analysis of the results show that the scanner can digitise the craquelure in three-dimensions, contributing to the exact measurement research on craquelure patterns.

**Index Terms**—3D imaging, fringe encoded stereo imaging, craquelure, girl with a pearl earring, high resolution

## 1 Introduction

The advances in 3D scanning technologies have resulted in a greater understanding of oil paintings. For example, 3D scanning and printing techniques have enabled museums to preserve and provide physical access to 3D printed reproductions of famous oil paintings [1] [2]. Furthermore, digital 3D confocal microscopy has been applied to scan small patches of a painting. The digital microscope assist art conservationists with microscopical resolution 3D images of textured samples, such as visualising unmounted paint flakes or by revealing the boundaries between varnish and paint with UV light [3] [4].

This paper investigates the feasibility of a high resolution topography and colour scanner for oil paintings. High resolution topography and colour imaging enables capturing of the fine brush strokes, which describes the colour usage, style and work flow of the painter [5]. Apart from the brush strokes, the high resolution also enables the registration of the craquelure pattern of oil paintings.

### 1.1 Topography: Craquelure

Many paintings suffer from fractures in the paint, or varnish. These fracture patterns, called cracks, crackle, or craquelure, can be caused by ageing, drying, and mechanical factors. The craquelure pattern is one of the main factors in determining the quality of a painting and is a broad indicator of authorship [6] [7]. Moreover, delamination of the paint from the canvas often takes place at the edge of craquelure [8]. Therefore, being able to monitor the craquelure of the entire painting is a valuable tool for conservators and restorers to determine the quality of the painting or possible preserving procedure [9] [10].

The state of a painting is also used to determine the response of the painting to environmental influences, like the humidity, light and temperature [11] [12] [13]. An accurate estimation of the craquelure would aid to objectively research these environmental influences on the degradation of a painting. Moreover, the finer craquelure is of great interest with respect to the detection of transportation damages [14] [15]. For example, a proprietor of a painting could register the craquelure of the painting to safeguard the quality of the painting. Even though craquelure is a useful property of the painting, there is little literature on the exact dimensions of the cracks. The craquelure literature is primarily about the classification on type of craquelure and the type of patterns they create [16].

### 1.2 Colour

The colour of a painting changes over time, depending on the type of paint pigment [17] and environmental influences [18] like microbes [19], light [20], temperature or humidity [21]. Accurately capturing the colour of a painting enables conservationists to further study the environmental impact on specific paint pigments. Furthermore, combining the colour with the topography information provides a more complete understanding of the features in the paint [22]. For example, a sudden increase of paint height can easily be clarified to be a drub of paint or excess dirt with the colour information. The relation between topography and colour assists restorers and conservationists in understanding, visualising and monitoring features in the painting [23].

Current state-of-the-art devices either lack the scanning range or the resolution to capture the craquelure pattern of the entire painting. In this paper, a novel high resolution topography and colour scanner capable for capturing the craquelure pattern of the entire painting is introduced.

### 1.3 Design requirements: high resolution scanner

As mentioned, there is a sparsity in the literature on exact dimensions of craquelure. Therefore, most concrete information on craquelure was retrieved from research performed on 'The girl with a pearl earring' in 1994 [24]. Using high-resolution images it was estimated that the width of craquelure varied from 50 to 100 µm, and the paint layer thickness from 25 to 300 µm based on paint samples of the research. The profile of ageing cracks, observed from an intersection of the paint layer, often approximate a rectangular shape [25].

To determine the spatial resolution and depth precision required to register the finest craquelure, signal analysis theorem was applied. To approximate a rectangular shape in signal analysis, the sampling frequency has to be minimally 5 times higher than the signal frequency [26] [27]. Since the width or wavelength of the signal is the inverse of the frequency, the width of the sampling signal has to be 5 times smaller than the width of the craquelure to register the shape. Applying this to describing the finest craquelure of 50 µm wide would require a spatial resolution of 10 µm (XY). Similarly, a depth precision of 5 µm is required to describe the depth of the shallowest crack.

These requirements are easily met with the capabilities of 3D microscopes [28] and just beyond the capabilities of most full painting scanners based on digital photography, which can achieve a spatial resolution of 25 µm [29].

The available time of museums is often the limiting factor in scanning entire paintings. In ideal cases a working day of 8 hours is allocated to scan the art work. High resolution scanning is generally accompanied with long acquisition time since more information has to be captured for the same area. Therefore, the acquisition time should be reduced, while preserving a high resolution. The design requirements of the high resolution scanner are set to scan the painting within the 8 hours of a working day. The acquisition time is difficult to determine since it varies with the painting size, the type of device, the resolution and device settings. In general, the high resolution scanners, like 3D microscopes, require multiple days to scan an entire painting. Techniques based on digital photography usually have a lower resolution and are usually capable of scanning a painting within the 8 hours [29], however this varies per device.

Additionally, development and purchase are often limited by the price of the device. In summary, the aim is to develop a quick 3D high resolution scanner capable of capturing the craquelure pattern of paintings, while minimising the costs.

Finally, as mentioned, it is important to reproduce the colours accurately to analyse colour degradation. Guidelines set by the Federal Agencies Digital Guidelines Initiative (FADGI) state that an accurate colour reproduction is achieved when the average colour accuracy is lower than 3 ΔE-2000 [30].

The design requirements for a scanner capable of capturing the craquelure pattern of 'The girl with a pearl earring' are summarised in table 1.

TABLE 1: Design requirements for an economical and quick high resolution topography and colour scanner

| | |
|---|---|
| Spatial resolution | <10 µm |
| Depth precision | <5 µm |
| Acquisition Time | <8 hours |
| Colour Accuracy | <3 ΔE-2000 |

### 1.4 Case: 'The girl with a pearl earring'

'The girl with a pearl earring' was painted by Johannes Vermeer in 1667. The painting has an area of about 1750 $cm^2$ (44.5x39 cm) and the craquelure is very fine and rectangular shaped as a result of ageing [24]. The fine cracks make the painting a suitable subject to test the capability to capture craquelure for the entire painting.

The Mauritshuis in the Hague initiated a research project on their world famous painting 'The girl with a pearl earring'. Several types of research were carried out from February 26 to March 11, 2018. The research was conducted in a glass enclosure, specially built to display the research being done on the painting, while the public was visiting the museum. The opportunity arose to scan 'The girl with a pearl earring' from 12 pm to 18 pm, on March 3, 2018. The research project gained broad media attention [31] with, in some cases, our scanner as centre point of attention [32].

## 2 Literature

Literature shows several state-of-the-art colour and topography capturing techniques with the potential to scan the craquelure pattern of the entire painting. The four techniques with the highest potential are discussed below. Since there is no commercial image sensor that can capture the entire painting at the required resolution, the painting is divided in multiple smaller tiles to reach its maximum resolution. The tiles are stitched together to digitally construct an image of the entire painting.

Digital 3D confocal microscopy can achieve a sub-micron resolution, but it requires the depth of field to be as small as possible to determine the depth accurately [28]. The depth of field is the area parallel to the capturing device that is perceived to be in focus. A small depth of field is achieved by increasing the resolution, which increases the acquisition time [33]. Moreover, until March this year (2018) there was no confocal microscope with the range to scan the entire painting [4]. Apart from a long acquisition time and limited range, a confocal microscopy is an expensive technique.

Photogrammetry is widely used to determine the dimensions of 3D objects due to its user-friendliness [34]. The limitation is its dependence on having to match areas to determine its position and orientation relative to the previous location and then also reconstruct the 3D image on these matched areas. Areas that do not have sufficient features to match, like the background in a painting, can not be reconstructed [35]. The technique is economical since it only requires one camera, but the depth precision is limited to 50 µm with the current technology [36].

Photometric stereo imaging estimates the depth from the shadows and highlights created by varying the direction of the indicent lights [37]. An algorithm extracts the shade and highlights from the multiple images and estimates the depth

from the shape. The angle of the indicent light determines the trade-off between the depth precision and the depth estimation of craquelure, since larger shadows increase the depth precision, but decrease the ability to register in the cracks [38]. The primary drawback is the depth precision, since the shadow and highlights are difficult to extract and do not always describe the topography of the painting accurately [39]. Photometric stereo imaging needs a camera and several light sources, making it an economical technique, but the depth precision is limited.

Fringe encoded stereo imaging determines the depth by using triangulation with two cameras. Fringe projection uniquely labels each position on the painting to aid the triangulation and solve the correspondence problem in photogrammetry. In contrast to confocal microscopy, the technique requires a large depth of field to determine depth, reducing the required spatial resolution and thus the acquisition time. Fringe encoded stereo imaging requires two cameras and is therefore less economical than other digital camera techniques, but still affordable. Scanners based on this technology are primarily used to produce a 3D printed reproduction of a painting. Reproducing a painting, perceptually equal for humans to the real painting, requires a spatial resolution of $\pm 50$ µm [40]. Current scanners therefore have specifications of a spatial resolution of 25 µm, a depth precision of 10 µm and acquisition time of 4 minute per tile [41], which do not meet the requirements to scan craquelure. However, the technique has more potential when changing the hardware, which was provided by Océ a Canon company. These low initial costs in combination with the low acquisition time and high potential resolution is Fringe encoded stereo imaging the best technique for our scanner.

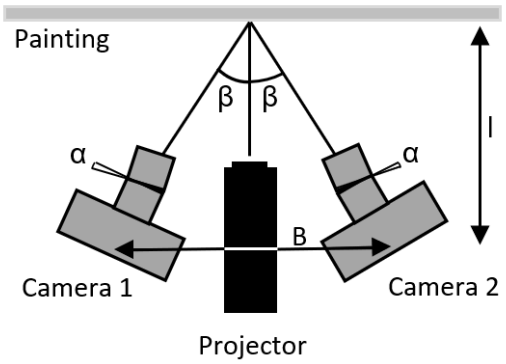



Fig. 1: The setup of fringe encoded stereo imaging as seen from above. The cameras are place 200 mm (l) from the painting with 140 mm (B) spacing between the cameras. The angle of the cameras relative to the painting's normal vector $\beta$ is 21.5°. The tilt-shift lens is tilted with the angle $\alpha$ of 8° to focus on the painting's surface.

## 3 Materials and Methods

Scanning devices based on fringe encoded stereo imaging devices currently lack the resolution required to capture the craquelure pattern. To meet the requirements, changes were made at both the hardware and the software. The changes in the hardware were focused at increasing the spatial resolution. The depth precision is disregarded since it changes proportionally with the spatial resolution when the camera angles relative to the painting's normal vector remain equal, based on the geometry in stereo imaging. The existing software was adjusted to reduce the acquisition time, keeping in mind that the acquisition time scales two-dimensionally with the resolution.

### 3.1 Hardware

The hardware consists of two cameras, a projector and a linear XY-stage that translates the device parallel to the painting. Each component of the hardware of the original design [2] was changed in order to primarily enhance the spatial resolution, but also lower the acquisition time. The hardware changes are summarised in table 2.

#### *Camera and Lens*

The key specifications of the camera and lens are the amount of pixels of the camera sensor, the pixel pitch, and the magnification of the lens. An increase in amount of pixels increases the tile size, which determines the number of tiles required to scan the painting and thus the acquisition time. The pixel pitch is the distance from pixel centre to its neighbour's centre and thus proportional to the spatial resolution. The magnification of the image is primarily determined by the camera lens. Moreover, the camera angle relative to the painting's normal vector ($\beta$) and the spatial resolution determine the depth precision. Increasing the camera angle enhances the depth precision, but it makes the scanner more prone to occlusions.

The two Nikon E800D cameras and Nikkor PC-E 85mm lenses were replaced with two Canon EOS 5Ds cameras combined with a 90mm TS-E f/2.8 tilt-shift lens, which lowers the pitch from 4.87 µm to 4.1 µm and improves the total amount of pixels from 36.3 megapixels to 50.6 megapixels, achieving a higher spatial resolution of 15.4 µm (-38%) and reducing the acquisition time respectively (-38%). In addition, a 25 mm extension tube is placed between the camera and the lens to theoretically achieve a spatial resolution of 7 µm. An extension tube magnifies the image by elongating the divergence of the light without compromising on the quality, but the longer divergence results in less light striking the sensor. A longer exposure time is required to compensate the loss of light, increasing the acquisition time (+20%) compared to no extension tube. Moreover, the enhanced spatial resolution reduces the tile size, also called the camera window, from 170x100 mm to 60x40 mm per tile and reduces the shared depth of field of the cameras. A smaller shared depth of field increases the difficulty to get the entire surface of the painting in focus for the cameras. Furthermore, the camera angle is decreased from 40° to 21.5° to decrease the susceptibility to occlusions, worsening the depth precision from originally 10 µm to 17.5 µm as determined by geometry of stereo imaging[1]. However, the depth precision enhances proportionally with the spatial resolution with a factor 0.28 [2] from the 17.5 µm to about 5 µm.

1. $(10 \cdot \sin(40))/\sin(21.5)$
2. 7 µm/25 µm

The magnification also results in a smaller camera working distance (l) compared to the original hardware, which means that the cameras have to be placed closer to each other, to keep the relative angles to the painting's normal vector ($\beta$) equal. The space between the cameras (B), where the projector is positioned, becomes smaller.

### *Projector*

The important specifications of a projector are the dimensions, intensity and throw distance for the design requirements. The projector should fit between the cameras in order to perpendicularly project the fringe pattern, which limit the shape distortions from the projector lens and also reduce the shading on the painting's surface. The distance between the cameras is about 140 mm for our hardware, but depends on the camera dimensions, the working distance and the camera angle. A high intensity projector reduces the influence of the environmental lighting and more importantly reduces the required camera exposure time and thus the acquisition time. The working distance depicts the required throw distance of the projector, which is about 200 mm for the selected hardware.

The Acer X113H 2800 Lumens Halogen projector was replaced with the AXAA M6 1200 Lumens LED pico projector, to meet the short throw distance and the dimensions. As the name suggests the pico projector is smaller, but it still has to be rotated 90° to fit between the cameras.The intensity of the projector was sacrificed to fit the projector between the cameras and be able to focus at the working distance, increasing the minimal shutterspeed from 0.5s to 0.6s and thus the minimal acquisition time (+20%). Moreover, polarisation filters are applied to the cameras and projector, which is called cross polarisation to remove any specularities introduced by the projector or the painting's surface.

### *XY-stage*

The essential specifications of a XY-stage are precision, range, rigidity and portability. The XY-stage needs to precisely translate the scanner parallel to the surface of the painting, to ensure the overlap is consistent. The required precision depends on the tile size and the minimal overlap. The selected hardware results in a tile size of 60x40 mm and a minimal overlap of 20% (8 mm) to stitch the images together. This results in a minimal precision of 0.8 mm to position the scanner with sufficient certainty. The range of the XY-stage should enable the scanner to capture the surface of most paintings, which is a range of about 1-by-1 m for 90% of paintings [42]. The stage also has to be rigid enough to minimise vibrations, while being portable to transport to any museum. The vibrations influence the quality and the consistency of the image. A rigid XY-stage is therefore essential for accurate topography and colour reproduction.

We used a high precision and rigid XY-stage developed by Océ a Canon company, seen in Fig. 2, which can be fully disassembled. Two linear drivers move the device vertically and one linear driver horizontally for a range of 1 by 1 metre with an effective precision of 0.01 mm. Moreover, the XY-stage is designed with a wide base to enhance the stability to minimise vibrations.

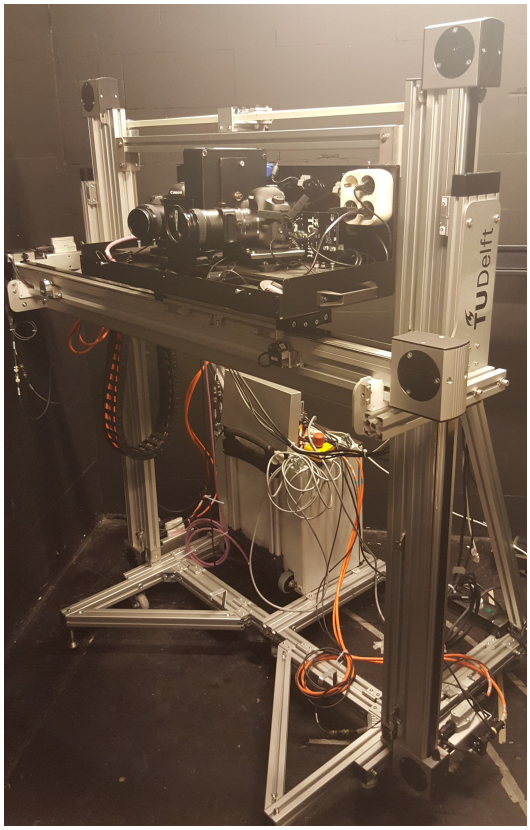



Fig. 2: The XY-stage translates the two cameras and projector parallel to the painting. Two vertical linear drivers and a horizontal linear driver provide the precise movement, while minimising the vibrations with a rigid structure.

The specifications of the original and selected hardware are summarised in table 2. The hardware changes result in an estimated spatial resolution and depth precision of 7 µm and 5 µm respectively, which is an increase of a factor 3.5 compared to the spatial resolution of the original design. As mentioned, the acquisition time scales two-dimensionally with the resolution. Therefore, the acquisition time increases with a factor 12 ($3.5^2$). When taking the extension tube, pico projector and camera sensor into account the acquisition time is increased with a total factor of around 13. These estimates are validated in practise to determine whether the requirements are met, which is further explained in validation methods.

TABLE 2: Comparison of the hardware differences between the original hardware [41] and the hardware of our scanner.

| | Original hardware | Our hardware |
|---|---|---|
| Camera | Nikon E800D | Canon EOS 5Ds |
| Cam Resolution | 36.3 MP | 50.3 MP |
| Pixel pitch | 4.87 µm | 4.14 µm |
| Cam. Angle | 40 ° | 21.5 ° |
| Camera Lens | Nikkor PC-E 85 mm | Canon TS-E 90 mm |
| Spatial Resolution | 25 µm | 7 µm |
| Depth Precision | 10µm | 5µm |
| Tile size | 170x100 mm | 60x40 mm |
| Projector | Acer X113H | AXAA M6 |
| Intensity | 2800 Lumens | 1200 Lumens |
| Resolution | 800x600 px | 1920x1080 px |
| xy-stage precision | ±1 mm | 0.01 mm |
| Automated Range | 1.3x1.3 m | 1x1 m |

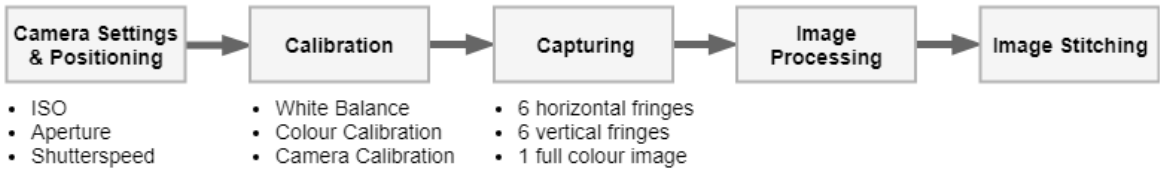



Fig. 3: The work flow of our scanner when completely assembled. The camera settings are set and positioning is done at the beginning. Next, the white balance is customised and the camera and colours are calibrated. The painting will be captured with 13 images per tile. After the scan the images are processed for each tile and finally stitched together into one image.

### 3.2 Work flow

Fig. 3 shows the overview of the entire work flow of scanning a painting when the device is assembled. The work flow starts with the camera settings, which determine the quality and salience of the image. The aperture is set quite small at f/16 to increase the depth of field, while minimising the diffraction. The light sensitivity of the camera sensor, ISO, introduces noise when the sensitivity is increased. The ISO is therefore kept at the lowest value of 100 to minimise sensor noise, requiring a longer exposure time to compensate. The shutterspeed varies from 0.6s to 3.2s, depending on the intensity of the painting and the available scanning time. A lower intensity painting requires a longer shutterspeed to capture a salient image.

When the resolution is increased the depth of field will reduce, making it more difficult to focus the image of both cameras for the entire painting [43]. The XY-stage has to be positioned nearly parallel to the painting at the working distance of the cameras and camera lens, which is 200 mm ($\pm$ 1 mm) for the Canon EOS 5Ds with the 25mm extension tube and Canon TS-E 90mm lens .

After positioning the system, the cameras are calibrated to determine the optical ray map per camera. Calibration is done by having the cameras capture around 20 images of an orientation changing checkerboard with a known square size. Since the checkerboard is planar and lies in a fixed grid, the intrinsic and extrinsic parameters of the cameras can be determined. The intrinsic and extrinsic parameters generate the optical ray map for each camera [44], which relates the light ray that strikes a pixel back to the 3D position. It takes about 3 minutes to capture the camera calibration checkerboard in several orientations rotated about all axes. The calibration results are used to estimate the depth information from stereo imaging and to correct for any lens distortions. Usually, the projector's position and orientation are calibrated for fringe projection, however fringe encoded stereo imaging bases its depth estimation on stereo imaging, eliminating the need for projector calibration.

Customising the white balance and calibrating the colour is the last step before the scan. An image of a grey chart is used to customise the white balance with the built-in software of the cameras. The customised white balance correct the non-uniformity of the projector's illumination over the projection and the vignetting by the camera lenses. The colour is calibrated with a micro colour reference chart to automatically generate an ICC-profile for the cameras, with the projector as the illuminant. The ICC-profile corrects the colour information and converts the colour information to a general colour space.

The capturing sequence, also called scanning, consists of capturing the projected fringe patterns and moving to the next tile. The fringe projection is projected horizontally and vertically to label each area of the painting uniquely to help solve the correspondence problem in stereo imaging. The projected fringe pattern is a sinusoidal pattern in grey scale that shifts with an off-set related to the amount of phases. A fringe pattern with a wavelength of 12 pixels and 6 phases obtained a more consistent result compared to the original 3 phases [45], but also doubles the total images and the total acquisition time. Moreover, even the 6 phases fringe pattern showed that the fringe pattern could not fully be removed from the colour map. Therefore, the software was adjusted to capture an extra image without the fringe pattern to determine the colour map, increasing the acquisition time (+8%).

To compensate the added acquisition time from the hardware changes and the extra image, the capturing sequence was optimised. The images were stored off-line instead of real-time image downloading, reducing the acquisition time with a factor of about 6, but eliminating the possibility to monitor the image quality during the scanning sequence, the capturing process therefore becomes 'blind'. The exact acquisition time is validated during a pilot study explained in validation methods.

After the fringes have been captured, they are wrapped into a wrapped phase map. The phase map is unwrapped to generate the continuous phase map, uniquely labelling each position on the image. To correlate the cameras, about 1500 features from both unwrapped maps are matched with the sparse matching algorithm SIFT [46]. The correlated phase maps match each unique value of both cameras, enabling the computation of the 3D position of each point. The 3D position of each point can be computed through ray tracing by taking in account the information from our camera calibration. The construction of this 3D point cloud takes about 15-30 minutes per tile in our MATLAB implementation[3].

The increased high resolution introduces issues in stitching and visualising the entire painting, because of the sheer size of data. Therefore, image stitching and visualisation requires specialised software and hardware, which is outsourced to prof. dr. R.G. Erdmann of the Rijksmuseum [4].

3. Developed by T. Zaman at the Technical University of Delft
4. Prof. at Faculty of Social and Behavioural Sciences at University of Amsterdam

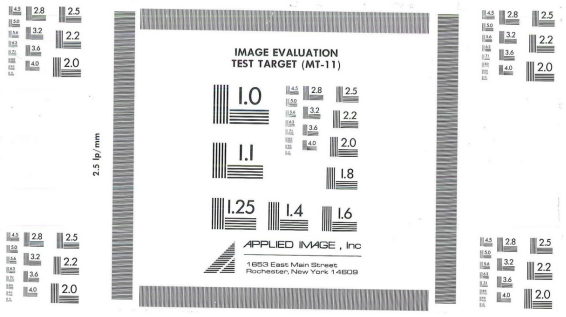



Fig. 4: The Applied Image Inc. MT-11 MTF chart contains several arrays and small patches of fine lines with varying line densities. The long 2.5 line/mm MTF bars positioned as a larger square, are used to determine the spatial resolution of the scanner for different parts of the processed image.

### 3.3 Validation Methods

The requirements for a 3D scanner able to capture the craquelure pattern of paintings are: spatial resolution, depth precision, colour accuracy and acquisition time, as summarised in table 1. Each of the four requirements is experimentally validated with a different method. Finally, the craquelure from 'The girl with a pearl earring' is analysed. The fringe pattern is kept equal for the consistency among all validation methods.

*Spatial Resolution*

The spatial resolution of any imaging system is defined as its ability to distinguish two points as separate in space. In fringe encoded stereo imaging, the spatial resolution is the distance that neighbouring pixels capture of the real world object. Commonly, to determine this real world distance, an object with known spacing is captured and processed, like a ruler. In this study a MTF chart was used to determine the spatial resolution, which is an accepted tool in digital photography [47]. A MTF chart contains several arrays of fine bars, which are typically used to determine the quality of the camera lens. The chart is an effective method to determine the spatial resolution, since the number of pixels describing the known amount of bars per millimetre determines the spatial resolution . Fig. 4 shows the Applied Image Inc. MT-11 MTF chart used to determine the spatial resolution. To determine the spatial resolution in horizontal (X) and vertical (Y) direction of the entire processed image, three horizontal and three vertical MTF bars with 2.5 lines per millimetre equally spaced over the tile were captured. The amount of pixels of the processed image describing the distance between the lines depicts the spatial resolution of the scanner and will be displayed in a boxplot.

*Depth Precision*

The depth precision is the ability to repeatably estimate the depth within a range of certainty. The measured depth of the painting is compared to the true depth information, to compute the depth precision. In stereo imaging, the depth precision depends on the spatial resolution and the camera angle ($\beta$). As mentioned, with the hardware of our scanner the estimated depth precision is about 5 µm, which is too small for the common depth calibration plates.


Fig. 5: The paint sample with the three areas that both scanners capture. The areas have a size of 6.5x4.8 mm. Area 1 is a bright part of the vase with very fine craquelure. Area 2 and 3 are the dark background with craquelure.

Therefore, the depth precision of the presented scanner is determined by comparing the depth measurements of our scanner to the more precise Bruker Nano Surfaces Scanner, which has a depth precision of 1 nm [48]. The large difference in depth precision makes the depth map from the Bruker nano scanner approximate the true depth value compared to our scanner. Both scanners capture the exact same three areas of a specially created section of a painting, shown in Fig. 5. This section was cut from an old painting with fine craquelure and stuck to a wooden panel to eliminate deformations of the canvas. When the section of the painting is scanned, the resulting depth maps of the three areas of our scanner are cropped to match the size of the Bruker's depth maps of 6.5x4.8 mm. The pixels not registered from both scanners are removed from the depth information. Next, the depth maps of these three areas are digitally translated and rotated in the optimal orientation to compute the difference in depth between the two depth maps. For each of the three areas the depth difference will be displayed in a boxplot, in which the standard deviation depicts half the depth precision.

*Colour Accuracy*

The colour accuracy is the performance of an imaging device to approach the true colour. The perception of colour depends on multiple factors, like the illuminant, gamma, and observer angle. To correct for these factors the colours are calibrated with a colour gauge chart. Due to the small camera window of the cameras a matte micro colour gauge target from $IS_A$ was used to calibrate the colour, as shown in Fig. 6. The calibration creates an ICC-profile by comparing the captured colour patches to the true colour value of the patches. The ICC-profile corrects the colour information and

Fig. 6: The micro colour gauge chart with 30 colour patches used to calibrate the colour information.

converts the colour information to a general colour space. The colour accuracy is measured with the standardised ΔE-2000 metric, which is the perceptual difference between two colours as observed by the human eye. The colour values are extracted from the colour corrected image of the micro colour gauge and compared to the true colour value to compute the ΔE-2000 for each colour patch [49]. The mean ΔE-2000 over all patches has to be lower than 3 to accurately reproduce the colour of the painting according to the FADGI guidelines [30].

### *Acquisition Time*

The total acquisition time depends on the amount of images per tile, the amount of tiles, the shutterspeed and the movement time between the tiles. The aim is to determine the average movement time between tiles and the average time to capture an image without the shutterspeed. To determine the movement time and time per image, the data of the two full scans of 'The girl with a pearl earring' are used. The first scan captured the girl with a pearl earring with a 45 % overlap and a shutterspeed of 0.6s, the second scan had a 25 % overlap and 3.2 second shutterspeed. The shutterspeed and overlap differ, but the number of images per tile remain equal at 13 images per tile.

The acquisition time is determined from the time stamp stored when the image is captured. The acquisition time per tile is computed by dividing the total acquisition with the number of images. The time per image is the average time between the first 13 images of each tile. Subtracting the average time per image from the time between the 13th image and 1st image of the next tile results in the movement time to the next tile.

### *Case: Craquelure Analysis*

'The girl with a pearl earring' is used as case study to determine the capability of our scanner to capture the craquelure of the entire painting. The topographical and colour maps are combined to analyse the profile of the craquelure. 3D intersections of the craquelure are extracted from the scanning data to determine the width and depth of craquelure but also to visualise the profile of frequently encountered cracks. The width and depth of craquelure are determined to validate whether the design requirements and current specifications are sufficient to accurately describe the craquelure pattern of a painting.

## 4 Validation Results

### *Spatial resolution*

Fig. 7 shows the spatial resolution for horizontal (X) and vertical (Y) MTF bars dispersed over the processed image. The

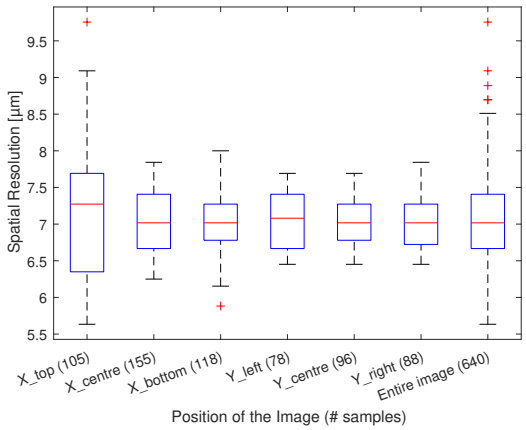



Fig. 7: The spatial resolution in horizontal (X) and vertical (Y) orientated MTF bars dispersed over the image.

spatial resolution in X and Y direction do not significantly differ. Apart from the horizontal spatial resolution at the top of the image, the spread of the spatial resolution is similar for all areas of the image. The average spatial resolution over all the areas of the image, shown on the right side of the figure, is 7.03 µm ($\sigma$ = 0.49, n = 640).

### *Depth precision*

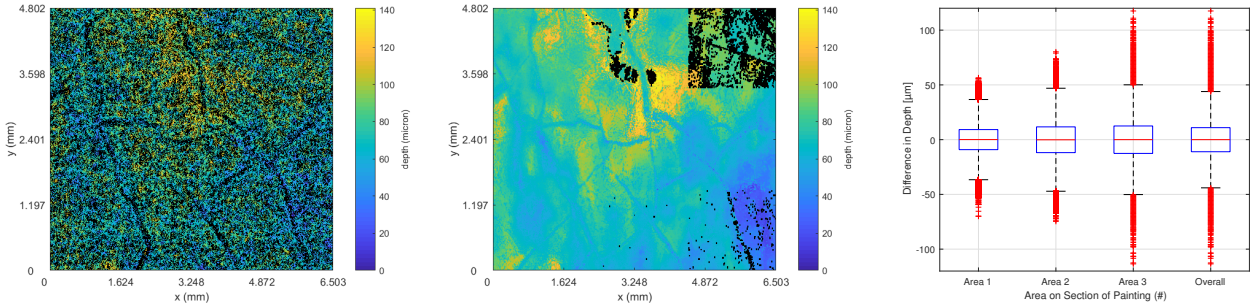
Fig. 8 shows examples of the depth maps of the presented scanner and the Bruker nano surface scanner of the same area of the sample of a painting. The height maps of both devices shown in Fig. 8a and Fig. 8b show the topography of the painting. Black depicts pixels from which the height was not determined. The depth map from our scanner is less complete than the Bruker scanner, especially the performance of the depth estimation of the craquelure is less. However, the depth information is sufficiently dense to interpolate a complete depth map. Even the Bruker nano surface scanner suffers from not determined depth information. The depth difference between our scanner and the Bruker is shown in Fig. 8c. The standard deviation differs per area from 13 µm to 20 µm and over all areas the standard deviation is 17 µm. The depth precision is double the standard deviation, thus the average depth precision is 34 µm of the presented technique.

### *Colour accuracy*

Table 3 shows the colour accuracy per colour patch in the ΔE-2000 metric for a shutterspeed of 3.2 seconds. Analysing the table using the colour gauge as shown in Fig. 6 reveals that the colours less accurately reproduced are white, light grey, dark grey and black. The average ΔE-2000 over the 30 colour patches is 1.54 ($\sigma$ = 2.32), which does not meet the guidelines from FADGI due to the large standard deviation.

### *Acquisition time*

Table 4 shows the average acquisition time per tile, the time per image and movement time from both scans. The difference in movement time was due to a faulty file naming line, which was removed in the second scan. Moreover, the robustness of image capturing was increased for the




(a) Area 3 of our scanner (FAST)

(b) Area 3 of the Bruker nano surface scanner

(c) The depth difference of each of the four area of the paint sample depicts the

Fig. 8: The height maps from our scanner and the Bruker nano surface scanner are compared with each other to determine the depth difference. The depth difference depict the depth precision of our scanner.

TABLE 3: The colour accuracy expressed with the ΔE-2000 metric for each colour patch of the Colour Gauge Micro Target.

| | | | | | |
|---|---|---|---|---|---|
| 1.07 | 1.43 | 0.15 | 0.52 | 0.22 | 0.33 |
| 1.76 | **5.04** | 2.28 | 2.11 | **6.11** | 0.15 |
| 0.44 | 1.83 | 0.24 | 0.26 | 0.11 | 0.12 |
| 0.05 | 1.25 | **9.87** | **6.00** | **3.05** | 0.24 |
| 0.16 | 0.27 | 0.21 | 0.20 | 0.19 | 0.48 |

second scan, resulting in 0.6 s extra per image. The minimal acquisition time per tile is reduced from 4 minutes per tile to 40 secondes per tile, which is a reduction of a factor 6. This means that a painting of 1750 $cm^2$ can be captured with a shutterspeed of 2 seconds and a overlap of 25% in just under 2 hours.

TABLE 4: Acquisition time of the two scans of 'The girl with a pearl earring' with the shutterspeed taken apart.

| | Time/Tile | Time/Image | Movement time |
|---|---|---|---|
| Scan 1 | 39.2 s | 1.66 s (+ 0.6 s) | 9.34 s |
| Scan 2 | 74.3 s | 2.26 s (+ 3.2 s) | 3.39 s |

### *Craquelure*

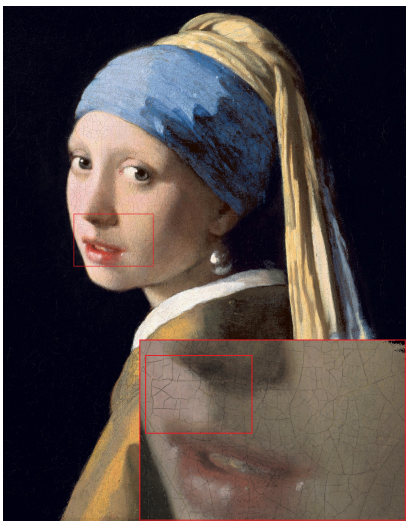
Fig. 9 shows 3D examples of the frequently detected craquelure of 'The girl with the pearl earring'. The 3D sections are enlarged and the height information is amplified to illustrate the profile of the craquelure examples. The convergent boundary like craquelure, shown in Fig. 9c, is the most encountered type of craquelure. One side of the crack is elevated and has a sharp decrease in height to the other side. The expected rectangular shaped craquelure is rather convex shaped as shown in Fig. 9d and is rarely detected in the depth maps of the entire painting. To give an indication, the width of the convex like craquelure is around 200 µm, measured from the beginning of the convex. The depth of the crack depends on when it is classified as a crack, but in the example varies from 20 to 80 µm deep.

## 5 Discussion

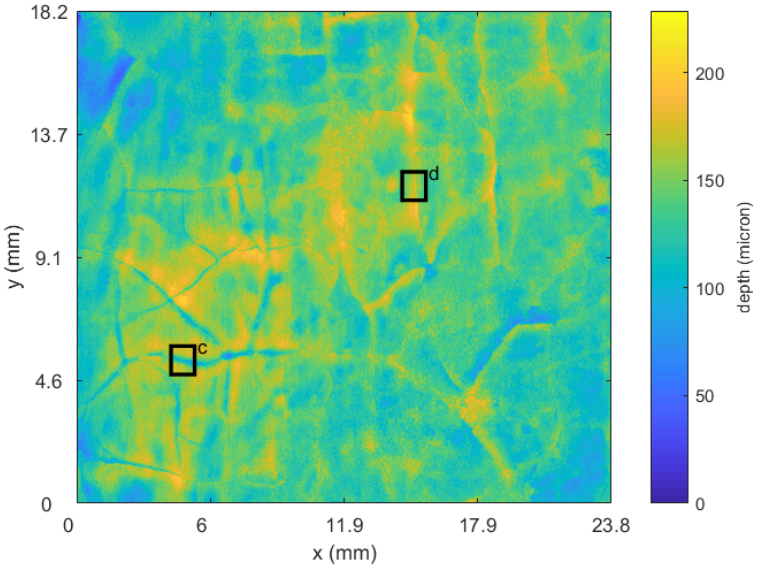
The results show that a 3D high resolution imaging device based on fringe encoded stereo imaging is an effective method to capture the surface of a painting in high resolution, but an imperfect technique to register the depth of craquelure. However, a primary estimation of most craquelure patterns and the beginning of cracks and thus the delamination, are accurately registered. The depth precision has shown to be 34 µm, which is worse than the estimated 5 µm based on the chosen hardware and the specifications of the technique [41] [2]. The difference in depth precision could be explained as result of a difference in validation method compared to the original design [45]. The depth precision could be increased with a larger camera angle ($\beta$). Still, the depth precision is better than other camera based scanning techniques mentioned in the literature. Moreover, the spatial resolution is with 7 µm ± 0.49 µm well below the required 10 µm. At this resolution, this scanner has a large potential of research possibilities like the environmental influences on the quality of a painting or the analysis of brushstrokes as mentioned in the introduction.

The colour accuracy does not meet the FARGI guidelines, because it is reduced by the projector's LED light source. The light spectrum of LED is less complete compared to other light sources and therefore has shown to poorly render colour [50]. The LED spectrum seems to primarily reduce the colour accuracy of the darker tones, which is compensated with a longer shutterspeed. However, the longer shutterspeed overexposes the white tones, reducing their colour accuracy. Even with the slight overexposure the overall colour accuracy is still better than for shorter shutterspeeds. To enhance the colour accuracy the light source of the projector should be replaced with a fuller spectrum light source.

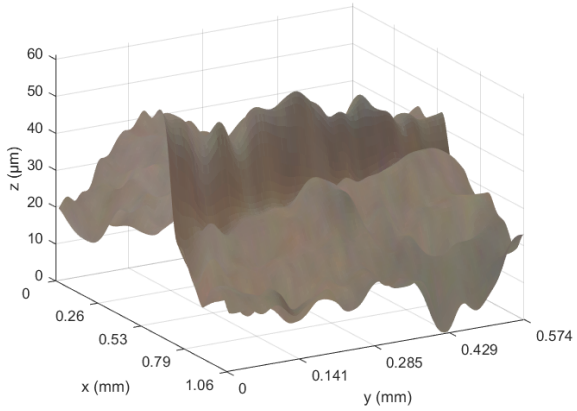
The analysis of the 3D information of 'The girl with a pearl earring' showed that the presented scanner is capable of capturing the craquelure pattern of the entire painting. Moreover, the profile of craquelure is more frequently an elevation resembling a convergent boundary. The elevation can be clarified as dirt that piled up in the crack during one of the restorations [51] or that the paint layers curve upwards at crack edge due to delamination at the edges. The craquelure that did have an inward profile rather ap-



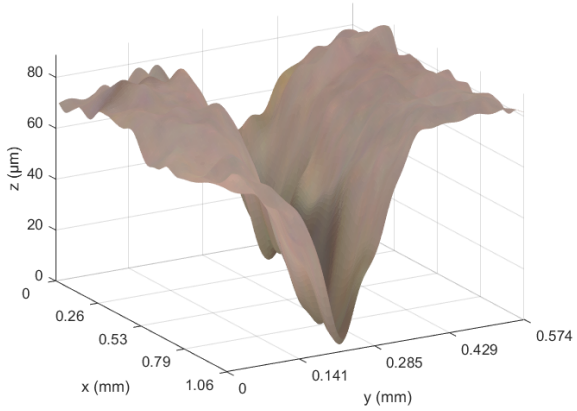
(a) Selected section of 'The girl with a pearl earring'




(b) An interpolated section of the height map of the face is enlarged to indicate the selected craquelure patterns for detailed visualisation.




(c) Convergent boundary like craquelure pattern




(d) Convex like craquelure pattern

Fig. 9: High resolution topographical and colour maps of 'The girl with a pearl earring' generate a 3D digitisation of the paintings surface. A 3D intersection of the profile of 2 typically occurring craquelure patterns are visualised.

proximates a convex shape than the expected right angled rectangular shape, which is the result of occlusions from stereo imaging. Both cameras need to register the same area to determine the depth and is therefore sensitive to occlusions in craquelure, which make the profile more convex. It requires more research to determine the exact measurement, type and origin of the profile of the craquelure. The 3D digitisation of a painting's surface will aid the classification of craquelure patterns in further research [16], in combination with new digital extraction and visualisation techniques [52].

The main drawback of increasing the resolution is a proportional reduction of the shared depth of field, increasing the difficulty to focus the images of both cameras. The images have to be focused sufficiently to be able to match the features of both images and to accurately compute the depth map. The minimal required amount of focus for fringe encoded stereo imaging to compute the depth map should be investigated in future research.

## 6 Conclusion

In this paper, we presented a high resolution colour and topography scanner capable of capturing the craquelure pattern of 'The girl with a pearl earring'. We replaced the hardware and software of the latest scanner based on fringe encoded stereo imaging to increase the spatial resolution, the depth precision, and the acquisition speed. Experimental validations show that the scanner reaches a spatial resolution of 7 µm $\pm$ 0.49 µm and a depth precision of 34 µm. The performance of digitally reproducing the colour information is just short of meeting the digitisation guidelines. The colour accuracy can be increased by replacing the light source of the projector with a light source with a more complete spectrum than LED.

The two scans of 'The girl with a pearl earring' indicated that a minimum acquisition time of 40 seconds per tile is required, resulting in an acquisition time of 2 hours for a painting of 1750 $cm^2$ with a 25% overlap and 2 seconds shutterspeed. The scans determined that the profile of the craquelure rarely has the expected inward convex-like shape. The profile of craquelure rather frequently has an outward profile approximating a convergent boundary. The origin and variation of craquelure profiles display an interesting path for future research.

## ACKNOWLEDGMENTS

The authors would like to thank Océ a Canon company for sponsoring the hardware required to develop the high resolution scanner, and the Mauritshuis for the opportunity to scan the painting 'Girl with a Pearl Earring' by Johannes Vermeer.

## REFERENCES

[1] Willemijn S Elkhuizen, Tim Zaman, Wim Verhofstad, Pieter P Jonker, Joris Dik, and Jo MP Geraedts. Topographical scanning and reproduction of near-planar surfaces of paintings. In *Measuring, Modeling, and Reproducing Material Appearance*, volume 9018, page 901809. International Society for Optics and Photonics, 2014.

[2] Tim Zaman, Pieter Jonker, Boris Lenseigne, and Joris Dik. Simultaneous capture of the color and topography of paintings using fringe encoded stereo vision. *Heritage Science*, 2(1):23, 2014.

[3] SD Jane, RE Barker, and JE Chad. Confocal microscopy and art conservation. *USA microscopy and analysis*, 31(55):31, 1996.

[4] Abbie Vandivere. Up close and personal, 2018. Accessed: 2018-05-29.

[5] Morteza Shahram, David G Stork, and David Donoho. Recovering layers of brush strokes through statistical analysis of color and shape: an application to van gogh's self portrait with grey felt hat. In *Computer image analysis in the study of art*, volume 6810, page 68100D. International Society for Optics and Photonics, 2008.

[6] Spike L Bucklow. A stylometric analysis of craquelure. *Computers and the Humanities*, 31(6):503–521, 1997.

[7] Olga Lelekova. Expert examination of icons. In *ICOM committee for conservation, 11th triennial meeting in Edinburgh, Scotland, 1-6 September 1996: Preprints*, pages 367–370. James & James (Science Publishers) Ltd., 1996.

[8] Sabine Stanek Elke Oberthaler, Jaap J. Boon and Martina Griesser. *The Art of Painting by Johannes Vermeer. History of treatments and observations on the present condition*. Wien: Residenz Verl, 2010.

[9] Joyce Hill Stoner and Rebecca Rushfield. *Conservation of easel paintings*. Routledge, 2012.

[10] Konrad Laudenbacher. Reentelado, parches y remiendos: mucha tela. In *Interim meeting, international conference on painting conservation: canvases, behaviour, deterioration and treatment: preprints*, pages 111–118. Editorial UPV, 2005.

[11] Alain Roche. Comportement mécanique des peintures sur toile: évaluation de la stabilité mécanique aux variations d'humidité et de température= comportamiento mecánico de la pintura sobre lienzo: evaluación de la estabilidad mecánica ante las variaciones de humedad y temperatura. In *Interim meeting, international conference on painting conservation: canvases, behaviour, deterioration and treatment: preprints*, pages 189–212. Editorial UPV, 2005.

[12] Jesper Stub Johnsen. Conservation of cultural heritage—european standards on the environment. *Climate for Collections Standards and Uncertainties*, page 35, 2013.

[13] Carla Arcolao. La craquelure o cavillatura superficiale: un'alterazione caratteristica delle superfici dipinte: cause, effetti, evoluzione tra storia e sperimentazione. In *Sulle pitture murali: riflessione, conoscenze, interventi. Atti del convegno di studi, Bressanone, 12-15 July 2005*, pages 549–556. Edizioni Arcadia Ricerche, 2005.

[14] Andreas Burmester, John Cupitt, Hervé Derrien, Nikolaos Dessipris, Anthony Hamber, Kirk Martinez, Manfred Müller, and David Saunders. The examination of paintings by digital image analysis. In *Proc. 3rd Int. Conf. Non-Destructive Testing, Microanalytical Methods and Environmental Evaluation for Study and Conservation of Works of Art*, pages 201–214, 1992.

[15] E Geake. Computer spots new cracks in old art. *New scientist*, 132(1798):20, 1991.

[16] Spike Bucklow. Consensus in the classification of craquelure. *Hamilton Kerr Institute Bulletin*, 3:61–73, 2000.

[17] Elzbieta Szmit-Naud. Brillance et couleur des retouches: causes de changements= glans en kleur van retouches: oorzaken van verandering. *Bulletin APROA/BRK*, 2:5–15, 2006.

[18] Gianluca Pastorelli, Costanza Cucci, Oihana Garcia, Giovanna Piantanida, Abdelrazek Elnaggar, May Cassar, and Matija Strlič. Environmentally induced colour change during natural degradation of selected polymers. *Polymer Degradation and Stability*, 107:198–209, 2014.

[19] Orio Ciferri. Microbial degradation of paintings. *Applied and environmental microbiology*, 65(3):879–885, 1999.

[20] Emilio Marengo, Maria Cristina Liparota, Elisa Robotti, and Marco Bobba. Monitoring of paintings under exposure to uv light by atr-ft-ir spectroscopy and multivariate control charts. *Vibrational Spectroscopy*, 40(2):225–234, 2006.

[21] Zdeňka Čermáková, Silvie Švarcová, Janka Hradilová, Petr Bezdička, Adriana Lančok, Vlasta Vašutová, Jan Blažek, and David Hradil. Temperature-related degradation and colour changes of historic paintings containing vivianite. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 140:101–110, 2015.

[22] Lucia Arbace, Elisabetta Sonnino, Marco Callieri, Matteo Dellepiane, Matteo Fabbri, Antonio Iaccarino Idelson, and Roberto Scopigno. Innovative uses of 3d digital technologies to assist the restoration of a fragmented terracotta statue. *Journal of Cultural Heritage*, 14(4):332–345, 2013.

[23] Topi Mäenpää and Matti Pietikäinen. Classification with color and texture: jointly or separately? *Pattern recognition*, 37(8):1629–1640, 2004.

[24] The Mauritshuis. Sample during treatment: girl with the pearl earring. Unpublished research, 1994.

[25] Spike Bucklow. The description of craquelure patterns. *Studies in Conservation*, 42(3):129–140, 1997.

[26] B Asady, M Tavassoli Kajani, A Hadi Vencheh, and A Heydari. Solving second kind integral equations with hybrid fourier and block–pulse functions. *Applied Mathematics and Computation*, 160(2):517–522, 2005.

[27] Ronald Newbold Bracewell and Ronald N Bracewell. *The Fourier transform and its applications*, volume 31999. McGraw-Hill New York, 1986.

[28] Katsumasa Fujita, Minoru Kobayashi, Shogo Kawano, Masahito Yamanaka, and Satoshi Kawata. High-resolution confocal microscopy by saturated excitation of fluorescence. *Physical review letters*, 99(22):228105, 2007.

[29] Willemijn S. Elkhuizen, Tessa T. W. Essers, Boris Lenseigne, Clemens Weijkamp, Yu Song, Sylvia C. Pont, Jo M.-P. Geraedts, and Joris Dik. Reproduction of Gloss, Color and Relief of Paintings using 3D Scanning and 3D Printing. In Tobias Schreck, Tim Weyrich, Robert Sablatnig, and Benjamin Stular, editors, *Eurographics Workshop on Graphics and Cultural Heritage*. The Eurographics Association, 2017.

[30] Federal Agencies Digitization Initiative and el. Technical guidelines for digitizing cultural heritage materials: creation of raster image master files, 2010.

[31] Nina Siegal. Uncovering the secrets of the 'girl with a pearl earring', 2018. Accessed: 2018-05-18.

[32] Rebecca Laurence Alastair Sooke, Alan Stockdale. The secrets of girl with a pearl earring, 2018. Accessed: 2018-05-18.

[33] Joyce H Townsend. Microscopy and paintings. *Microscopy and analysis*, 39:11–13, 1994.

[34] Ilaria Barbetti, Alberto Felici, Donata Magrini, Rachele DEL FA, and Cristiano Riminesi. Ultra close-range photogrammetry to assess the roughness of the wall painting surfaces after cleaning treatments. *International Journal of Conservation Science*, 2013.

[35] Thomas Luhmann, Stuart Robson, SA Kyle, and IA Harley. *Close range photogrammetry: principles, techniques and applications*. Whittles, 2006.

[36] PhotoModeler. Factors affecting accuracy in photogrammetry, 2017. Accessed: 2017-12-08.

[37] Brian Owens and Andrew R Scott. Software: Picture perfect. *Nature*, 545(7654):S12–S12, 2017.

[38] Robert J Woodham. Photometric method for determining surface orientation from multiple images. *Optical engineering*, 19(1):191139, 1980.

[39] Svetlana Barsky and Maria Petrou. The 4-source photometric stereo technique for three-dimensional surfaces in the presence of highlights and shadows. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(10):1239–1252, 2003.

[40] Christine A Curcio, Kenneth R Sloan, Robert E Kalina, and Anita E Hendrickson. Human photoreceptor topography. *Journal of comparative neurology*, 292(4):497–523, 1990.

[41] Yu Song Clemens Weijkamp Sylvia Pont Willemijn S.Elkhuizen, Tessa T. W. Essers and Joris Dik Jo M. P. Geraedts. Gloss, color and topography scanning for reproducing a painting's appearance using 3d printing. Unpublished, 2018.

[42] Elizabeth Moodey. The flemish primitives: catalogue of early netherlandish painting in the royal museums of fine arts of belgium.

[43] Sean T. Mchugh. *Understanding Photography - Core Concepts and techniques with digital image capture*. No Starch Press, 2016.

[44] Zhengyou Zhang. A flexible new technique for camera calibration. *IEEE Transactions on pattern analysis and machine intelligence*, 22(11):1330–1334, 2000.

[45] Tim Zaman. Development of a topographical imaging device for the near-planar surfaces of paintings. *TU Delft Repository*, 2013.

[46] David G Lowe. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2):91–110, 2004.

[47] Yukio Okano. Mtf analysis and its measurements for digital still camera. In *Is and T annual conference*, pages 383–387. The society for imaging science and technology, 1997.

[48] Bruker Inc. Contourgt-k 3d optical microscope datasheet, May 2018.

[49] Gaurav Sharma, Wencheng Wu, and Edul N Dalal. The ciede2000 color-difference formula: Implementation notes, supplementary test data, and mathematical observations. *Color Research & Application*, 30(1):21–30, 2005.

[50] Elodie Mahler, Jean-Jacques Ezrati, and Françoise Viénot. Testing led lighting for colour discrimination and colour rendering. *Color Research & Application*, 34(1):8–17, 2009.

[51] Rian MH Deurenberg. Examination and treatment of a set of klismos chairs, attr. to john and hugh finlay. *Journal of the American Institute for Conservation*, 47(2):97–117, 2008.

[52] Kai Lawonn, Erik Trostmann, Bernhard Preim, and Klaus Hildebrandt. Visualization and extraction of carvings for heritage conservation. *IEEE transactions on visualization and computer graphics*, 23(1):801–810, 2017.

# II

# Literature Study

# 1

# Introduction

The interest in art research has intensified over the past two decades. The technical progress in scanning technologies has enabled researchers to uncover new insights into the creative process of an artist. Especially scanning technologies focused at the material aspects of oil paintings have resulted in a deeper understanding of the painting, often extending into the understanding of our cultural heritage. Advanced techniques like high-intensity X-ray fluorescence can observe the underlying layers of a painting, which has lead to the discovery of a secondary oil painting underneath the surface [1]. Or techniques, like photometric stereo imaging , which digitises the surface of oil painting in 3D to increase the accessibility to the art work [2]. Other techniques are specialised on capturing the painting to reproduce the painting with 3D printing [3]. All these techniques give a better insight in the mind of the artist, which is important for study, restoration and conservation of the art work.

The state-of-the-art non-invasive 3D scanning techniques either capture a small area with ultra-high resolution, like confocal microscopes [4], or the entire surface of the painting with a medium resolution [3]. Few devices can accurately capture near-planar surface in 3D ultra-high resolution for the entire surface of paintings.

The primary visualisation of the degradation of paintings is the fine cracks in the paint, which is named craquelure. The craquelure pattern is one of the main factors in determining the state of a painting and broad indicator of authorship [5]. It arises as part of natural ageing of the paint [6]. The paint molecules reorganise over time due to chemical reactions initiated by environmental factors, like change in temperature, light or humidity [7]. The reorganisation results in surface tension in the paint layer, which eventually results in a rupture. Moreover, at the edge of craquelure the paint can delaminate and thus become detached from the canvas. Therefore, being able to track the craquelure pattern of the entire painting in 3D, is extremely useful for conservationists to determine the restoration process.

### Environmental conditions

The environmental condition is one of the main causes of the degradation process. Currently, it is difficult to determine the objective impact of the environment of the painting. The existing knowledge is based on model and individual cases [8]. Being able to capture the craquelure pattern for the entire painting would allow research of the environmental influence on the quality of the painting.

### Baseline digitisation

Being able to digitise the colour and near-planar topography of the entire painting at ultra high resolution, will enable owners to digitise the state of the painting as a baseline for future reference. In the case the painting gets damaged during a loan or storage period, the owner can compare the baseline craquelure with the current state.

So state-of-the-art devices either lack the ultra-high resolution or the manoeuvrability to capture the craquelure pattern of the entire painting. This literature study will aim at laying the foundation for the design of a 3D scanner able to register the craquelure for the entire painting.

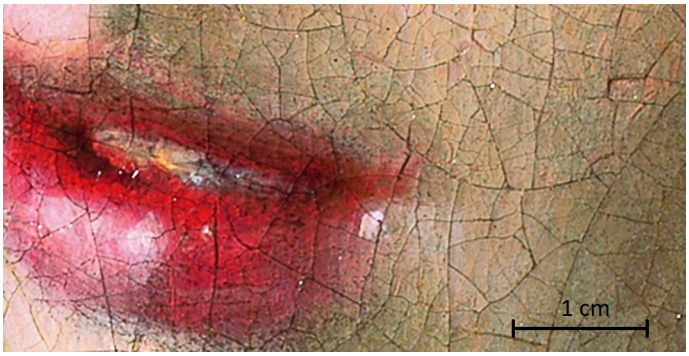



Figure 1.1: An enlarged image of the lips and cheek of the girl with a pearl earring. The very fine craquelure is clearly visible from this image [9].

### The girl in the spotlight

The applied scientific research techniques have intensified over the past two decades. The impact of research on the material aspects of the painting often extends into the understanding of our cultural heritage. Therefore, after 24 years, the Mauritshuis in the Netherlands decided to allow state-of-the-art techniques to research the world famous painting 'the girl with a pearl earring' from Johannes Vermeer. The Mauritshuis hopes to discover more secrets of Johannes Vermeer's most famous painting [9]. The research project took take place from end February until beginning of March 2018. Each research technique is given a time slot. The Mauritshuis did not want to take the painting away from public display, so all research will take place in open view of the museum visitors. For this reason a glass enclosure will be build to protect and shield the painting, techniques and researchers from the public.

This unique research possibility became known in the beginning of my project (September 2017). It is not only an ideal opportunity to scan a craquelure pattern of a famous painting, but especially this painting makes it interesting because of its very fine craquelure, shown in figure 1.1 and relatively flat surface. The scanner has to be functioning within 6 months, which gives a challenging time restraint. The research also has to take into account the time slot of 15 hours, in which the complete painting (39 x 44.5 cm à 1750 $cm^2$) must be captured. Obviously the safety of the painting should be guaranteed in any situation.

### Craquelure

In order to capture the craquelure pattern with sufficient detail, the requirements of the scanner should be designed to the depth and width of an average rupture. However, the dimensions of craquelure are dependent on several factors, like the thickness of the paint layer, the type of paint, the age of the painting, or the underlying layers [10]. Therefore, a conclusive definition on the average dimensions can not be given. As mentioned, 'the girl with a pearl earring' is the ideal case to design the scanner to, because of its fine craquelure. Being able to accurately scan very fine craquelure, confirms the capability to capture the craquelure of most paintings.

In 1994 the Mauritshuis conducted the last research project on the girl with a pearl earring, during which they took samples from the painting shown in figure 1.2. These samples show the thickness of the top paint layer to be 25 up to 300 µm (h), which determines the height of the craquelure. The width (w) of craquelure is determined with high resolution image and is measured at about 50 up to 100 µm [9].

Craquelure tends to rupture in smooth lines and patterns as can be seen in figure 1.1. The profile of the craquelure rather approximates right angles than a convex shape, as a result of ageing [12]. In signal analysis a rectangle can be approximated with a sum of multiple sinusoids with a sampling frequency of at least five times larger than the signal [13] as illustrated in figure 1.3. The figure shows how a simplified profile of right angle shaped craquelure is estimated with a minimal of 5 samples. We want to estimate the even the smallest craquelure of 25 µm depth and 50 µm wide. The dimensions of such a crack can be related to signal analysis by imaging the width and depth as the wavelength ($\lambda$) and amplitude (A) of the wave. The frequency on which the sampling frequency is based, can be determined with equation (1.1). The highest frequency required to describe the craquelure profile is

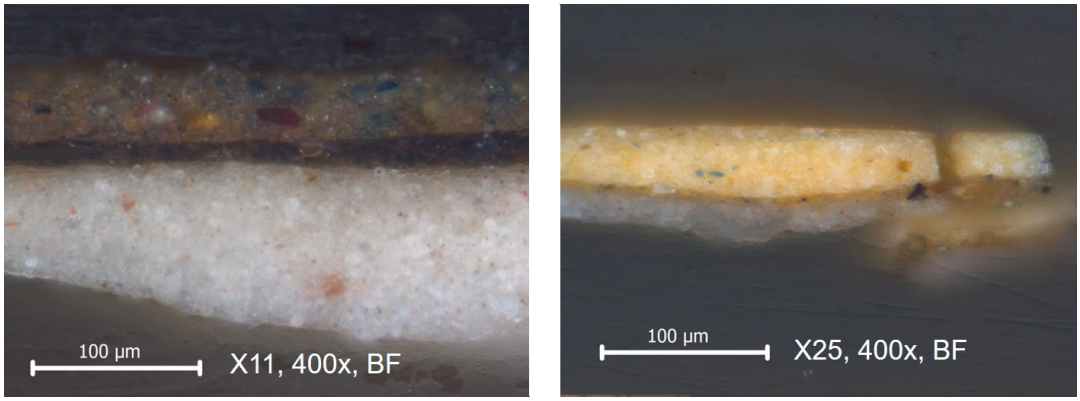



Figure 1.2: Samples of the paint layers taken from several areas of the 'the girl with a pearl earring'. The thickness of the top layer can be determined from these samples, which gives insight of the craquelure depth [11].

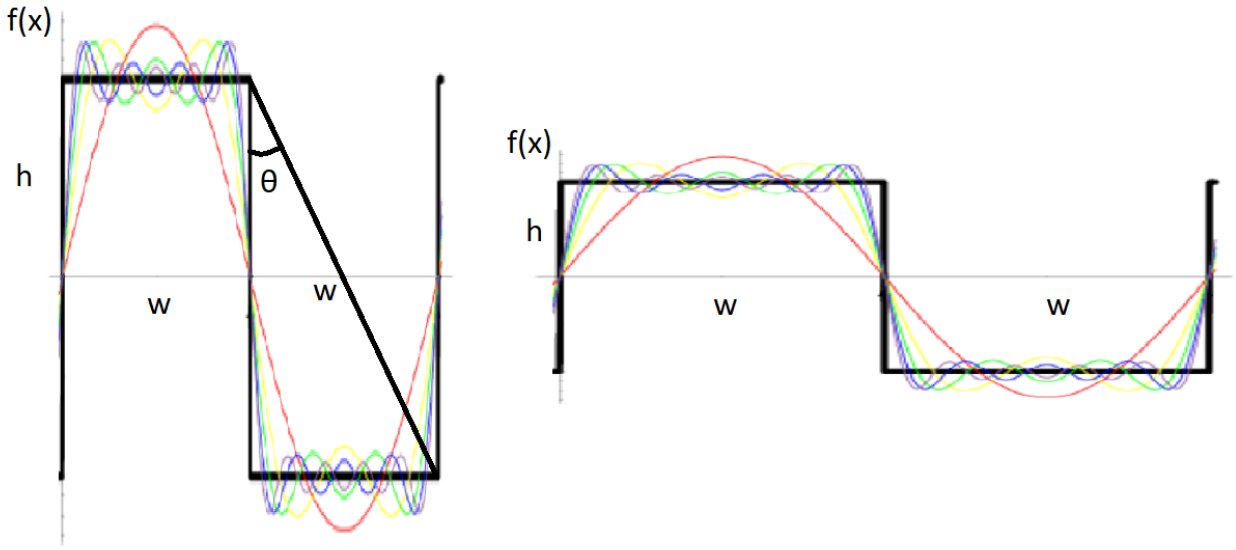



Figure 1.3: Fourier analysis approximates a rectangle with multiple sinusoids. Based on the minimal sample frequency to approximate the rectangle is the spatial resolution and depth precision determined [13].

20.000 Hz for a width of 50 µm. As mentioned, to describe the right angle profile in signal analysis, the sampling frequency has to be at least 5 times larger than the frequency describing the wavelength [13]. The sampling frequency should therefore be 100.00 Hz, which is converted to a wavelength of the spatial resolution of 10 µm for a width of 50 µm. Simply put, there should be 5 measurement points within the crack to give a basic description of the smallest crack. Applying the same reasoning to the required depth precision gives a depth precision of 5 µm for the smallest depths of 25 µm. So based on findings from 'the girl with a pearl earring', the design requirements of a scanner capable of scanning the craquelure pattern are a spatial resolution of 10 µm and a depth precision of 5 µm.

$$f = \frac{1}{\lambda} = \frac{1}{50 \cdot 10^{-6} m} = 20.000\ Hz \tag{1.1}$$

The example of craquelure on the left in figure 1.3 also indicates one of the issues of scanning craquelure, which are deep cracks. Most scanning techniques will have trouble registering depth when the crack is small and deep, due to occlusions, loss of shading, or the depth becomes out-of-focus. In the worst case a crack would be 50 µm wide and 300 µm deep (1:6), which is 3 times deeper than displayed in figure 1.3. Any scanning device could register the depth of such an crack only with an angle ($\theta$) of smaller than 10°. However, the lack of literature on the exact dimensions of craquelure

gives no clear picture on the profile of craquelure. The surface of 'the girl with a pearl earring' is relatively flat and thus deep crack are expected to rare. In the end, the subject is still an important design requirement for the scanner.

Furthermore, oil paintings tend to sag in the middle of the painting as result of shrinking and stretching of the canvas. Among other, the factors that also play a role in sagging are the size of the painting, the type of listing and environmental influences [14]. Moreover, scanning the painting horizontally might be more stable, but increases the sagging of the painting compared to vertical scanning.

## 1.1. Literature goal

The goal of this report is to design a high resolution topography and colour scanner for the craquelure pattern of paintings based on available literature. The requirements are set by the Mauritshuis and the dimensions of the craquelure, which are:

| | |
|---|---|
| Spatial resolution | 10 µm |
| Depth precision | 5 µm |
| Acquisition time | <15 hours for 1750 $cm^2$ |
| Ability to register craquelure (1:6) | [-+] |

Table 1.1: The design requirements for a 3D high resolution scanner applied to the craquelure pattern of paintings.

## 1.2. Literature review outline

This literature study aims to analyse the current colour and topography capturing techniques that could accurately capture the craquelure pattern. In chapter 2 the study begins with a holistic view of different areas of techniques that can be used in capturing colour and topographical information of the painting. A thorough review of the technique best suited to capture the craquelure pattern is given in chapter 3. In addition, the associating issues of the technique are examined. The final two chapters review literature for methods to resolve theses issues.

# 2

# Non-Invasive Topography Capturing Techniques

The goal of this literature report is to develop a high resolution topography and colour scanner for the craquelure pattern of paintings. This chapter discusses the-state-of-the-art techniques and devices able to capture topography (i.e. texture) and/or colour information. Invasive techniques (i.e., based on contact) are excluded from the discussion, since the painting's surface cannot be touched. These techniques are accessed on performance criteria and goals shortly mentioned in the previous chapter. The (hybrid) device should be able to capture the topography and colour of the near-planar surface of paintings. Adding to the spatial resolution and depth precision stated in the design requirements (table 1.1), a fund of maximum 20.000 euros was assigned to the development of the device. The device is required to capture the entire painting of 1750 $cm^2$ within 15 hours as discussed in 1, without physical contact with the painting. Furthermore, the estimated costs to develop the device and provided resources by external parties, reducing the costs, are determined and taken into account. Lastly, the technique is graded at the ability to registering depth of (deep) craquelure, giving an indication of the drawback of the techique.

This chapter gives an holistic view of different areas of techniques that can be used in capturing colour and topographical information of the painting. Next, the currently existing devices that capture topography and colour simultaneously are discussed. The section related techniques shortly discusses considered techniques that are unsuitable as a solution. Finally, the devices are compared, and a conclusion is made.

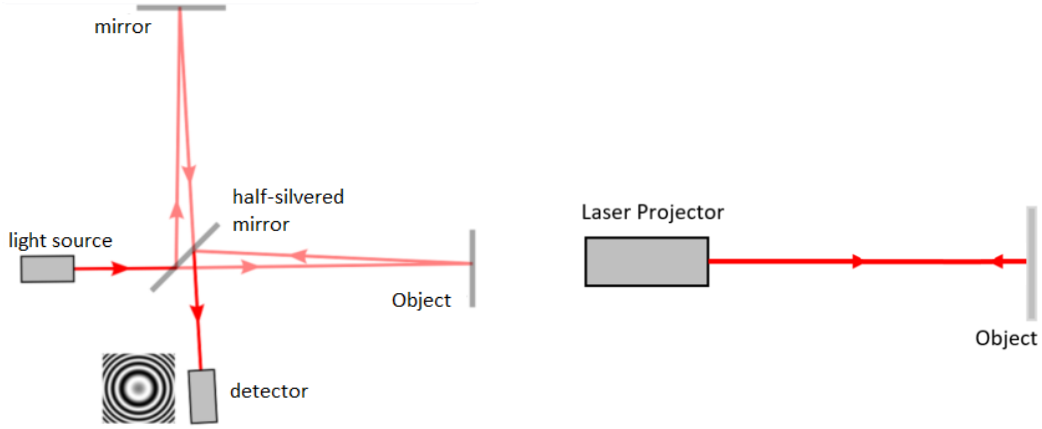
## 2.1. Topography Capturing Techniques

In this section a broad review is done on topographical capturing devices, which in combination with a colour image would reproduce an accurate presentation of the painting's surface. The working principle of each topographical device is shortly explained, followed by an assessment on:

- The depth precision and spatial resolution estimation of the surface [µm]
- The acquisition speed of the depth registration [h]
- The cost estimation of the device [€]
- Ability to register (deep) craquelure

### 2.1.1. Ultrasound Imaging

Ultrasound imaging is a technology that uses high frequency sound waves. The ultrasound waves are reflected on transition between layers based on density difference, the reflection is proportional to the difference in impedance of these layers. In other words, if the difference in density decreases, the reflected sound is also decreased, and thus transmitting sound waves are increased. With the reflected light there are two different methods to determine the distance to the object: Time-Of-Flight (TOF) and interferometry both are shown in figure 2.1.




(a) Interferometry based on a light beam (b) Time Of Flight based on a light beam

Figure 2.1: Figure (a) displays interferometry, in which the source is split into two beams. One part is reflected by a mirror, the other reflects on the object. Both beams are recombined causing interferometry, which contains the information of the distance. The figure is adopted from [19]. Figure (b) displays a Time Of Flight (TOF) based devices, which determines the distance by the time it takes for the emitted beam to return to device.

### Time-Of-Flight

The emitted beam reflects on the surface of the object back to the source and observer. The time it takes for the reflection to return is known as the TOF. Depending on the emitted source the time of flight determines the distance to the object [15]. In ultrasound imaging, TOF is called echolocation.

### Interferometry

In interferometry, light is split into two rays that travel different paths and are combined again to determine interference. The interference gives information about the difference between the optical paths [16] and is highly accurate. This technique, based on ultrasound is commonly used to determine the thickness of thin plates [17] or for the detection and characterisation of thin layers between two much thicker media [18].

This technique preferably uses a higher density liquid as medium between the device and the object, but a painting can only use air as medium. Since, the difference in density between the painting and air is enormous, roughly all sound waves are complete reflected on the surface. Thus less information will be retrieved about the underlying layers and most information about the surface, which is exactly the area of interest. However, to reach the best results, the device should be connected to the object, but physical contact is excluded from the options for paintings.

Moreover, for TOF the depth precision depends on the distance to the object and the quality of the device. For example a 10MHz soundwave, the receptor is required to measure with an accuracy of tens of nanoseconds in order to measure with 8,5 µm depth precision [20]. Increasing the frequency would increase the depth precision, however it also increases the attenuation and hence lowers the signal-to-noise ratio at the receptor. On the other hand, the working distance, which is the distance from the device to the object, limits the spatial resolution of TOF to about tenths of millimeters. The depth precision of interferometry is limited by the precision of sound-waves and receptor. Moreover, ultrasound interferometry is mainly used in thin plates, which can reach sub-micrometer depth precision [21]. The spatial resolution of ultrasound interferometry is theoretically limited increments of the linear stage moving the device, which could reach an accuracy of 0.1 µm [22]. In practise, the spatial resolution is limited by the laser diameter and the working distance required for painting scanning to estimated 4-10 µm [23].

The acquisition speed of interferometry based ultrasound is much lower than TOF based devices, since the depth is computed for multiple points per second. TOF based ultrasound can register depth for a line, increasing the acquisition speed relative to interferometry. Furthermore, increasing the spatial resolution will drastically increase the acquisition time, which applies to both techniques. Moreover, both ultrasound imaging based techniques require a hybrid solution with colour images of the painting to reproduce a 3D image, which also increases the acquisition time.

The estimated cost of an ultrasound imagine device varies from 5.000 to >100.000 depending on

the quality of the device [24]. Only the highest price segment of ultrasound TOF devices can reach the required depth precision. Techniques based on interferometry are expensive and are estimated to cost about a half ton per device, making ultrasound imaging a costly technology to suit our criteria.

The ability of ultrasound based devices to register the depth and profile of craquelure is primarily limited to scan directly above the craquelure. As mentioned, the high density difference causes most sound-wave to reflect on the surface, therefore scanning a crack at an angle will result in occlusions, which depend on the indicent angle of the sound-wave. Ultrasound imaging is therefore less ideal, but still a good option for craquelure registering.

### 2.1.2. Optical Coherence Tomography

Optical Coherence Tomography (OCT) determines the depth with interferometry, but it uses light-waves instead of sound-waves. The technique uses low-coherence interferometry to determine a two dimensional image of light reflected from up to 2 mm depth. Applying the technique to paintings has been done successfully to determine the varnish (protection) layer of paintings [25].

The current OCT based painting scanner reaches a spatial resolution of 4-10 [22] [26] micron and up to about 2 millimetres in depth [25]. However the spatial resolution is limited due to aberrations induced by the imaging optics of the specimen [27]. In theory, light based interferometry is able to achieve picometer accuracy with a moving light source [28]. Furthermore, the acquisition speed differs per OCT device, in a ideal situation this technique stakes 6 point measurements per second [29]. Based on the required spatial resolution of 10 µm, it would take about 100 hours to scan 1750 $cm^2$. The spatial resolution depends on the step size of the optical coherence tomography, and as mentioned, the spatial resolution is proportional to the acquisition time. Thus in most cases the available scanning time limits the spatial resolution. Moreover, like most techniques based on interferometry the cost is quite high. An Optical Coherence Tomography device is estimated to cost about 40.000 euros [30].

The ability of OCT to register the craquelure isl similar to ultrasound imaging techniques, most of the depth information is obtained from the surface. Therefore, OCT is also sensitive to occlusions when the incident angle is not perpendicular. OCT often scans under a very small angle ($< 5°$) to reduce reflections from the surface, but this also introduces occlusions. The indicent angle of the OCT scanner is smaller than the angle stated in the introduction to register the depth of craquelure. Therefore, OCT scanning is a good option for craquelure registering, but the colour and topography are captured separately, which requires mutual information matching the colour image with the height map.

### 2.1.3. (Line) Laser Projection

This section discusses Line Laser projection based on Time Of Flight and triangulation. The depth from triangulation method originates from the position of a single laser dot, which can be determined by triangulation when the orientation and position of the laser source and observer is known[31], shown in figure 2.2. This technique is also able to determine depth from multiple points in a line [32][33][34], also called laser profilometry.

Laser projection can also estimate the depth based on Time OF Flight. Similar to ultrasound imaging [35] the time the light takes to return determines the depth. For TOF, it is not possible to scan the surface with a line at high spatial resolution and is therefore limited to point measurements.

The depth precision of laser scanning techniques based on TOF are limited due to the high speed of light. Currently, this method can ideally reach sub-millimeters of depth precision. Similar to echolocation the spatial resolution depends on the working distance and the minimum step size. Theoretically the spatial resolution could reach sub-micrometer when combined with a high precision linear stage, which cost about 500 euros [22] for smaller stages. Very little linear stages can cover the size of a painting. However, the depth precision of TOF based technologies is insufficient for the sub-millimetre measurements, where triangulation and other techniques are often used.

For triangulation, the depth precision depends on the accuracy of the estimated position and orientation of the laser source and the observer. Similar to ultrasound imaging, a smaller working distance increases the spatial resolution and depth precision. Moreover, the resolution of the laser source and camera influence the accuracy of the depth registration. In theory, the depth precision for triangulation based techniques could reach the 0.4 micrometer [36], since the wavelength is the limiting factor. The spatial resolution is for such a line laser triangulation device is 10 µm [37] [38].

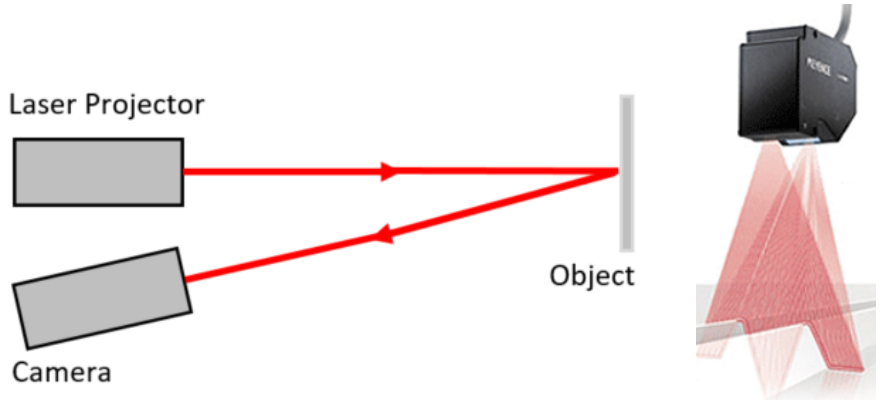



Figure 2.2: Laser Triangulation: A laser point is projected at the object and the reflection is captured by the camera. The position and orientation of the Laser Projector relative to the camera is used to determine the distance to the object with triangulation.

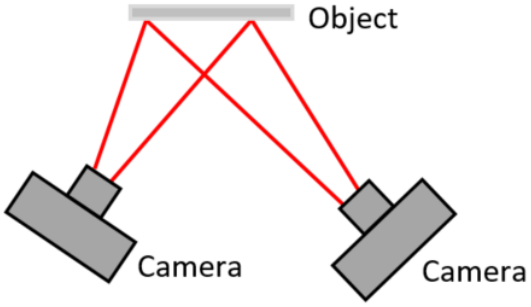



Figure 2.3: Stereo imaging captures the object with two cameras with a different viewpoint. The depth is determined by triangulation with the relative position and orientation from both cameras.

The acquisition speed of laser projection depends on the technique. On a near-planar surface multiple laser lines with triangulation speeds up the process. Using a single line to register the height of the entire painting takes a considerable amount of time, mainly due to requested the spatial resolution. For example, it would take a Line Laser triangulation scanner 50 hours to scan 1750 $cm^2$ (A.1.1). Moreover, both triangulation and interferometry would require a hybrid solution with colour images, resulting in a longer acquisition time [39]. In summary, Line Laser projection triangulation is a method with the potential to scan the entire painting at the required depth precision and spatial resolution, but is quite slow.

The laser triangulation and Time-Of-Flight are relatively economical techniques. Several start-ups have developed affordable 3D scanners based on Line Laser projection, like Eora 3D or Altas 3D. Laser Time-Of-Flight devices range from 100 to 1000 euro, depending on the depth precision and range. The estimated costs for Line Laser triangulation devices are over 1000 euros a piece [40].

The ability of line lase projection to register the profile of craquelure decreases as the angle of the line laser increases, as illustrated in figure 2.2 on the left. The depth of the deepest craquelure can not be determined when the angle of the laser is larger than 10°, as mentioned in the introduction. For laser line projection the ability to register craquelure is a trade-off with the acquisition time since a wider line projection allows for more surface to be scanned. Line laser projection is, therefore, less ideal to scan the deepest craquelure patterns.

### 2.1.4. Stereo Imaging

Stereo imaging constructs a three dimensional image by matching two images with a relative difference viewpoint of the scene as shown in figure 2.3, which is similar to how the human eyes determine depth. The relative distance and orientation between these viewpoints is essential to triangulate the position in the three dimensional space. Moreover, the camera and lens determine the spatial resolution, because they determine the magnification and pixel density of the camera sensor.

The depth is estimated by the difference in position of matching areas of each camera. The stereo

image registration process matches areas in the images to generate the disparity map [41][42]. Image matching can be either with sparse matching by matching features in the image [43] or with dense matching [44] by computing the lowest entropy for the entire images. The disparity map contains the difference in position between the stereo images, which is used to compute the absolute height information of the scene [45].

The main drawback of stereo vision is that it can only triangulate positions where the data in the image is salient and where there are features that can be matched with enough certainty. Where there are good and identical features, stereo matching and triangulation can be highly accurate. The background of paintings is often not salient and would thus cause errors in matching, making stereo imaging not a suitable technique for a scanner. Moreover, stereo imaging is sensitive to occlusions, shading and reflections of the surface, these artefacts create a mismatch between the cameras.

The depth precision of stereo imaging depends on the accuracy by which the position and orientation to both cameras is known, since the relative positioning of both cameras is used to triangulate the depth information. Furthermore, the depth precision and mainly the spatial resolution are limited by the pixel size of the camera sensor and the magnification of the lens. A greater magnification lens will result in more pixels capturing the same area, enabling a higher spatial depth resolution and depth precision. However, there is a limit to the magnification of a lens, the best macro lenses theoretically could reach the 1 µm per pixel spatial resolution and 0.5 µm depth precision (A.1.2).

Furthermore, the capturing speed and costs primarily depend on the camera and the lens. The required image resolution and exposure time to capture the detail of the scene determine the acquisition speed. Stereo imaging could scan the entire painting at the 1 µm per pixel spatial resolution with specialised magnification lenses. A scan of 1750 $cm^2$ with this technique is estimated at 1 hours with a full-frame camera, magnification lens and 30 % overlap in the images, since only 1 capture of each camera is required for each position of the painting.

The camera and lens able to achieve the required resolution are quite expensive. The estimated cost is about 5000 euro for the cameras and an additional 4000 euro for the lenses.

The ability of stereo imaging to determine the profile of craquelure is limited due to occlusions. Both cameras need to capture the same area to triangulate the depth information. The indicent angle, mentioned in the introduction, for both cameras becomes 5° instead of 10° to determine the depth of deep craquelure with stereo imaging. However, the depth precision of stereo imaging depends on a larger indicent angle. The ability of stereo imaging to register the profile of craquelure is a trade-off with the depth precision. For example; stereo imaging with a common indicent angle (20°) can register depths up to 70 µm for the narrowest cracks (50µm). Stereo imaging is therefore not the ideal technique to register deep craquelure.

### 2.1.5. Shape from Focus and Defocus

Shape from focus and defocus determines depth by using different focus levels. The object is translated in front of the camera and with the sum-modified-Laplacian operator local measures of image focus are determined. The height map is generated by an ad-hoc interpolation strategy of the in-focus areas from each focus level [46]. The shape from focus is improved by using the defocus bokeh (blur) of the image data to retrieve an enhanced estimate of the structure [47][48][49].

The resolution of the colour image is limited by the resolution of the camera and lens. The depth precision of shape from focus depends on the step size by which the linear stage with the object is moved and the size of the Depth Of Field (DOF). The parallel area in front of the camera considered to be in focus is called the DOF. For example, with a full-frame camera and high magnification lens, the spatial resolution and a small depth of field is achieved. However, for this ideal situation the DOF is still only 0.0463 millimetre (A.1.3). The depth precision is roughly assessed to be at the Depth OF Field of about 50 µm, which is not sufficient for the depth requirement.

The acquisition speed is limited by the amount of images required for the height map. It becomes a trade-off between capturing time and depth resolution. Imagining the technique could reach the precision, it would take about 14 hours to get the depth increments for the entire painting. Shape from focus and defocus is one of the most economical techniques. The technique only requires a camera and a linear stage, costing an estimated 5000 euro for the camera and lens, and about 300 euro for the linear stage [22]. However, vibrations can become an issue for the sharpness of the microscopic

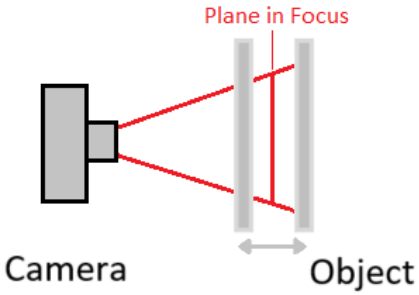
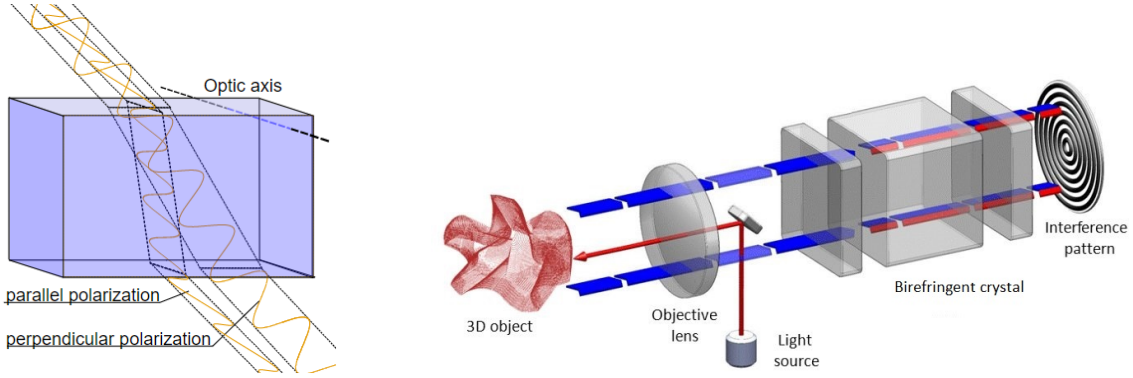



Figure 2.4: Shape from Focus: The object (or camera) is moved with small increments by a linear stage, while taking photographs. Shape from Focus and Defocus determines depth from when the object is in focus and when it is out of focus.




(a) Birefringent crystal [51] (b) Conoscopic holography [52]

Figure 2.5: (a) The birefringent crystal refracts light ray depending on the direction of the light ray. (b) Light is directed to the object with a dichroic mirror and the reflected light passes through the birefringent crystals. Light with a different direction is refracted with a different angles. The combined light creates a inference pattern on the sensor take is translated to a distance with frequency analysis.

resolution images. Having a translational table introduces another stage that generates vibrations that need to be minimised. The shape from focus uses the captured colour images to determine the depth information, which eliminates the need for a hybrid solution with a colour camera.

The ability of shape from focus and defocus to register the profile of craquelure depends on the Depth Of Field and the lighting in the craquelure. The Depth Of Field is not small enough to determine the depth of the shallow craquelure, but the deepest cracks could be registered when lighted well, which in turn introduces shading. Moreover, some occlusions might occur at the edge of the camera window, but this is limited since the indicent angle is still small. In summary, shape from focus and defocus is not good for capturing the craquelure pattern of paintings.

### 2.1.6. Conoscopic Holography

Conoscopic Holography determines depth by frequency analysis of the diffraction pattern created by the birefringent crystals. These differ in refractive index depending on the direction of the light ray as shown in figure 2.5a. Similar to interferometry the light is projected onto the object with a dichroic mirror. The reflected light is passed through the birefringent crystal, resulting in a interference pattern that is used to compute the distance to the object by frequency analysis [50], shown in figure 2.5b.

In Conoscopic Holography, the frequency analysis and birefringent crystal determine the depth precision. Therefore, it is difficult to give an exact number on the techniques depth precision, but it is comparable to OCT [53][54]. Like with OCT the spatial resolution is mainly limited by the translational capabilities of the linear stage.

Furthermore, conoscopic holography can simultaneously capture the colour information and depth estimation. This lowers the acquisition time, but is still limited to a point measurement. Thus like OCT is acquisition time the limiting factor for scanning the entire painting. Moreover, Conoscopic holography is

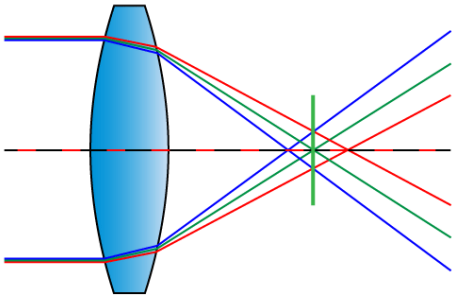


Figure 2.6: The refraction of the lens affect the longer wavelengths more, creating chromatic aberrations. Chromatic aberrations is an undesired phenomenon in photography, because it separates the colours.

not a popular choice for companies to use for depth estimation, since most focus on confocal microscopy mainly for its greater acquisition speed. Therefore, the price fairly difficult to determine, so based on comparable techniques, like OCT and confocal microscopy, the price is estimated to be about 50.000 euro .

Cononscopic holography is good method to register the craquelure of paintings. This method uses point measurements and perpendicular light source to determine the craquelure profile, which takes more time to scan, but it eliminates occlusions. Moreover, the light source projects perpendicular, minimising shading. Conocopic holography also allows for simultaneous capture of the depth and colour information eliminating the need for a hybrid solution with a colour image.

### 2.1.7. Depth from Chromatic Aberration: Nano Point Scanner

The Nano Point Scanner (NPS) determines the height of a single point by using chromatic aberrations. This is a lens phenomenon, in which refraction affects longer wavelengths more, as shown in figure 2.6. Similar to OCT, light is directed to the object with a dichroic mirror. The lens refracts the light, forming chromatic aberrations, which have a different focus distance at each wavelength used to determine the depth [55] [56].

The strength chromatic aberrations determine the total measurable depth and the just notable difference. The accuracy depends on the algorithm that determines at what light wave the depth point is in focus and the accuracy by which the aberrations are approximated. In theory depth from chromatic aberrations can achieve high depth precision of sub-micrometers [57], because it only limited by the wavelength of the light.

Furthermore, the nano point scanner can scan multiple points per second. Similar to OCT the depth estimation is based on point measurements and thus also a similar acquisition time (100 h). The spatial resolution and depth precision limit the acquisition speed. Theoretically the NPS can scan the entire painting at the requested spatial resolution and depth precision, however the height deviation of an average painting (>1mm) will go beyond the range of the NPS. The distance to the object would have to be adjusted during scanning, which is possible with a linear stage, but will introduce new issues as how to determine the required working distance. The price of the NPS is not yet known, but it can be estimated by the costs of other devices with a similar build, which would set the estimated price at about 50.000 euros.

Similar to conoscopic holography, the Nano Point Scanner uses point measurements to determine the craquelure profile, which is ideal for eliminating occlusions, but requires a long acquisition time. In summary, the device is not yet on the market, quite expensive, and requires a long acquisition time. Moreover, only depth information is registered, introducing the need for a hybrid solution with a colour camera. Therefore, the Nano Point Scanner is not ideal to use as depth estimation device for the craquelure pattern of paintings.

### 2.1.8. Structured Light

Structured light is based on the disparity between the projector and the camera due to a different viewpoint. The technique has similarities with laser scanning discussed in section 2.1.3, where the position of a laser dot can be triangulated when the position and orientation of the source and observer

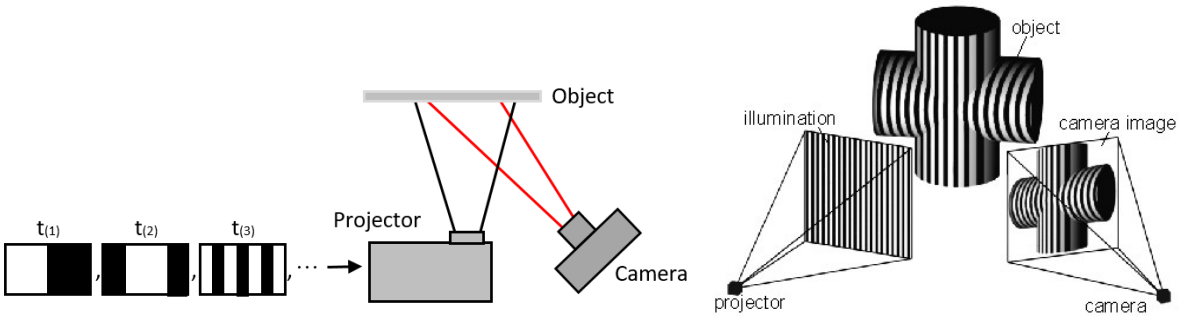



Figure 2.7: Structured light: The projector projects multiple structured patterns ($t_1$, $t_2$, ... $t_n$,) on the object. A camera with a different viewpoint captures the object with the projected structured pattern, which is perceived completely different due to the topography. The changed structured light is used to extract the depth information of the object. [62]

are known. In turn, the dot was replaced by a line to accelerate depth acquisition. With structure light, the laser line is replaced with coded structured light emitted by a projector [58][59][60]. Similar to triangulation with Line Laser projection the depth is determined with the line created by the projection. When viewed at a different angle, the topography changes the projected pattern, from which the depth information can be determined. An example is given of typical structured light in figure 2.7, however there is an abundance in varying coded structured light methods [61]. As can be seen by the multiple start-ups that use some kind of structured light to determine depth like Artec Eva and Einscan.

The depth precision and spatial resolution are limited by the accuracy of the position and orientation estimation and the resolution of mainly the projector. A high accuracy position and orientation estimation enables the device to reach the depth precision criteria. However, the spatial resolution is insufficient to meet the requirements, because even high resolution projectors would only be able to reach 50 µm (A.1.4). Furthermore, compared to laser line scanning technology, is the acquisition speed significantly increased, since multiple line can be projected. However, for good depth registration multiple images are required. In the end, the entire painting can be captured in 15 hours assuming 20 structured images are sufficient to reach the maximum depth precision. Moreover, structured light is an affordable and accessible technology since only a projector and camera are required, estimated to cost about 2500 to 10000 [63], depending on the projector.

Structured light simultaneously captures the depth and colour of the image, eliminating the need for a hybrid solution. However, the reconstruction of the colour image has a risk of slightly mismatching the structured projections for the colour image. Furthermore, the colour image is also restricted by the projector and camera resolution.

The ability of structured light to accurately determine craquelure is primarily limited by the shading introduced by the projector. The occlusions are minimised when the camera captures the painting perpendicularly,but that increases the shading from the projector. Moreover, the depth precision is proportional to the disparity between the camera and projector. Similar to stereo imaging, structured light has a trade-off between depth precision and registering craquelure due to shading. However, structured light is not dependent on two cameras and can therefore register twice as much depth (140 µm) for the narrowest craquelure (50 µm) as stereo imaging. Structured light is therefore a better option, but still not ideal to register the craquelure patterns of paintings.

### 2.1.9. Fringe Projection

Fringe projection uses phase shifting structured light and fringe pattern analysis to determine the height of the object. Similar to structured light, the disparity between the projector and camera, due to difference in viewpoint, is triangulated, as long as the location and orientation of the projector and camera are known. Instead of changing the pattern of the structured light the phase of a sinusoidal pattern is shifted [64][65][66], as can be seen in figure 2.8.

The phase shifting sinusoidal pattern projection is called stereo imaging. Structured light is base on black and white and thus the projectors resolution depict the spatial resolution. A smooth sinusoidal pattern has a unique value on each position, allowing for low sub-pixel resolution [67]. Analysing stereo

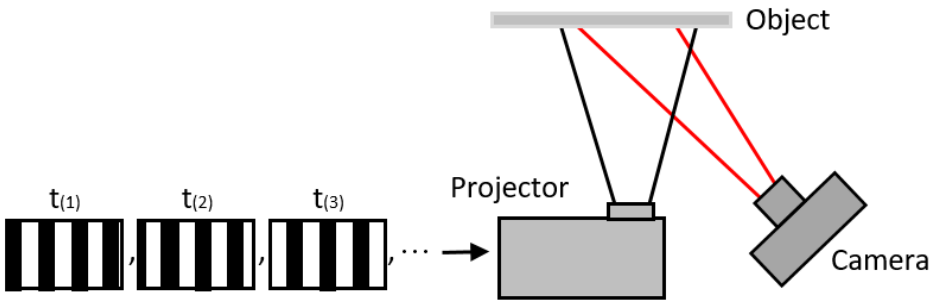



Figure 2.8: Fringe Projection: Similar to structured light, the projected pattern enable the extraction of topographical information, due to the disparity between the projector and camera. In stereo imaging, the projected pattern is a sinusoid, that is phase shifted.

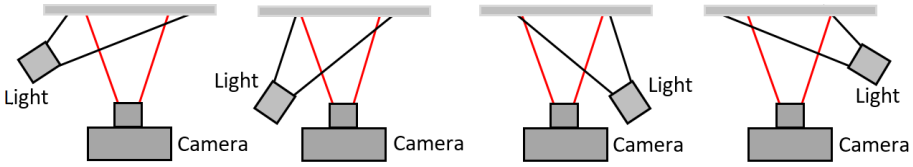



Figure 2.9: In Photometric stereo imaging, the direction of the light changes for each photograph. The shadows, highlights and gradient are used to extract the topographical information of the object.

imaging generates a wrapped phase map with signal to noise selection for higher accuracy. The phase map is then unwrapped into depth data to obtain the depth map of the object. All based on accurate position and orientation information of the observer and projector.

Current devices based on stereo imaging can reach sub-pixel resolution of the projector, however it is still limited by the resolution of the camera and the accuracy by which the orientation and position are estimated. Current techniques with stereo imaging can reach sub projector pixel spatial resolution of 25 µm [68][69], which is still insufficient to reach the criteria. In theory, a spatial resolution of 1 µm could be reached with a high-tier full-frame camera and high magnification lens.

The acquisition speed mainly depends on the required depth precision. A higher depth precision would require more phases, which requires more images to be captured. It is a trade-off between acquisition time and depth precision. Fringe projection requires roughly half the amount of images to achieve the same depth precision as structured light, because a sinusoid can be used to extract twice as much depth information. Moreover, the estimated costs of stereo imaging are lower compared to structured light methods, because it doesn't require a high resolution projector. Even for lower resolution projectors there still is a difference in pricing, therefore are estimated to cost 2500 - 5000 euro for stereo imaging.

The ability of fringe projection to capture the profile of craquelure is, like with structured light, limited by shading introduced by the projector. Equal to what is concluded in structured light, the technique is less ideal to register the craquelure pattern of paintings, especially the fine deep cracks are difficult.

### 2.1.10. Photometric stereo imaging

Photometric Stereo imaging determines the shape from shading by capturing multiple images with an illumination that varies in direction and incident angle as shown in figure 2.9. Shading, highlights and gradient hold information of the shape of the object [70][71]. Using these clues from a single image to extract the height map is an ill-determined problem. Multiple images each with a varying direction and/or indicent angle of the light source creates a solvable problem. From photometric stereo imaging a surface orientation map is generated and by integration of the gradients the height map is estimated [72].

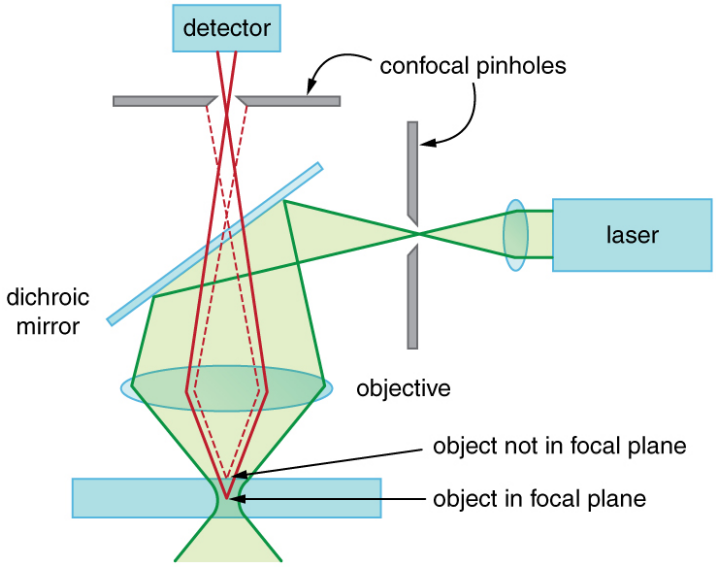



Figure 2.10: A Confocal Microscope directs the light at the object with a dichroic mirror. Two translatable confocal pinholes determine the focal distance, enabling the in-focus plane to be at different distances [77].

The true colour accuracy depends on how well the photometric clues can be removed. The depth precision is less than to stereo imaging. Stereo photometric imaging computes the height map with integration, which is highly sensitive to occlusions, discontinuities and miscalibration [73]. Moreover, the illuminate source is a point source and the surface a Lambetian surface, which is rarely the case in reality. The depth precision could reach accuracy of a camera pixel [74], which is about 4µm for a full-frame 50 MP camera.

Photometric stereo imaging is a quick technique. The acquisition time is similar to stereo imaging. Where stereo imaging required several images with a phase shift, this technique requires several images with a changing light direction. The requested depth resolution primarily depicts the amount of images, limiting the acquisition time. A device based on photometric stereo imaging would need an estimated 15 hours for the required depth precision. On the other hand, the technique only requires a camera and multiple lights, which makes it one of the cheapest techniques. A high resolution camera and multiple light sources are estimated to cost about 5000 euro. The colour and topography images are both captured with the stereo photometric imaging, eliminating the need for a hybrid solution.

The ability of photometric stereo imaging in registering the profile of craquelure is limited by the introduced shading, which is the primary source for depth estimation of this technique. In order determine the depth of the finest cracks, the angle of illumination sources has to be smaller than the 10° stated in the introduction. However, angle of illumination sources also depicts the depth precision since larger shading is easier to determine the depth from. Similar to other techniques, registering the craquelure is a trade-off with the depth precision. The ability of photometric stereo imaging to register the craquelure pattern is not ideal, especially since the artefact shading is the source of the depth estimation.

### 2.1.11. Confocal Microscopy

Commonly the working principle of a confocal microscope consists of capturing multiple images at different focal distances by changing the pinhole position as shown in figure 2.10. An algorithm then determines the areas in focus and uses those for the perpendicularly viewed colour map. The height map is constructed by the known difference in focal length between those areas in focus. In other words, a confocal microscope captures slices of the to be captured near-planar object [75][76].

The colour resolution is defined by the magnification of the lens attached to the confocal microscope. The height map accuracy depends on the predetermined step size and range of focal distances. However, the surface of painting usually has propagating curvature exceeding most focal distance ranges. Therefore, the relative height between painting surface and microscope needs to be controlled. Over-

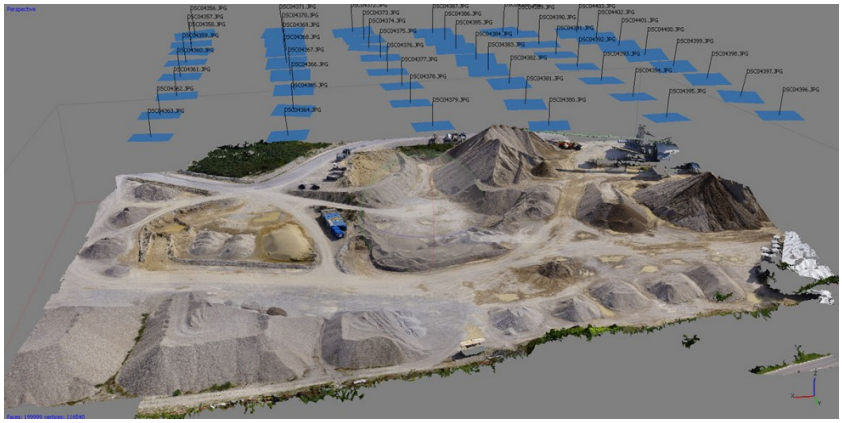


Figure 2.11: An example of photogrammetry used to create a 3D virtual scene of a construction site. The depth information is extracted by matching pixel neighbourhoods in several images [81].

all the spatial resolution and depth precision are well within the confocal microscopic range [78]. In essence, confocal microscopy computes depth similar to shape from focus and defocus, the technique requires a small Depth Of Field to determine depth differences. For shape from focus and defocus the depth of field was too large for the relatively flat surface of a painting, however a confocal microscope can reach much higher magnification where the painting's surface in relatively not flat. This also indicates the issue with confocal microscopy, to have accurate depth information a high magnification is required, which in turn results in a long acquisition time.

The acquisition time is determined by the amount of confocal images required to capture the height of the painting. Assuming the minimum height difference of 'the girl with a pearl earring' and the specifications of a popular confocal microscope [56], the entire painting could be captured within about 50 hours. At the beginning of this project there was no confocal microscope with a larger xy-stage range than 300mm [22], and therefore not capable to capture the entire painting, but such a xy-stage was developed for 'the girl with a pearl earring' by Hirox Europe. A confocal microscope, like all interferometry techniques, is quite expensive. For example a Hirox RH-2000 is estimated to cost about 100.000 euro [79], but they are very private about their prices. That is without the varying prices for microscope lenses.

The ability of a confocal microscope to register the profile of craquelure is good. The light source is projected perpendicular to the painting, minimising shadow forming. It is not a point measurement so occlusions might occur at the edges of the image, but these are minimal due to the small camera window. As mentioned the image has to be magnified a lot to obtain depth information about the craquelure profile, which results in a long acquisition time. In summary, the technique has enormous potential for depth precision and spatial resolution, but requires a long acquisition time and is expensive.

### 2.1.12. Photogrammetry

Photogrammetry is based on Structure from Motion in which 3D structures are extracted from 2D image sequences that are coupled with local motion signals. The matches of similar pixel neighbourhoods are identified in multiple photos and create points in the sparse cloud. If matching pixel neighbourhoods are found in two, or preferably more, photos, the areas occupied by the pixel neighbourhoods in the respective photos are projected into the virtual 3D scene. The sparse cloud of points in the virtual scene represents the real-world environment containing the imaging subject [80]. The technique is similar to stereo imaging, but it calibrates the changing camera position with the images, which are also used to determine the depth map with. An example of the technique is shown in figure 2.11.

Photogrammetry is great at digitising a solid 3D object, however it has issues with small depth details in the surface of the object, especially with near-planar surfaces. There are multiple factors that influence the depth from motion precision, like the algorithm tends to misread shadow of the object, intrinsic camera calibration or the angle and redundancy of the images. So in an ideal situation, a full-frame camera can reach depth precision of 50 µm at best [82]. On the other hand, the method is medium fast and cheap to implement, since only one camera is required. As with previous techniques, the spatial

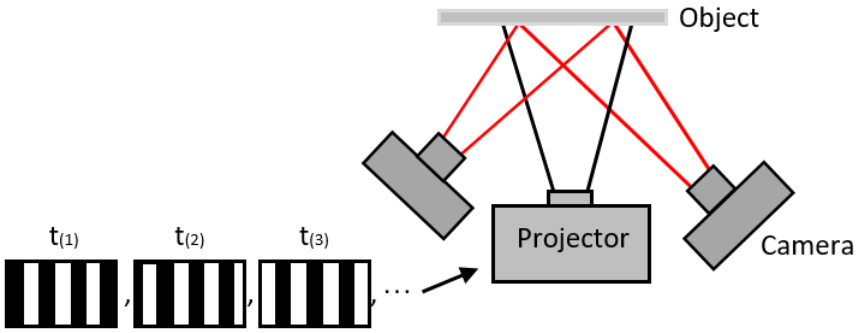



Figure 2.12: Fringe encoded stereo imaging: The projector projects the fringes on the object. The Fringe Projection is captured in stereo with the cameras. The Fringe Projection uniquely labels each position on the object, solving the correspondence problem of stereo imaging.

resolution is tied to the camera and thus is sufficient.

The acquisition speed depends on the capturing time and the amount of images required for the minimal resolution and accuracy of the 3D rendering. The technology is a very accessible to anyone with a camera and has been applied to many platforms and devices like for pixelio and 123D catch. To reach the theoretical depth precision of 50 µm it would take a very high overlap and thus in total more images than techniques as stereo imaging. Roughly estimated it would take about double the acquisition time of stereo imaging. On the positive side photogrammetry only requires a camera to capture the object and a computer to compute the 3D image. A camera able to reach the spatial resolution would be about 2500 euros.

The ability of photogrammetry to determine the craquelure profile is limited by shading and occlusions. An additional illumination source is required to register the depth of the crack, which introduces shading. Moreover, the difference in perpendicular position of the camera has to be much smaller than 5° to minimise occlusions, which is about 15 mm for a working distance of 200 mm. So to register the depth of craquelure, the painting has to be captured from multiple positions within the 15 mm. However, apart from the image exactly above the crack, most of these images still contain some occlusions reducing the accuracy of the craquelure profile. In summary, photogrammetry is a more economical technique and medium fast, but it is not as accurate in surface and craquelure registering.

### 2.1.13. Fringe encoded stereo imaging

Fringe encoded stereo imaging is a hybrid solution, that uses fringe projection as encoder to solve stereo imaging. The projector projects the fringe patterns onto the painting and two cameras capture the pattern in stereo. Fringe Projection processing is done by sparse stereo matching to encode the images. After correcting for the phase, the encoded images help solve the correspondence problem in stereo matching. This results in an accurate topographical map and colour image. The technique is primarily used to reproduce a painting with 3D printing [83][84].

Combining stereo imaging and Fringe Projection increases the accuracy of stereo imaging by using the fringes to match and triangulate. Furthermore, the hybrid solution gives Fringe Projection another viewpoint to estimate the distance with more certainty, resulting in a remarkable depth precision of 10 µm for existing devices [3]. The resolution of the colour and topography images is limited by the lenses of the cameras. The greater the magnification ratio of the lenses the more detail the sensors will capture. Furthermore, the spatial resolution in height map increases proportionally to the magnification, but the height range decreases also proportionally. As long as the height of the near-planar surface stays within the depth of fields of both cameras, the height can be determined. The current devices based on this method can reach 25 µm spatial resolution and 10 µm depth precision [85], which is insufficient. In theory, a full-frame camera and magnification lens would be able to capture the object at about 1 µm spatial resolution and 0.5 µm depth precision (A.1.2).

As with Fringe Projection, the acquisition speed is limited by the depth precision, since it influences the amount of images required. The device needs double the amount of time of Fringe Projection, because it requires horizontal and vertical fringes to uniquely label every position on the object. There-

fore it is estimated it would take 15 hours to scan 1750 $cm^2$, based on capturing speed of the current device [85]. Moreover, the hybrid solution requires two cameras and a projector making it less economical than other camera based techniques. The estimated costs are 10.000-15.000 depending on the cameras, lenses and projector quality. However, the technique was developed at the Technical University of Delft and can therefore be provided without any costs.

The ability of fringe encoded stereo imaging to estimate the profile of craquelure is limited by shading and occlusions. The depth is estimated with stereo imaging and thus, as mentioned, occlusions occur and for finest craquelure a depth of 70 µm can be registered. The projector introduces shading on the painting's surface, but it is minimal since the projector projects perpendicular to the painting. In summary, fringe encoded stereo imaging has a high depth precision, spatial resolution and is quite quick, but the technique is sensitive to occlusions and thus less ideal to register the profile of craquelure.

## 2.2. Comparison

All techniques in this chapter are compared on the criteria mentioned in the beginning with some notes. The main performance criteria are the depth accuracy and resolution of the technique, the acquisition speed considering an entire painting and finally the cost of the technique. The metrics are defined as their theoretical maximum potential.

| Technique | Spatial Resolution [µm] | Depth Precision [µm] | Acquisition Speed [h] | Estimated Costs [€] | Craquelure [+/-] | Notes |
|---|---|---|---|---|---|---|
| Ultrasound TOF | 100 | 5 | 5 | >100.000 | - | |
| Ultrasound interferometry | 4-10 | <1 | 100 | 50.000 | - | |
| OCT | 4-10 | <1 | 100 | 40.000 | + | |
| Line Laser TOF | 100 | 100 | 100 | 500 | - | |
| Line Laser triangulation | 10 | 0.4 | 50 | >1000 | - | |
| Stereo Imaging | 1 | 0.5 | 1 | 9.000 | - - | Available at TU Delft |
| Shape from Focus | 1 | 50 | 14 | 5300 | - - | |
| Conoscopic Holography | <1 | <1 | 100 | 50.000 | + | |
| Depth from Chromatic Aberration | <1 | <1 | 100 | 50.000 | + | |
| Structured Light | 1 | 50 | 15 | 8000 | - | Available at TU Delft |
| Fringe Projection | 1 | 25 | 7.5 | 5000 | - | Available at TU Delft |
| Photometric Stereo | 1 | 4 | 15 | 5000 | - - | |
| Confocal Microscopy | <1 | <1 | 50 | 100.000 | + | |
| Photogrammetry | 1 | 50 | 30 | 2500 | - - | |
| Fringe encoded Stereo | 1 | 0.5 | 15 | 12.000 | - - | Available at TU Delft |

Table 2.1: A holistic view of different areas of techniques that can be used to capture topographical information. The techniques are accessed on their their theoretical potential.

The techniques with a camera as observer have a maximum theoretical spatial resolution of 1 µm with a small pixel size and highest magnification macro lens. Some of the techniques are available at institutions, making the estimated costs for that technique an ignoreable criterion.

From table 2.1 can be deduced that the devices capable of sub-microscopical spatial resolution and depth precision are very expensive and require a long acquisition time. Excluding the techniques unable to reach the required spatial resolution and depth precision, leaves stereo imaging, photometric stereo imaging and fringe encoded stereo imaging. Stereo imaging was argued to be unsuited for a painting scanner, because of its sensitivity to oclussions and non-salient areas. From the remaining two techniques the fringe encoded stereo imaging holds the most potential for resolution, acquisition time, and requires no initial cost since the technique has been made available by the TU Delft and Océ, a Canon company. The technique is expected to perform less ideal at registering the profile of the deepest craquelure due to occlusions. However, as stated in the introduction, 'The girl with a pearl earring' is a relatively flat painting minimising the depth of craquelure. In addition, little research has been done on the measurement of craquelure, it is difficult to estimate the exact performance of the technique.

## 2.3. Discussion

In the beginning of my previous project, the possibility arose to use one of the confocal microscopes (Hirox RH-2000) from the Rijksmuseum in the Netherlands. I could use the microscope to test and eventually build a frame that enabled the microscope to scan the entire painting. Since the microscope can achieve much greater magnifications than photography based technologies, vibration would be my main focus. The structure has to be very rigid and minimise self-induced vibrations.

However, about four months in the previous project, it came to our attention that Hirox, the confocal microscope company, had been developing a very similar device as I had been working on. They had already made a frame that would enable the microscope to scan a 1 by 1 meter area. Furthermore, as mentioned in section 2.1.7, Hirox was developing a Nano Point Scanner (NPS) that would be able to scan the curvature and depth of the painting's surface. In summary, Hirox had already developed the stage I was researching. Since there is no point in inventing the wheel twice, the focus was shifted to a device that can capture the craquelure pattern.

The fringe encoded stereo imaging technique was developed by the Delft University of Technology. Océ/ Canon developed a professional device based on the technique called Fine Arts Scanning Technologies (FAST). Océ donated one of their older FAST models to the Delft University of Technology for further research. The current technique has been developed in order to reproduce paintings with a 3D printer. Thus a spatial resolution of 50 µm is sufficient since the human eye [86] can not distinguish higher resolution. So the potential of the technique is not yet unravelled.

## 2.4. Conclusion

This chapter researched a holistic view of different areas of techniques that can be used to capture topographical information and accessed them on the criteria stated in the beginning. Based on the comparison it can be concluded that the hybrid of fringe projection and stereo imaging holds the highest potential. Apart from being able to reach an ultra-high spatial resolution and high depth precision, the main advantages are the low acquisition time and having state of the art cameras and high tech frame directly available at no initial cost. However, the potential needs to be unlocked. The next chapter analyses the technologies in depth, to determine how best to meet the requirements to capture the craquelure pattern of oil paintings.

# 3

# Fringe encoded stereo imaging

Chapter 2 concluded that fringe encoded stereo imaging is the best technique to capture the craquelure pattern of paintings. The aim of this chapter is to analyse the technology further in-depth, to determine how the spatial resolution and depth precision of resp. 25 µm and 10 µm can be enhanced to reach at least 10 µm and 5 µm to capture the craquelure pattern.

The first section reviews the settings and calibration, essential for stereo imaging and fringe projection. Followed by an in-depth review of both technique separately. Thirdly, both techniques combined as fringe encoded stereo imaging is discussed. After which the hardware of the hybrid technique is discussed.

## 3.1. The basics: Settings and Calibration

The camera settings are essential to capture in-focus and well-exposed images. A colour and camera calibration is required to accurately reproduce the colour and topography of a painting. This section discusses the settings and calibration before the capturing process starts.

### 3.1.1. Setup Fringe encoded Stereo Imaging

Figure 3.1 illustrates the setup of fringe encoded stereo imaging. The setup is a combination of stereo imaging with 2 cameras and fringe projection with one camera and a projector. The tilt angle $\alpha$, shown in the figure, enables the camera to focus on flat objects at an inclined angle, which is discussed in detail in section 3.5.3. The distance to the painting (L) depends on the working distance of the camera lens. The distance between the cameras (B), also called the Baseline, depends on the camera angle relative to the painting's normal vector ($\beta$), which is covered in the next section. The exact angles and distances depend on the hardware that will be discussed in chapter 5.

### 3.1.2. Camera and Lens Settings

The cameras are positioned at the same height and directed at the same point on the object. The painting is placed at the focal distance of both cameras to obtain the highest magnification. The camera angles relative to the painting's normal vector ($\beta$) do not need to be exactly equal or known, however the camera angles have influence on the accuracy of the depth registration. Larger camera angles relative to the painting ($\beta$) results in a more accurate the depth triangulation, however it also increases the chance of occlusions [45], since there are more areas that are not registered by both cameras. This trade-off between depth precision and occlusions is important for capturing craquelure with fringe encoded stereo imaging and will be discussed further in chapter 5.

After positioning the cameras, the lens and camera settings need to be optimised to get the painting in-focus. A MTF board, shown in figure 3.2, simplifies the focus calibration, but any high resolution image could be used. The MTF board is placed at the minimum focus distance and the centre area is brought in-focus [87]. Next, the tilt angles (explained in section 3.5.3) are adopted such that all corners are in-focus.

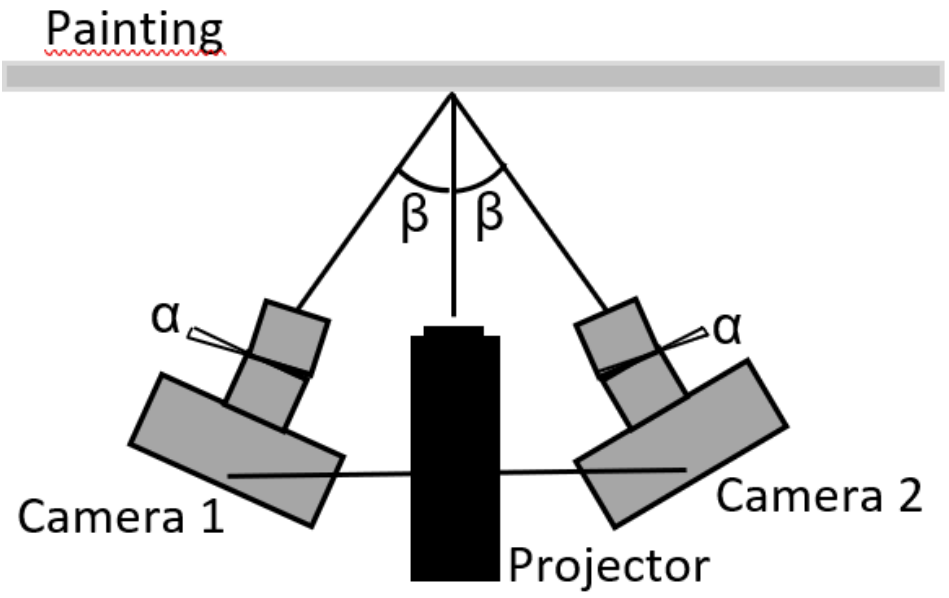



Figure 3.1: The setup of fringe encoded stereo imaging seen from above. The distance to the painting (l) and space between the cameras (b). The angle of the cameras relative to the painting's normal vector ($\beta$). The tilt-shift lens is tilted with the tilt angle ($\alpha$) to focus on the painting's surface

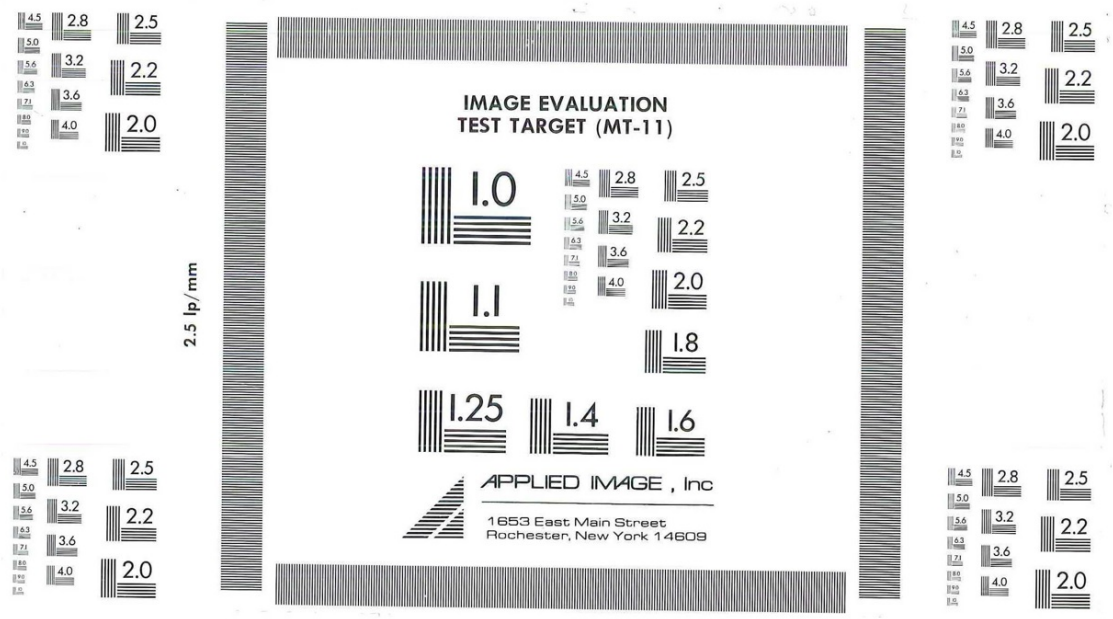



Figure 3.2: MTF board is used to determine the spatial resolution for the entire image. It is a semi-standarised method to test a camera and lens quality for the entire image [87]

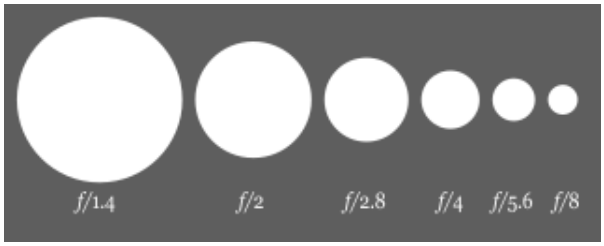



Figure 3.3: The diameter of the lens is defined with the aperture number. As illustrated the aperture is given by f/n, where n is the aperture number and f the focal length [92].

The White Balance (WB) removes unrealistic colour casts, so that the perceived white is equally white in the captured images. A camera White Balance takes the 'colour temperature' of the light source into account. Most digital cameras have basic 'warmth' or 'coolness' pre-settings for the White Balance corrections. The most accurate digital White Balance is set with a custom White Balance, because it sets the White Balance to the current environmental lighting condition. At the beginning of the calibration process, the custom White Balance is set by a capture of an optical whiteboard or grey chart, and then manually selected on the cameras [88][89]. The light condition of the environment should be constant to obtain a constant image quality and high colour accuracy.

After the sharp focus plane has been aligned with the painting and the White Balance is customised, the correct exposure time has to be determined to get most information without overexposing the photograph. Exposure is an interaction between the shutter speed, aperture and ISO.

The light intensity is called the ISO on cameras and is preferably kept low, since the higher the ISO , the more noise is introduced [90]. Therefore, the aperture and shutter speed remain to determine the exposure. The aperture number is inversely proportional to the diameter of the camera lens, as illustrated in figure 3.3. In other words, a larger aperture reduces the amount of light striking the camera sensor. A larger aperture number also increases the Depth Of Field (DOF) [90], but when the aperture number becomes larger than 13 [91], diffraction can occur, because the diameter of the lens becomes too small.

The shutterspeed can be adopted to still achieve the exposure time for a salient image. In the case of a large aperture number the shutter speed is lowered to get sufficient exposure. Moreover, the correct exposure differs per camera and lens, so it often is an iterative process to find the optimum of aperture and shutterspeed.

### 3.1.3. Camera Calibration

Stereo imaging requires an accurate model of each the camera and the relative position and orientation to each other, to triangulate the depth of the painting. The values required to generate the camera model are called intrinsic and extrinsic parameters.

The intrinsic parameters describe the focal length (f), scale factors ($k_i$), the skewness ($\tau$), and the optical centre ($o_i$) of the camera with a pinhole model shown in equation 3.1. The pinhole model describes every point on the sensor to a location in the real world. Furthermore, the pinhole model can then be used to correct the radial lens distortions. However, when a tilt-shift lens is used, the pinhole model does not define the system perfectly [93][94].

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} fk_x & \tau & o_x \\ 0 & fk_y & o_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \tag{3.1}$$

The extrinsic values describe the position and orientation of the camera in a 3D space. They can be defined in a rotation matrix (R(3x3)) and a translational matrix (T(3x1)). With both the intrinsic and extrinsic values accurately determined by camera calibration, each pixel from the camera sensor be related to the beam of light it was struck by.

There are several photogrammetric calibration methods to estimate the camera parameters [95]. Cur-

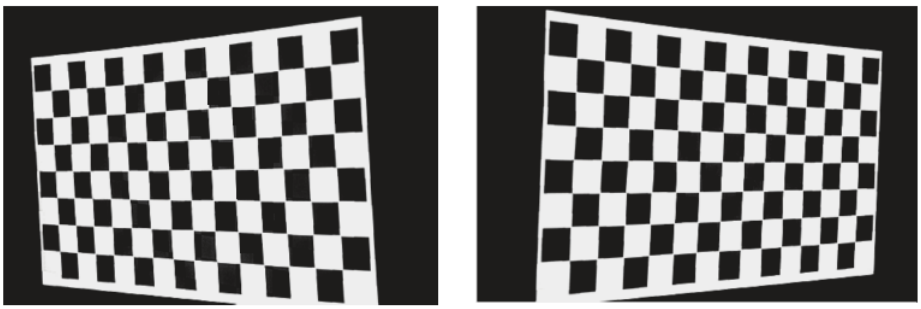


Figure 3.4: Camera calibration: An example of the photographs of the left and right camera during the camera calibration. The orientation of the checkerboard is changed and thus every image will be different. During the calibration the four corners are selected from which the orientation and position are determined.

rently, a flexible closed-form solution is used with a nonlinear refinement based on likelihood criterion [96], based on the intersections of checkerboard's squares. The size of the squares of the checkerboard are used to determine the intrinsic and extrinsic parameters of the cameras. Multiple images with altering orientation of the checkerboard will increase the accuracy of the calibration of the cameras. An example of stereo capture of the checkerboard is given in figure 3.4. The corners of the checkerboard are selected by hand, but recent techniques automated the checkerboard calibration [97]. As mentioned, the flexible closed-form solution does not yet include the tilt-shift lens, which is possible with current technology [94][93].

### 3.1.4. Projector Calibration

The phase shifting sinusoidal fringe pattern allows an exact geometric reconstruction of the surface shape, when the position and orientation from the projector relative to the camera are known. Therefore, an accurate calibration of the projector is essential for fringe projection.
The position and orientation of the projector relative to the camera is inversely computed compared to the camera calibration. A checkerboard is projected onto a flat surface like a small white board. Like in camera calibration the orientation of the board is changed for each image. To increase the accuracy of the coordinates of the projector relative to the camera. The projected checkerboard can be combined with the camera calibration, to create an quick calibration method. Nowadays there are many similar methods that achieve higher accuracy calibration, like sub-pixel edge detection method [98] or accurately determination of marker points of a calibration target in terms of projection transformation [99][100].

For the fringe encoded stereo imaging, the fringe projection is used as encoder and not as technique to determine depth. Therefore, is the projector calibration irrelevant for fringe encoded stereo imaging. However, the projector could used to determine the distance and orientation of the painting relative to the scanner, which will be discussed in chapter 5.2.5.

### 3.1.5. Colour Calibration

Reproducing objective colour information is essential for an accurate digitisation of art work. However, colour reproduction depends on various factors like the illumination source, the colour space, or the display of the device. For example, a RGB image is going to look different from one display device to another. The RGB colour space is a device-dependent colour space and thus less ideal, but is still often resorted to due to its ease in comprehension. This selection discusses some common colour spaces and ICC profiling as general colour space.

#### Colour Space

There are multiple colour spaces, each developed with a different goal in mind. RGB colour model is a linear combination of the red, green and blue primaries, which is usually displayed as a cube. As mentioned, it is the most comprehensive colour space, but device dependent [101]. Almost all cameras capture their photographs in the RGB colour space or a similar version of the RGB space. Therefore, the photographs are often converted into another colour space to better fit its purpose. For example, when the colour scans are used to (3D) print the painting, the colour images need to be converted

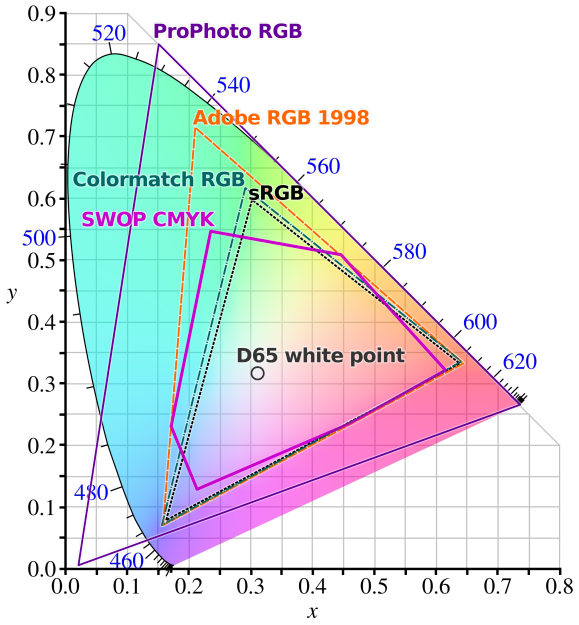



Figure 3.5: The CIE chromaticity diagram with all visible chromaticities. Several colour spaces are marked in the horse shoe shaped CIE region to indicate their colour space. In the CIE chromaticity diagram is D65 marked, which is one of the white point.

to the CMYK colour space. The CMY(K) colour system is used in printers, in which the abbreviations stand for black (K), cyan (C), yellow (Y), and magenta (M). The manner in which the inks creates colour is called subtractive colour mixing. The inks selectively absorb wavelengths of incident light and reflect the remainder [102][103]. HSV colour model can be divided in three independent dimensions corresponding to hue (H), saturation (S), and value (V) [104][101] This colour model is commonly used in image processing and editing software. This colour model has two visualisation representations. One is a cylinder with black at the bottom and full-intensity colours on the top. The other is a representation by a cone, with black at the peak and white at the base. Since this colour space focuses on image processing, it fits well with the image processing part of the device, but the final goal is to digitally visualise the painting on a a computer monitor. Other common colour systems are sRGB, developed for monitors, printers and the internet, or Adobe RGB which improves the gamut of sRBG in cyan-green. Adobe RGB is also primarily used for monitor display [105][101]. Therefore, would these colour spaces better fit the end goal of this literature report.

The International Commission on Illumination or abbreviated CIE (Commission International de L'Eclairage) is responsible for standardisation of data in colour space. They created device-independent colour models like the CIEXYZ, CIELAB , and CIELUV [106]. CIE-XYZ colour model of colorimetry is created based on visual experiments of colour matching. The colorimetry values depend on the observer's field of view. Therefore, the CIE defined a standard observer and a set of three colour-matching functions, called X, Y, and Z [107]. CIELAB colour system is designed to approach a perceptually uniform colour space in terms of colorimetry values of XYZ. The CIELAB space is believed to be useful for evaluating perceptual error [108].

In the end, can the digital reproduction of the painting be used for varying purposes, like 3D printed reproductions, image processing, image analysis, or digital display on a monitor. The goal of the scanner is to capture and digitally visualise the craquelure pattern of paintings. The preferred colour space to digitally visualise the data on a monitor is sRGB or adobeRGB. The input colour space of the cameras is raw RGB thus the colour information often requires a colour convertion. Therefore, has the colour information be able to convert to any colour space, which brings us to the colour management system. We chose for the most popular colour management system for converting colours: the ICC-profile [109], which is widely used for printing and colour visualisation on monitors.

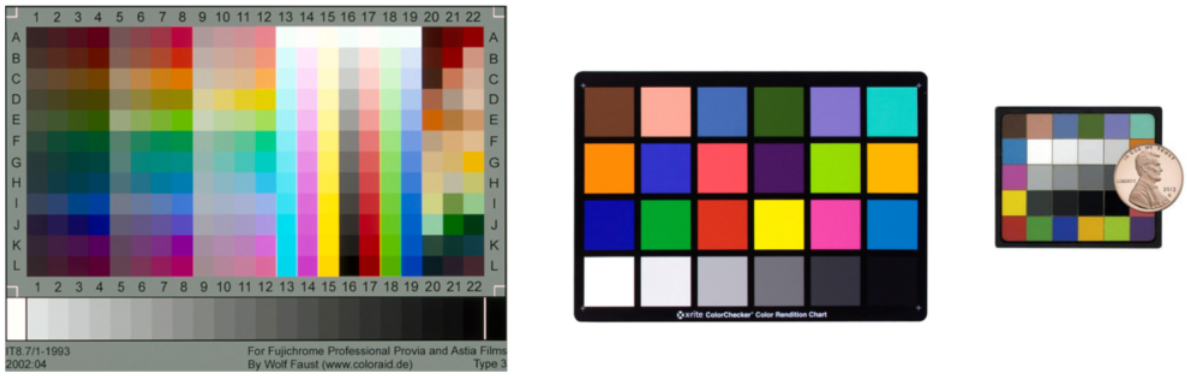



Figure 3.6: Examples of colour checker charts to calibrate colour used to create an ICC profile. The images of the colour checker chart determine the difference to the true colour. The small colour chart on the right is ideal for fine art scanning, since the camera window is too small for the larger colour charts.

### ICC Profile

Depending on the application, colour spaces have their individual optimised uses. However, to objectively capture the colour information of the paintings, independent of the environmental lighting conditions, the colour information should be transferable to any colour space. The International Colour Consortium (ICC ) has been successful in standardising device independent colour spaces between displays, printers, and capture devices [110]. ICC profiles store colour transformation profiles to and from different colour devices [111]. The current technologies uses the IIC profile to convert the colour images in to the CMYK colour space for 3D printing [112]. The ICC profile is determined by capturing a colour chart, from which the difference between the known colour value of the colour chart, and the captured colour value is computed.

There are several colour checker charts that are used to determine an ICC profile, shown in figure 3.6. Each chart differs in purpose or focus of colour system. The colour chart on the right in figure 3.6, is a small colour chart for fine art scanning that uses higher magnification photography, since the larger colour chart do not fit within one photograph when in focus. The colour checker chart is captured and the colorimetry values of each colour patch is extracted. From the extracted colour values the colour transformation profiles are constructed as ICC -profile.

To obtain higher colour accuracy, the conditions should remain constant during the entire capturing process. Therefore, a dark room is preferred to retrieve accurate colour information. The colour accuracy is measured with the distance metric ΔE, which is computed by the average difference of the captured CIELAB values compared to the reference values [113]. However, the transformation of the RGB to CIELAB with ICC profiling is complicated since it depends on various factors like the illuminate, gamma and observer angle. To make colour calibration and its performance measurements even more complicated, the ΔE metric knows different methods of computation that result in significantly different values. The way of computation has changed over the decades due to the fact that the lab colour space turned out to be not as perceptually uniform as intended, which modern ΔE computations account for. Most common and simple is the '1976' method, which is just the Euclidean distance between two lab values. In the end, we chose a more accurate and complex colour accuracy metric the ΔE '2000' version of the metric [114][115], which is also more perceptually uniform. To make a clear distinction, values obtained with the latter computation will be referred to as Δ-2000.

## 3.2. Stereo imaging

After the basic settings and an accurate camera and colour calibration, image capturing process of stereo imaging can be initiated. The stereo images are matched by their corresponding areas. Matching the areas can be done with two types of matching algorithms; sparse or dense matching. From the location differences of the matching areas, the disparity map is generated. The disparity map is used to determine the depth map of the painting. Each component of stereo imaging is discussed in the following sections.

### 3.2.1. Dense Stereo Matching

Dense image matching aims at computing the best match for each and every pixel of an image [44]. This facilitates the generation of an accurate and highly detailed digital surface model. There are several variations in dense stereo matching algorithms like difference in cost functions [116], locally or globally minimisation, or window size [117]. However, most dense algorithms come down to optimising cost computation.

The performance of dense stereo matching methods is highly dependent on the captured images. Dense matching will attempt to match each area of both images, but will fail when the images differ, are not salient enough, or have repetitive areas. Inevitably differences occur in stereo imaging due to occlusions, slight colour differences or spectral reflection. Furthermore, most paintings contain areas with little salient information, like the background. So just matching the entire images to one another is not ideal for painting scanning.

### 3.2.2. Sparse Stereo Matching

Sparse stereo matching obtains and matches features (keypoints) in both images [43]. Features can be edges, corners or gradients in the images. From all tentative keypoint matches only the valid matches are selected based on geometry.

The performance of sparse matching depends on the amount of features. Similar to dense matching if areas do not have salient information the match and thus the height can not be determined. The speed and accuracy of sparse stereo matching is high [118]. Furthermore, sparse matching is more resilient to mismatches due to occlusions, colour differences and irregularities. Therefore, spare stereo matching is more suitable for processing images of paintings. In the end, we decided to used the speed of sparse stereo matching to retrieve initial matching areas. These matching area are used to speed up dense stereo matching and obtain a dense and accurate disparity map.

### 3.2.3. Disparity Map

Figure 3.7 shows the working principle of the disparity map to the depth information. Disparity is the difference in location of an object in the image [41][42]. Objects closer to the observers have larger location differences than objects in the background. You can measure the apparent motion in pixels for every point and make an intensity image out of the measurements. In figure 3.7a the objects in the foreground are brighter, denoting greater location difference.

Figure 3.7b illustrates the ray tracking of a matched pixel or area. The camera calibration determined the position and orientation (extrinsic) of the cameras relative to each other, and the lens behaviour (intrinsic) of each camera lens, which construct the ray tracking map. The matching area or pixel is triangulated with the ray tracking map by both cameras into a point in space. Doing this for all pixels results in a depth map of the painting's surface. So simply put, the depth position (Z) can be determined with equation (3.2) from the disparity map (d in pixels), when the focus length (f) and baseline (B) are known [119][45].

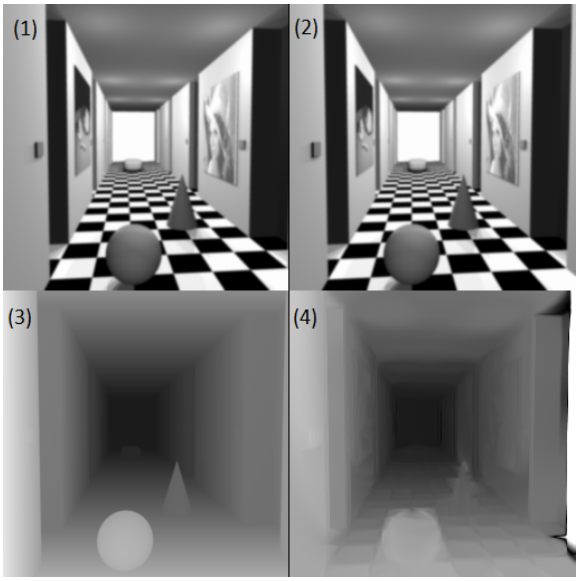
$$Z = B \cdot f / d \tag{3.2}$$

### 3.2.4. Discussion stereo imaging

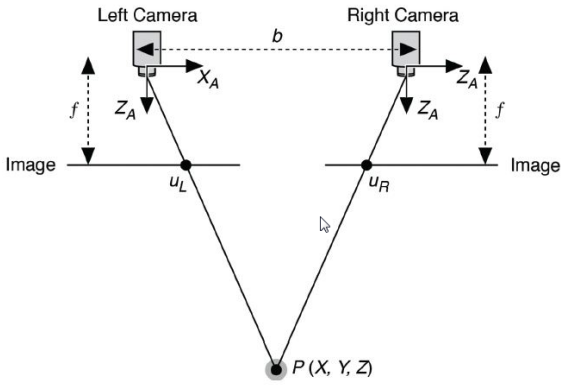
Stereo imaging computes depth by triangulation of the matching features in both images, but when the object does not contain salient information, no features will be found. A painting contains multiple areas where there is insufficient information for features, like the background in paintings. One could increase the visibility of available features by increasing the exposure time of the cameras or by creating new features to match by using a projector to encode the painting.

## 3.3. Fringe Projection

As discussed in chapter 2, fringe projection is similar to the structured light method. Fringe projection is faster and more accurate than typical structured light and is less limited by the resolution of the projector. Moreover, instead of projecting beams of black and white, fringe projection uses a shifting sinusoidal pattern called fringes. Since fringe projection uses phase shifting sinusoid to measure the topography or profile of the surface [64][65][66], is it also called phase shifting profilometry.




(a) Digital example of Stereo vision [119]




(b) Disparity geometrics [120]

Figure 3.7: (a) The top figures (1,2) show a simulated stereo pair and the bottom figures the exacted disparity map (3) and the computed disparity map (4). Similar to how human eyes determine depth, a larger difference in position in the images, means it is closer to the cameras. (b) The depth information is generated when the focal length (f), baseline (b) of both cameras and the disparity/distance between corresponding points ($U_L$ and $U_R$) are known.

There is a large variety in fringe patterns each with different results. This section first discusses the parameters creating a fringe pattern, the influence of the parameter and the chosen pattern. The second part reviews the fringe processing, which can be summed up as exacting the wrapped phase map from the captured images and unwrapping the wrapped phase map.

### 3.3.1. Fringe pattern

The amount of fringes and the fringe parameters; amplitude (modulation), wavelength (frequency), and off-set defining the sinusoid, influence the digitisation of the painting. In practise, fringes are not a continuous sinusoid due to among others the non-linearity between the cameras and projector, which is partially corrected by the gamma correction. Furthermore, multiple fringe projection can also be used to enhance the accuracy of the depth estimation.

#### Modulation

In theory, the modulation of a 8-bit sinusoidal fringe can vary from minimal 0 (black) to maximal 255 in grey-scale, since 8-bits has $2^8$ = 256 unique values. In practise, this is rarely achieved due to; display settings, GPU conversion of the projector, or inaccurate dynamic ranging of the camera, resulting in flattening of the peaks of the sinusoidal. Therefore, the modulation is kept within a safety margin of about 20 to 230 to ensure a smooth sinusoid, but a lower modulation gives a lower signal-to-noise ratio (SNR ) in wrapping algorithm.

#### Gamma Correction

The GPU conversion of the projector and inaccurate dynamic ranging of the camera can also result in non-linearity of the fringes [121]. For example, a projector can have a larger amount of different darker colours than lighter colour, to enhance the experience of movies with a darker theme. This results in over representation of darker colours and thus a deformed sinusoid. The gamma correction counters the non-linearities with a non-linear response in the other direction. The effect of the gamma correction on the input and the resulting sinusoid is illustrated in figure 3.8.

The required gamma correction is accurately determined by analysing the wrapped phase map, from which the deviations can be used to correct future wrapped phase maps. If the gamma is incorrectly applied, it causes banding in the RGB image [122]. Another option is to convert the raw un-demosaiced image into a RGB profile that has a linear gamma [123]. However, the projector and GPU contribution would still need be estimated from the fringe images. Therefore, the gamma correction is determined

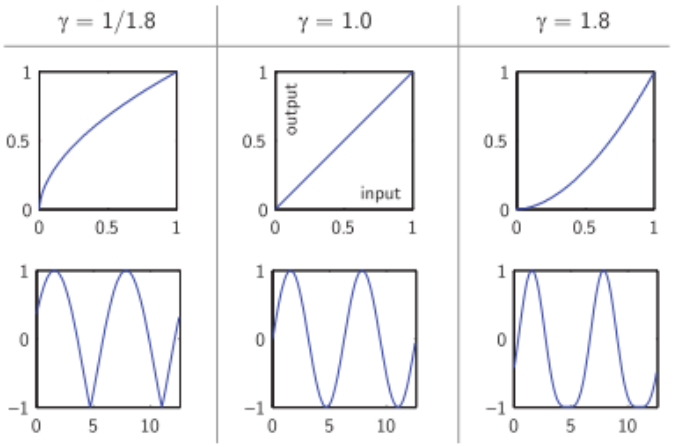



Figure 3.8: The gamma correction on the input signal, resulting in a non-linear output. The effect of the gamma correction is shown in the figure below the correction. The gamma correction counters the non-linearities due to the conversion of the Gpu and the projector or inaccurate dynamic range of the cameras. The figure is adopted from [3]

from the wrapped phase map. We determined that a gamma correction of 1/1.3 gave the best wrapped phase map and depth estimations.

#### Wavelength

The wavelength determines the distance between each fringe. A smaller wavelength results in an increased spatial resolution and the signal-to-noise ratio of the depth registration. A larger signal-to-noise ratio will distinguish subtle depth differences from the larger error deviation. fringe projection assumes the projected pattern is a continuous sinusoid, which can't be achieved for small wavelengths due to the limited resolution of projectors. In the end, a wavelength of 12 pixels was chosen since it was the smallest wavelength that still gave good results, even independent of the projector.

#### Amount of fringes

The off-set of the sinusoid is related the amount fringes, since more fringes equals more sinusoids within one wavelength requiring a smaller off-set. A larger amount of fringes lowers the dependency on a continuous sinusoid, but will introduce corresponding harmonics. The number of fringes is proportionate to the acquisition time, since each fringe need to be captured. Reducing the number of fringe will lower the acquisition time, but at least three phases are required to solve the wrapping algorithm [124].

#### Double step

The double three phase fringe projection method averages the result of two three phase fringe patterns. One of the fringe patterns has a off-set of one-sixth of a phase ($\pi/6$). The double three phase method performs well in practise and also reduces harmonics [125]. The study found that the double three phase method without correcting for the gamma performed better than the single-three step method with different accurately modelled gamma corrections. The three phases fringe pattern did not consistently produce a dense depth map even with the double step. Therefore the 6 phases fringe pattern was chosen for its more consistent results, but without a double step to reduce scanning time.

### 3.3.2. Fringe projection processing

Fringe projection, or phase shifting profilometry, is a less comprehensible depth perception technique than stereo imaging. Therefore, figure 3.9 shows a simplified side view of the working principle, to make the technique more understandable. The fringes are wrapped together into the wrapped phase map, containing the depth information of the painting. The depth is determined from the change in the fringe pattern due to the topography, when observed from a varying viewpoint. The wrapped phase map is still unrecognisable due to the arc-tangent operation (eq. (3.4)), which is used to compute the wrapped phase map [126]. Finally, the wrapped phase map is unwrapped and the carrier-component is removed to create the depth map, as illustrated in figure 3.10 with a simulated example.

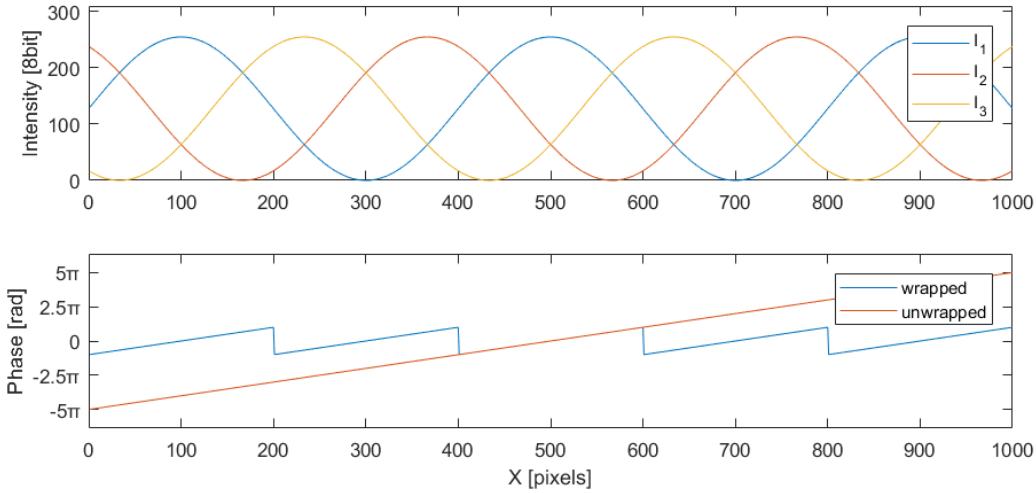



Figure 3.9: A side view of the working principle of a three phases fringe projection. The phases are wrapped together, creating the wrapped phase map. An unwrapping algorithm removes the arc-tangent operation, resulting in the unwrapped phase map. The bottom figure shows the wrapped (blue) and the unwrapped maps (orange) with the carrier-component.

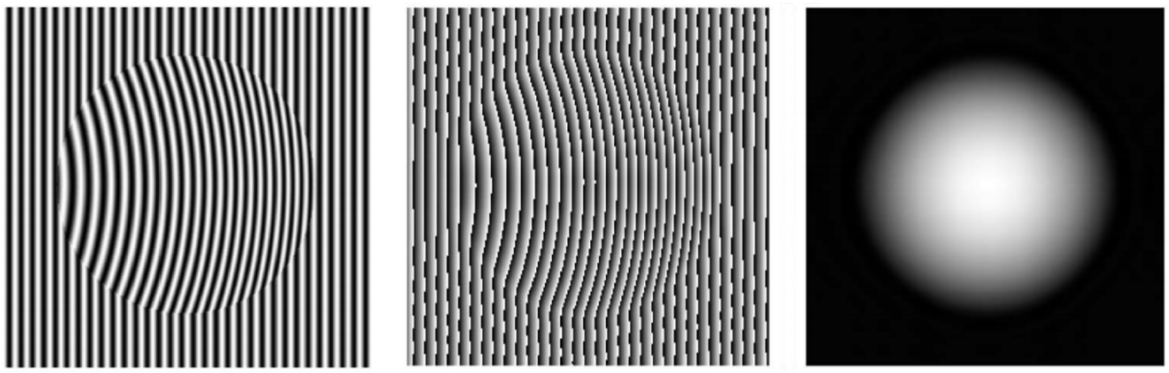


Figure 3.10: The entire fringe projection process shown with a simulated example. Firstly, the object with the fringe projections captured. Secondly, the three or more photographs of the fringes are wrapped together into the wrapped phase map. Lastly, the arc-tangent operation and carier-component are removed with an unwrapping algorithm, resulting in the depth map. The figures are adopted from [126].

The phase shifting profilometry algorithm is experimentally complex, but requires little processing time. When compared to other techniques like Fourier transform profilometry[127][128], continuous wavelet transform [129] or short-time fourier transform [67]. In phase shifting profilometry, the intensity-profile fringe pattern captured by the camera is defined with the following equation

$$I_n(x,y) = A(x,y) + B(x,y) \cdot cos(\phi(x,y) - 2\pi n/N) \tag{3.3}$$

where A ( x, y) is the average intensity and the background illumination. B (x,y ) is the intensity amplitude of the pattern contrast. The phase-shift index is defined ed as n = 0, 1, 2, ... , N-1, and $\phi(x,y)$ is the corresponding wrapped phase map which can be extracted [130] by the following equation

$$\phi(x,y) = \tan^{-1}\left(\frac{\sum_{n=0}^{N-1} I_n(x,y) \cdot sin(2\pi n/N)}{\sum_{n=0}^{N-1} I_n(x,y) \cdot cos(2\pi n/N)}\right) \tag{3.4}$$

As mentioned, since there are three unknowns in the wrapping algorithm, at least three phases are required to solve equations (3.3) and (3.4). As illustrated in figure 3.9, the wrapped phase map ranges from $-\pi$ to $\pi$ due to the arc-tangent operation. There are several phase unwrapping approaches as multifrequency (hierarchical)[131][132], multi-wavelength (heterodyne) [133][134] and number-theoretical [135][136], of which multifrequency is the most reliable unwrapping approach [137]. Phase unwrapping is in essence a process concerned with traversing through the wrapped phase vector sequentially in the x direction and adding or subtracting multiples of $2\pi$. This results in an unwrapped phase map which is given by (3.5), where $\phi$ (x y, ) is the unwrapped phase and k (x ,y ) is the integer number to represent fringe orders.

$$\Phi = \phi(x,y) + 2\pi k(x,y) \tag{3.5}$$

Finally, the carrier-components need to be removed from the unwrapped phase map. As seen in figure 3.9, the carrier keeps increasing the unwrapped phase. The carrier component is removed by fitting a plane to the wrapped phase map.

### 3.3.3. Discussion Fringe Projection

Fringe projection requires a continuous projected sinusoid, but the resolution of the projector causes the sinusoid to discretise, especially for smaller wavelength fringe patterns. To solve this the projector is set just a bit out of focus, which is similar to applying a Gaussian blur. The projected fringes approach a near continuous sinusoid, but some modulation is lost.

Furthermore, fringe projection is usually done with a projector and a camera. Increasing the number of cameras with different viewpoints will enhance the accuracy and reduce the occlusions. Similar to double step, for multiple cameras the depth and colour maps are constructed and averaged together into one more accurate image. However, non-linearities from the projector will persist in the final result.

The phase information from fringe projection can be used to encode the image. In the case of multiple cameras a matching area is required in all images from which the carrier removal begins. That way the absolute unwrapped phase map is generated from equal location for all cameras, but the algorithms should be able to handle the non-linearity due to difference in viewpoint [138]. If the fringe projection is used as encoder, each spatial location has a unique value, which could also solve the correspondence problem in stereo imaging. In practise, horizontal and vertical fringes are required to give each pixel a unique value. This hybrid solution of stereo imaging aided by fringe encoding eliminates the need for projector calibration and gamma correction, since the depth estimation is based on stereo imaging. The next section will review the hybrid solution in depth.

## 3.4. Fringe encoded stereo imaging

Combining stereo imaging and fringe projection solves the primary limitation of each technique. Stereo imaging reaches the required depth precision and spatial resolution, that fringe projection can not

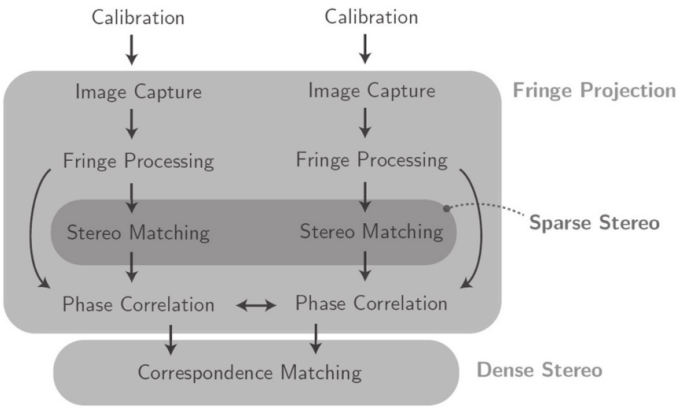



Figure 3.11: Both cameras photograph the object with the projected fringes and are then separately wrapped. The wrapped phase maps are unwrapped with multifrequency. The unwrapped phase maps are matched with sparse matching, from which the carrier component removal is initiated. Finally, the disparity map is generated by dense matching both unwrapped phase maps, from which the topographical and colour maps are computed. Adopted from [85]

achieve. Fringe projection solves the correspondence problem by encoding the surface of the painting, but this does require horizontal and vertical fringes.

The basic settings and calibration for the hybrid method are equal to the calibration explained in section 3.1. Apart from the projector calibration, which has become unnecessary, because the depth estimation is not dependent on the projector's 3D location.

After the cameras are calibrated and positioned well, the phase shifting fringes are projected and captured in stereo. The amount of fringes increase the depth precision and spatial resolution, but also doubles the acquisition time. As mentioned, the fringe encoded stereo imaging requires horizontal and vertical fringes to uniquely label each spatial location in the image. Thus the total amount of images is doubled when the amount for every added phase. Therefore, it is a trade-off of accuracy and acquisition time.

### 3.4.1. Combined Image Processing

The combined image processing uses the fringe projection as encoder to aid the stereo imaging. Figure 3.11 gives a broad overview of the entire image processing. After camera and colour calibration are set, the phase shifting fringe pattern is captured for with a minimal of three phases. The captured images are wrapped with the fringe projection algorithm discussed in section 3.3.1. The wrapped phase maps are unwrapped with multifrequency. The unwrapped phase maps are matched with with sparse matching, from which the carrier component removal is initiated. Finally, the disparity map is generated by dense matching both unwrapped phase maps, from which the topographical and colour maps are computed [85][3].

### 3.4.2. Discussion Fringe encoded stereo imaging

The fringe encoded stereo imaging has the potential to capture the craquelure pattern of paintings. Stereo imaging determines the topography and colour images. In stereo imaging, the depth precision is proportional to the spatial resolution, if the angles of the cameras relative to the painting remain constant, this will be explained in depth in the next chapter. Moreover, the spatial resolution in stereo imaging is set by the resolution of the photographs produced by the camera and the lens. So for fringe encoded stereo imaging to capture the craquelure, the resolution of the camera and lens need to be analysed.

## 3.5. Hardware

The hardware properties limit the range of potential of the technology. The main hardware components involved are the cameras, tilt-shift and lens, the projector, the xy-stage and the polarisation filters. In this section these components are discussed.

### 3.5.1. Projector

The projector creates the fringe pattern on the painting's surface and therefore influences the digitisation of the painting. This section discusses the preferred properties of a projector.

The projector should be able to focus at close distances of 0.1 to 0.5 meters, since the working distance of the cameras is within that range. Moreover, the projector preferable fits in the space between the cameras, which differs from 0.15 to 0.3 meters depending on the working of the cameras. The projector could project from a different position than in between the cameras, but that requires complex corrections of the projection and complicates the design of the scanner. Another option is placing the projector on its side to fit within the cameras and rotating the projection.

#### Intensity

The intensity of a projector, measured in ANSI Lumens, increases the modulation and thus the signal-to-noise ratio, but it requires a larger dynamic range of the camera. In other words, the device is less sensitive to the environmental illumination, which could open the possibility to scanning in daylight. Moreover, a higher intensity reduces the required exposure time and thus lowers the acquisition time of each capture. However, a high light intensity can damage paintings, and therefore the musea lighting is kept at about 50 lux (lumen per $m^2$).

The Dutch Foundation for Enlightenment Science [139] states that for a common restoration, a lamp with an intensity of 1000 lux is used for 4 weeks 10 hours a day. In frequency of occurrence, scanning a painting can be compared to restoration. For restoration they assume a constant enlightenment of the entire painting for 200 hours. According to their assumptions and computation, illuminating a small part of the painting for only 5 minutes, accounting for overlap, would allow a light source of 600.000 lux. So the intensity of a projector is preferably kept low, but it is not limiting since all commercial projectors stay well below 600.000 lux [63].

#### Resolution

The importance of the resolution of the projector is minimised, since the depth is determined stereo imaging instead of fringe projection. However, fringe projection is still used to encoded the surface of the painting. Fringe projection assumes a continuous sinusoid, which is achieved by defocusing the projector.Defocusing projector will decrease the modulation, but is necessary for the fringe processing. Moreover, if the projector is insufficiently defocused, the fringe pattern is not fully removed from the colour and topographical maps. A high resolution projector is favoured, because it requires less defocusing to achieve a continuous fringe.

#### Colour information

An object is perceived when light strikes the object. How the object is perceived depends on the illuminate and the surface of the object. Paintings are perceived by the emitted and reflected light from its surface. The reflection are removed with a polarisation filter, further discussed in 3.5.5. To perceive the colours of the painting in true value, the illumination source should emit a spectrum similar to the spectrum of daylight. The spectrum the projector emits is therefore important for the colour accuracy of reproduction of a painting [140][141]. Some of the spectra of light sources are shown in figure 3.12. The light source that emits the fullest spectrum are the projectors based on a halogen lamp, but these lamps degrade much faster. For fringe encoded stereo imaging the projector is preferably placed between the cameras, which limits the size of the projector. Most small projectors are based on LED lighting, which mainly excite 3 channels, red, blue and green, to resemble all colours. This will decrease the colour accuracy, fortunately the colour calibration somewhat corrects for the difference in light spectrum, but a complete spectrum is preferred [106].

#### Projector Requirements

Currently, the scanner uses a LED pico projector with a spectrum that is a combination of the cool and warm white LED spectrum. The highest intensity setting is rather approaches a cool white LED, but still only reaches 100 ANSI lumens [85], which requires a long shutterspeed of 5 to 10 seconds to get a salient image. The long exposure time makes the device slow and sensitive to the environmental illumination. Therefore, the projector should be replaced with a high intensity (>1000 Lumens) projector to reduce the shutterspeed and to possibly enable daylight scanning. Furthermore, it can be deduce

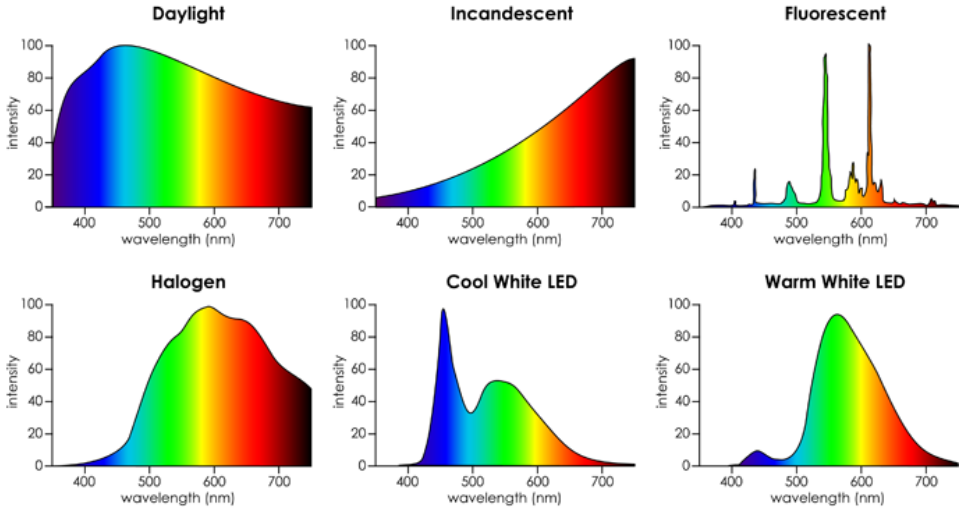



Figure 3.12: The spectrum of a light source is shown as the intensity of a light source as function of the wavelength. Projecting the object by a light source with a fuller spectrum produces a more colour accurate perceived object [142]. A halogen based projector is considered to reproduce the highest colour accuracy.

from the current projector that a fuller spectrum than LED is not essential for the technique to work, but it would enhance the colour reproduction accuracy. Similarly the projector's resolution is currently not limiting the technique, but an increase would only increase the smoothness of the fringe pattern.

### 3.5.2. Cameras

The cameras determine with what detail the fringe pattern and the painting's surface is captured. The settings discussed in section 3.1 greatly influence the captured information, but there is more than camera settings, like White Balance and exposure time. The camera sensor primarily depict the capabilities of the camera, like the dynamic range or the spatial resolution.

Each pixel of the camera sensor converts the amount of photons per time unit into an electronic signal. The difference between the minimum and maximal measurable amount of photons is called the dynamic range [143]. The smallest difference measurable within the dynamic range is called the just noticeable difference (JND). The just noticeable difference influences salience of the features, since more differences result in more features.

Furthermore, the size of the sensor and pixel density directly determine the spatial resolution and the acquisition time. A higher pixel density means more pixels registering the same object and thus a higher spatial resolution. The size of the sensor determines the acquisition time, because a larger sensor with the same pixel density, requires a lower amount of images to capture the painting.

So ideally the camera has a lot of pixels within a small area with a large dynamic range. However, the dynamic range is difficult to determine as a specification of the camera, since it depends on several factors, like camera settings, illumination source, or non-linearities in the dynamic range. Therefore, the focus is the resolution and pixel pitch of the camera sensor. The next chapter will discuss the camera sensor further in depth.

### 3.5.3. Tilt-shift adaptor

In stereo imaging, the cameras are positioned at an angle relative to the surface of the painting. Tilt-shift adaptors, based on the Scheimpflug principle [144][145], allow the Depth of Field to tilt and shift relative to sensor plane, as shown in figure 3.13. Tilting the depth of field would enable the in-focus plane to align with the inclined painting. Being able to tilt the depth of field requires a considerable smaller DOF to get the entire painting in focus compared to not using a tilt-shift adaptor. Since a set-up without a tilt-shift adaptor would need a larger aperture number to get the entire inclined plane in-focus, which requires a longer exposure time to get the same information.

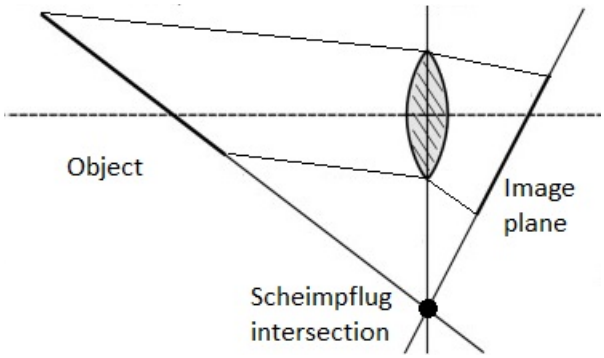



Figure 3.13: Simplified example of the scheimpflug principle of a tilt(-shift) lens [144]. In the case of this image, the light is bend faster at the top of the lens than the bottom, changing the normal vector of the imaging. In practise, the tilting is done by a set of lenses to obtain the scheimpflug principle.

Furthermore, the lens is often fixed to the tilt-shift adaptor, because their optical centres have to be accurately aligned. Therefore, a tilt-shift lens is preferred, since the alignment has been done during production. This does limit the choice of magnification lenses. The greatest magnification for the state-of-the-art tilt-shift lens is currently 1:2. For most cameras this magnification is insufficient to reach the spatial resolution criteria.

One could determine whether a lens without a tilt-shift adaptor able to reach the spatial resolution could capture the painting within the acquisition time. Otherwise a capable tilt-shift lens could be enhanced with a extension tube or close-up filter to realise the spatial resolution. However, both solutions will influence the quality of the image. Next chapter will discuss the options further.

### 3.5.4. The xy-stage

The xy-stage supports the cameras and projector and enables the technology to scan the entire painting. The xy-stage moves parallel to the painting in horizontal and vertical direction. The xy-stage is build to withstand vibrations of the environment and self-induced vibrations from translation of the device. Abrupt and fast movements induce greater vibrations, thus commonly a balance between movement speed and acceptable vibration is selected. Moreover, the acceptable vibration becomes smaller when the spatial resolution is increased, since the vibrations are also enlarged when the image is magnified. Furthermore, the entire scanner with xy-stage has to be transported from and to the musea. Therefore, it is important to design a xy-stage that is not only rigid to withstand the vibrations during scanning, but also portable.

Often the importance of the control of the xy-stage is underestimated. In most projects the movement speed is reduced in order to stay within the device dependent vibration acceptance, but the self-induced vibrations can also be minimised with a well designed controller [146][147]. An optimal controller considerably reduces the settling time of a vibration, resulting in a shorter movement time and thus a lower acquisition time.

So the xy-stage has to be rigid enough to withstand the spatial resolution dependent environmental vibrations, but also portable to transport to musea. The self-induced vibrations can be minimised by optimising the movement controller.

### 3.5.5. Cross-Polarisation

The paint of paintings is protected in most cases with a layer of varnish. The varnish on the painting's surface reflects projected light in multiple directions, resulting in specularities varying per viewpoint. The specularities causes issues in stereo image matching, because the specularities differ in each camera and thus do not match. In theory, this can be solved by applying polarisation filters on the cameras.

Linear polarisation filters lets only wavelengths pass though that are linear to the filter as is shown in figure 3.14. The undesired specularities that do not align with the desired light are blocked. The polarisation filter can be screwed on the camera lens.

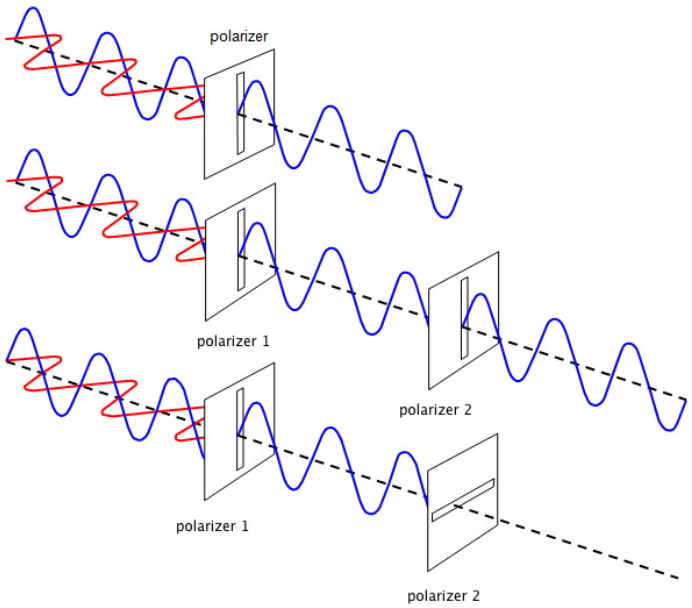



Figure 3.14: A polarisation filter only lets one direction of light through. So two polarisation filters in-serie would block all light when one is rotated 90 °, as shown in the lower example. Cross-polarisation works similar to the second example, where only 1 direction is let through. The emitted light from the projector is polarised and the light entering the cameras is polarised, filtering any specularities from the surface.

In theory, the camera polarisation filter blocks the specularities from the surface, but in practise, the projector introduces specularities, which in some cases align with the camera filter [148]. Therefore, cross-polarisation is used to eliminate specularities, which polarises the light from the projector [149]. A major drawback is that the polarisation filters reduce light intensity, lowering the modulation and thus signal-to-noise ratio, but it is essential for stereo image matching. The quality of the polarisation filter greatly influences the intensity of the light that passes through.

## 3.6. Conclusion

For a high colour accuracy the colour calibration and the entire scan should be done during a constant environmental lighting conditions. This is difficult to guaranty for a scan during the day, since the solar intensity differs during the day. Therefore, a dark room is preferred to obtain the highest colour accuracy.
In fringe encoded stereo imaging, the projector's resolution is not essential for the topography estimation, but does influence the amount of defocusing required to obtain a continuous fringe pattern. Less defocusing results in a higher modulation and thus a higher Signal to Noise Ratio. Moreover, a high intensity (>1000 Lumens) projector reduces the exposure time and overall acquisition time. So a high resolution and high intensity projector is favoured for the final design.
Reaching the spatial resolution and depth precision required to capture the craquelure pattern with Fringe encoded stereo imaging, can be achieved by adapting the hardware. The camera and tilt-shift lens are the limiting factor in this technique. The software can extracting the topography from the fringe aided stereo image for higher resolution photographs. Therefore, the next chapter will analyse the possibilities to adapt the camera and especially the lens to increase the spatial resolution from 25 µm to lower than 10 µm.

# 4

# Ultra-High Resolution

The previous chapter concluded that current fringe encoded stereo imaging achieves a high resolution of 25 µm (1200DPI ), which is insufficient to capture craquelure. The resolution required to capture the craquelure pattern is smaller than 10 µm (2500DPI ), which is near to microscopic resolution. So to emphasise the difference between high and the required resolution, is the required resolution defined as ultra-high resolution.

This chapter investigates the possibilities of the camera and lens to reach the spatial resolution and depth precision.

## Spatial resolution and depth precision

In fringe encoded stereo imaging, the spatial resolution and depth precision are connected. In theory, when the angle of the cameras relative to the painting are kept equal, increasing the spatial resolution will also increase the depth precision of the technique. For example, if the pixel density of the camera sensor is doubled, the depth precision is also doubled, because there are also twice as much depth points measured with, as illustrated in figure 4.1.

Assuming the angle of the cameras relative to the painting are kept constant, the geometrics state that the ratio of the spatial resolution to the depth precision is also constant. The current technique has a 25:10 ratio [85], thus a spatial resolution of 10 µm would result in a depth precision of 4 µm, which meets the requirements of the final design (1.1). So the focus of this chapter is to enhance the spatial resolution of fringe encoded stereo imaging.

## Camera and tilt-shift lens

As mentioned in the previous chapter, the spatial resolution is determined by the pixel density and the lens. The acquisition time is mainly influenced by the size of the camera sensor, assuming a constant pixel density. So when picking a camera sensor, the sensor size is leading over the pixel

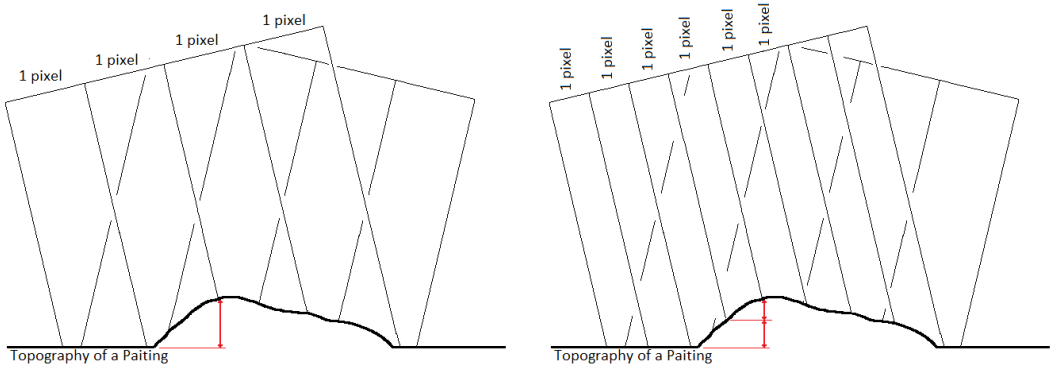



Figure 4.1: In stereo imaging, increasing the spatial resolution of the photograph with also increase the depth precision. When the spatial resolution is doubled, the precision by which the depth is determined is also doubled.

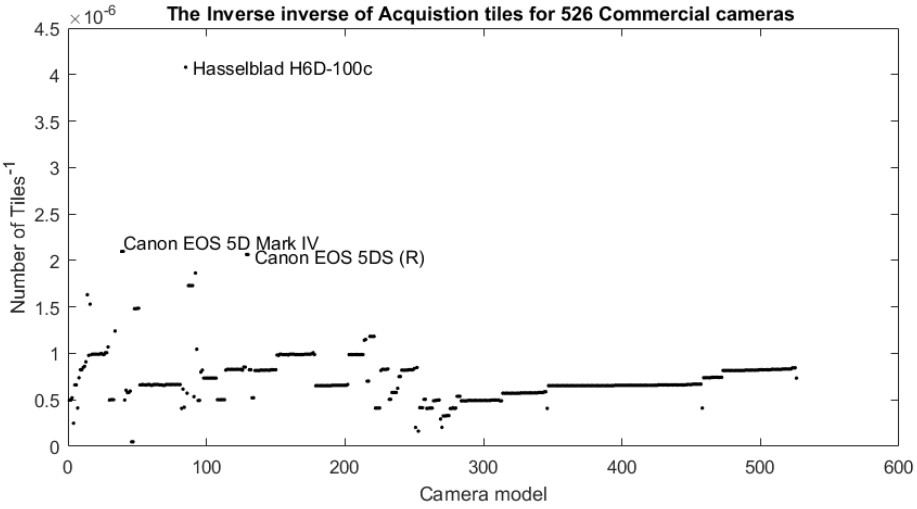



Figure 4.2: The inverted amount of tiles required to capture 'the girl with a pearl earring' based on the sensor size for several camera models [150]. A higher scoring camera model requires less acquisition time to capture 'the girl with a pearl earring'. The name of the three best scoring camera models are displayed in the graph.

density, because the spatial resolution can also be achieved with a magnification lens. In figure 4.2 the inverted acquisition time per camera model is shown. The acquisition time is proportional to the amount of tiles required to capture the entire painting. The amount of tiles is determined by the area of painting divided by the sensor size [150], and normalised to the required spatial resolution. The acquisition time is inverted to amplify the important information (A.2.1).

From the graph can be concluded that the Hasselblad H6D-100c achieves the lowest acquisition time, but each camera costs about 15.000 euros. The second closes camera models are the Canon 6D Mark II and the Canon EOS 5DS (R), of which the latter is currently available with a Canon TS-E 90mm tilt-shift lens at the university. The focus is therefore switched from the camera model and its sensor to the camera lens, to meet the spatial resolution requirement.

## 4.1. Requirements and Objective

The methods will be assessed on the theoretically spatial resolution in micrometers reached by the method. As mentioned in the introduction, the requirement to capture craquelure is a spatial resolution of 10 µm. The methods are also accessed by the costs attached to implement the method in euros. Furthermore, other practical information is taken into account as a note to the method, which are often difficult to quantify and/or define for each method. The objective is to achieve a spatial resolution of at least 10µm by enhancing the Canon 5Ds with or without the Canon TS-E 90mm tilt-shift lens, while minimising the costs of the method.

- Spatial resolution of 10 µm
- Minimise costs

## 4.2. Methods

The methods can be classified in two groups, replacing the Canon TS-E 90mm tilt-shift lens or adding parts to the currently available tilt-shift lens. The first subsections discuss the methods replacing the tilt-shift lens, of which options are limited, because lenses are often camera brand dependent. The second part consists of methods including the currently available lens.

### 4.2.1. (Macro) Lenses

The greatest magnification in photography is reached with a macro lens. There is a wide range of macro lenses available that can reach the spatial resolution, but a limited amount in combination with tilt-shift capabilities, and even less compatible with a Canon camera. This section therefore discusses

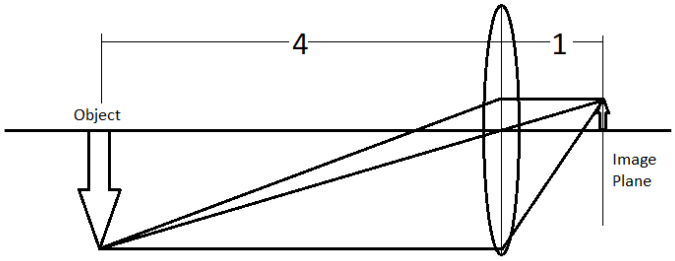



Figure 4.3: The working principle of a magnification lens. The magnification ratio of this lens is 1:4. The subject photographed in the real world is 4 times larger than received by the sensor [90].

Figure 4.4: Using a tilt-shift lens allows the alignment of the Depth Of Field with the painting's surface. The Depth Of Field can be smaller when aligned, increasing the aperture and thus reducing the exposure time [90].

high magnification macro lenses without a tilt-shift adaptor. This requires a larger Depth Of Field to get the entire window in focus, which means a higher aperture number and a longer exposure time. However, it is an interesting method to consider, because there are more options to choose from.

Compared to other lenses, a macro lens can enlarge the image of the object in the real world. In other words, the object is larger on the sensor than in the real world. The loss of image quality and distortions, due to the magnification, are minimised by specialised magnification optical lenses. They are able to focus on objects very close to the lens, which magnifies the image. However, high magnifications often result in (i) a small working distance, which is not preferred when scanning a painting, and (ii) in a very small Depth Of Field, which is undesirable especially without a tilt-shift adaptor [151].

In macro photography, the magnification is defined by the magnification ratio, which is the ratio between the size of the object on the sensor and the actual size of the object. For example, a magnification ratio of 1:2 means that an object of a pixel on the sensor, is 2 pixels in the real world. An example of a 1:4 magnification ratio is given in figure 4.3.

The greatest (compatible) magnification lens is the Canon MP-E 65mm f/2.8 1-5x, reaching a 5:1 magnification [152]. Theoretically being able to reach a spatial resolution of 0.826 µm (A.2.2). This high magnification lens costs 1140 euros per lens [153], resulting in a total of 2280 euros. There are more economical options for macro lenses with a lower magnification ratio still meeting the requirement.

A lens with a magnification ratio of 1:2, theoretically achieves a spatial resolution of 8.26 µm. There is a larger amount of macro lenses able to reach the 1:2 magnification, and is therefore a more economical solution. The price ranges from 300 to 400 euros per camera for the cheaper macro lens (Canon EF 50mm Macro Objective).

### 4.2.2. Tilt-shift Marco lens

Recently, the Canon TS-E 50mm f/2.8L MACRO tilt-shift lens has been introduced (12/2017 [154]). Similar to the macro lenses discussed in the previous section, it magnifies the subject with a certain ratio, while minimising the loss in image quality and distortions. The tilt-shift adaptor enables the Depth Of Field of the camera to be tilted and shifted to align with the painting, which decreases the aperture and thus reduces the exposure time. A tilted Depth Of Field is compared to a non-tilted Depth Of Field in figure 4.4. The combination of a macro lens and tilt-shift adaptor is quite unique, especially since it is compatible with the current camera.

This macro lens has a 1:2 magnification ratio, reaching a spatial resolution of 8.26 µm in combination


Figure 4.5: The close-up filter depends on the diameter of the lens of which it is screwed on. The close-up filters differ in refractive power indicated in dioptre (dpt).

with the Canon EOS 5Ds. However, since the lens is quite unique and new, the price per lens is 2200 euros. So implementing the tilt-shift macro lens in the fringe encoded stereo imaging would cost 4400 euros [154].

### 4.2.3. Teleconverters

A teleconverter is a set of lenses placed between the camera and lens. Even though it is typically used to increase a lens's focal length, it is also a popular way to increase magnification. The teleconverter multiplier directly correlates with the increase in magnification; a 2X teleconverter therefore doubles the maximum magnification. Disadvantages include a loss in image quality, introduces distortions, and an increase in the lens' minimum f-stop (max aperture of 32), which may prevent a camera from being able to auto-focus [155].

The teleconverter increases the magnification of the current set-up by a factor of 2. The current magnification approximately reaches 0.3x [156], so the with a teleconverter the magnification becomes about 0.6x. The spatial resolution is 6.883 µm with a 2x teleconverter. A compatible teleconverter costs 350 euro per piece, making teleconverters one of the more economical methods to increase the spatial resolution.

### 4.2.4. Close-up Filter

Close-up filters are special lenses that can be screwed onto the front of the lens like an camera lens filter. They are basically just a magnifying glass that is placed between your lens and the subject. It is for this reason that they are also often called close-up filters. Disadvantages include a loss in image quality and distortions, due to new optics introduced by the lenses [157].

The function of a lens differs greatly, and thus each lens has a differing design to best execute the function. The diameter of most lenses is therefore different, making the close-up filter dependent on the lens. Close-up filters have several variants with differing refractive power indicated in dioptre, as can be seen in figure 4.5. Dioptre is the inverse of the focal length of the lens. So similar to the teleconverter, a close-up filter increases the focal length of the lens.

Increasing the focal length with close-up filters also increases the distortions, therefore the minimum refractive power necessary is used as method. A close-up filter needs at least a 2 dpt to meet the spatial resolution as computed in the appendix A.2.3. Since a 2 dpt close-up filter gives a spatial resolution of 9.178µm, which just meets the requirement. The quality of the optics of the close-up filters not only influence the quality of the image, but also the price. A high quality close-up filter for the current TS-E 90mm lens costs about 140 euros per filter. So implementing the compatible close-up filters into the device would cost 280 euros.

### 4.2.5. Extension tube

The extension tube is placed between the camera and lens. The tube extends the length the light can diverge, as is illustrated in figure 4.6. The image is magnified on the sensor, but this also dilutes the light intensity. In other words, the image is magnified, but less light reaches the sensor. Therefore, a longer exposure time is required, but the exposure time is still small relative to the methods without a tilt-shift lens [158].

Moreover, the extension tube does not introduce new optics and thus only uses the optical system of the current lens. This reduces the loss of image quality and distortions compared to teleconverters

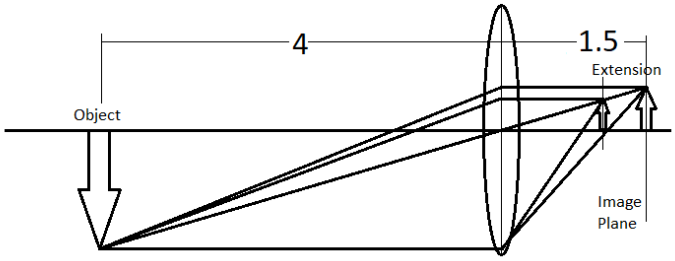



Figure 4.6: The extension tube is place between the camera and the lens to extend the time the light can diverge. A longer time to diverge results in an enlarged image on the sensor [90].

or close-up filters.

To meet the resolution requirement an extension tube of 25mm is needed. The 25mm extension tube enables the current system to reach a spatial resolution of 6.88 µm. Two compatible extension tubes for the current set-up would cost 270 euro in total.

## 4.3. Comparison

The methods can be classified in replacing the lens or implementing an attachment to reach the spatial resolution. The computations are based on the pixel pitch on the camera model Canon EOS 5Ds and in some methods on the Canon TS-E 90mm f/2.8 Tilt-Shift Lens. A holistic view of the methods is given in table 4.1. The depth precision is shown to give a more complete view, but is not part of the final assessment, since its value is directly related to the spatial resolution with a factor 2.5.

| Method | Spatial Resolution [µm] | Depth Precision [µm] | Costs [€] | Notes |
|---|---|---|---|---|
| Macro Lens w/o Tilt-shift (5:1) | 0.83 | 0.33 | 2280 | - Long exposure time |
| Macro Lens w/o Tilt-shift (1:2) | 8.26 | 3.30 | 600 | - Long exposure time |
| Tilt-shift Macro Lens (1:2) | 8.26 | 3.30 | 4400 | |
| Teleconverters | 6.88 | 2.75 | 700 | - loss of image quality |
| Close-up Filter | 9.18 | 3.67 | 280 | - loss of image quality |
| Extension Tube | 6.88 | 2.75 | 270 | - Less light |

Table 4.1: Different methods for enhancing the spatial resolution of the Canon EOS 5Ds camera with or without the TS-E 90mm f/2.8 Tilt-Shift Lens.

The highest quality image is achieved with the new Canon Tilt-Shift Macro lens, but at an high price. The table shows that the extension tubes are the best option to increase the spatial resolution for an economical price. Therefore, the extension tubes are selected as enhancement of the device.

## 4.4. Validation

The spatial resolution and depth precision achieved with the extension tubes need to be determined to validate the method. They will also be the primary specifications describing the device. This section discusses the methods to determine the spatial resolution and depth precision.

### Spatial Resolution

The spatial resolution can be defined as how the pixel pitch is related to the real world distance between pixels. In other words, what distance does the camera sensor measure in reality. Determining the spatial resolution over several pixels will give a more accurate result. One possibility is to photograph a ruler and count how many pixels fit in a millimetre, but a MTF board is preferred for photography. As mentioned in chapter 3 (fig. 3.2), a MTF board is used to determine the quality of the image, but

it can also be used to determine the spatial resolution for the entire image. So a MTF chart is used to determine the spatial resolution.

### Depth Precision

The depth precision is the precision by which the depth information can be determined by the device. In other words, the smallest difference in depth registered by the device. The depth precision is more difficult to determine, because the estimated depth precision is about 3 µm when using extension tubes. The depth precision can be computed by capturing depth information of an known object.

The most comprehensible method is using a staircase with thin layers to determine the depth precision. The thickness of the layers should be smaller than the estimated depth precision. The precision by with the staircase is digitally reproduced determines the depth precision of the device. However, the layer need to be thinner than 3 µm, which is far thinner than an average human hair of 60 µm [159] or a sheet of paper of 65 µm [160], making it not a viable method.

Another proposed method uses stacked layers of ink from a printer to validate the depth precision. Similar to a thin layer staircase, the layers the device can register. The thickness of one ink layer from a Ricoh MP C3001 reaches 2 µm, but it can vary up to 20 µm depending on several factors, like ink type, temperature, or type of paper [161][162]. So ink layers can not be used for depth precision validation, since they are not precise.

To validate the depth precision, the Contour GT-K 3D Optical Microscope is chosen to compare the 3D measurements with. The Contour GT-K 3D Optical Microscope has a precision of typically 0.02 nm (RMS) [163]. The same 3D section of a paint sample with craquelure will be scanned. The resulting height maps are optimally fitted and the difference in height depicts the depth precision of the fringe encoded stereo imaging device.

## 4.5. Discussion

The spatial resolution is computed with the theoretical specifications of the lens and camera, but these are determined in the ideal situation. So in practise the spatial resolution and thus depth precision will be larger than computed. Similarly, the spatial resolution and depth precision ratio is based on geometrics, which work in theory, but will not fully match the findings in practise. In practise, the results are influenced by multiple factors, like human error in placement and/or bringing the plane in focus.

The extension tubes are the best method based on the requirement. However, the new Canon tilt-shift lens is a good option, when the budget allows it. Interesting to note for future projects, is that the Canon tilt-shift lens can be combined with the extension tube to reach an even greater magnification, with limited loss of image quality.

Section 4.2.1 warned about the small Depth Of Field introduced by the higher resolution of the macro lens. The high resolution in general reduces the Depth Of Field, which can become an issue when the entire surface of the painting has to be within the Depth Of Field. In other words, the parallel area in which the painting is in focus has become thinner with the spatial resolution. This is an issue, since each part of the painting has to reach a certain amount of focus in order for the fringe encoded stereo imaging to work. The current painting scanner based on fringe encoded stereo imaging device has a working distance tolerance of ± 1 mm [164], so the tolerance will be even smaller than a millimetre when the resolution is enhanced. If the Depth Of Field becomes smaller than a millimetre, it becomes difficult to position the entire painting within the range of the Depth Of Field by hand.

## 4.6. Conclusion

This chapter concludes that the extension tube is the best method to enhance the spatial resolution of the current set-up. The extension tube is the most economical method with low image quality loss compared to the other methods.

Increasing the resolution reduces the Depth Of Field of the device. The entire painting has to be orientated in a thinner parallel area than for lower spatial resolution. Initial estimations place the Depth Of Field lower than a millimetre, which is difficult to position the orientation of the entire painting in the Depth Of Field by hand. So, the next chapter will investigate and discuss the Depth Of Field and the positioning of the device relative to the painting.

# 5

# Distance and Orientation Estimation

Chapter 4 concluded that increasing the resolution reduces the Depth Of Field (DOF) of the device. The device has to be positioned such that the painting is orientated within the Depth Of Field, which is determined first for ultra-high resolution fringe encoded stereo imaging. Next, the degrees freedom of the xy-stage are discussed, to determine the possibilities for positioning the device such that the entire painting fits within the Depth Of Field. The primary goal of this chapter is to determine what is the best method to estimate distance and orientation of the device relative to the surface of the painting.

## Depth Of Field

The Depth Of Field is the parallel area in front of the camera which is considered to be in-focus. The surface of the painting has to be in-focus to determine the depth information. Optical geometry can approximate the Depth Of Field of a single camera, as illustrated in figure 5.1. Where the aperture Number (N), focal length (f), lens-object distance (U), and Circle of confusion (C) approximate the total Depth Of Field ($D_1$ + $D_2$) with equation (5.1) [165]. The Circle of confusion, in this equation, is the circle that is still perceived to be a dot, which is dependent on the enlargement of the photograph, or the device the photograph is examined with. For example, the circle of confusion is larger when the photograph is seen with a projector than seen on a monitor, since the distance to the display is smaller. Commonly, the circle of confusion is the size of a pixel [90], since that is the smallest dot a the camera sensor can register.

$$DOF_{total} \approx \frac{2NCU^2}{f^2} \tag{5.1}$$

For ultra-high resolution fringe encoded stereo imaging, the Depth Of Field is tilted and reduced due to the extension tubes. Stereo imaging also requires both images to be in-focus and thus the DOF is

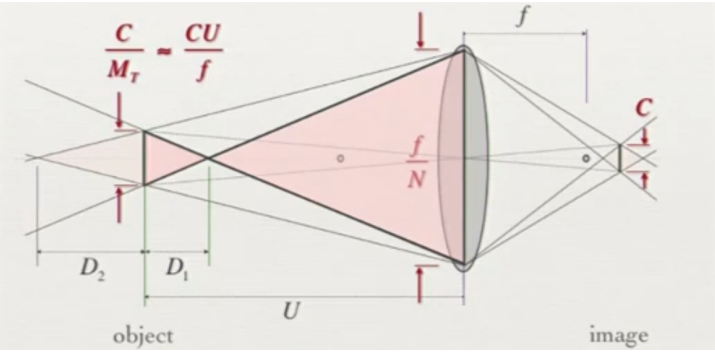
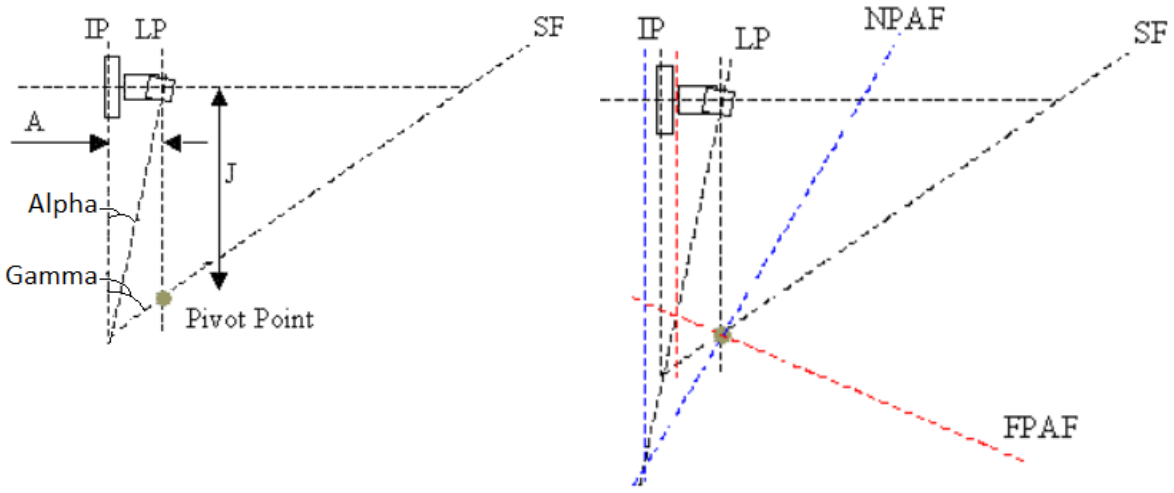



Figure 5.1: The Depth Of Field of a single camera can be determined with optical geometry. The geometry defines $D_1$ and $D_2$, which approximately result in equation (5.1). The total Depth Of Field is computed when the aperture Number (N), focal length (f) , Circle of confusion (C), and lens-object distance (U) are known. The image is adopted from [165].




(a) The geometry of a tilted in-focus plane. (b) A tilted Depth Of Field.

Figure 5.2: (a) The Image Plane (IP), Lens Plane (LP), tilt angle (Alpha), and focus length determine the Sharp Focus (SF) plane. (b) The Depth Of Field for a tilted Image Plane is illustrated with the Near and Far Plane of Acceptable Focus (NPAF/FPAF), which are determined by the acceptable focus Image Planes, based on the Circle of confusion. The images are adopted from [144].

where the DOF of both cameras overlap. Figure 5.2a illustrates the Sharp Focus (SF) plane, which is tilted from its parallel position. The Depth Of Field for a tilted focus plane is determined, similar to the DOF for a single camera, by the Circle of confusion, since it depicts the near and far acceptable image plane. The near (blue) and far (red) acceptable Image Plane (IP) determine the Near/Far Plane of Acceptable Focus (NPAF/FPAF), as illustrated in figure 5.2b.

The intersection of the Sharp Focus plane and the Lens Plane is called the pivot point. The distance (J) between the pivot point and the perpendicular Image Plane, and the angle Gamma of the Sharp Focus plane and the Lens Plane (LP) is determined with equations (5.2) and (5.3) [166][167]. An extension tube increases the distance between the Image Plane and the Lens Plane and is therefore added to the distance (A).

$$J = \frac{f}{sin(Alpha)} \tag{5.2}$$

$$Gamma = tan^{-1}\left(\frac{sin(Alpha)}{cos(Alpha) - f/A}\right) \tag{5.3}$$

As mentioned, the Depth Of Field of stereo imaging is where the DOF of both cameras overlap. Figure 5.3 shows an example of the shared area in-focus. Interesting for the positioning of the device relative to the painting is the minimal DOF and thus the smallest the Near and Far DOF. Assuming the cameras are positioned perfectly, one camera can be used to determine the Near and Far DOF.
With the pivot point-IP distance (J), the Circle of confusion, and the size of the tile the Near and Far Depth Of Field can be determined with triangle geometry. The NDOF and FDOF are resp. 0.369 mm and 0.370 mm, determined with the computations in A.3.1. The total Depth Of Field, where the entire painting's surface has to be positioned in, is 0.739 mm for a Circle of confusion of one pixel.

So to capture the entire painting in focus, the scanner has to be positioned parallel to the surface of the entire painting at the working distance within a range of certainty of 0.739 mm. The working distance is about 250 mm for the setup with extension tubes [154], which shows that the required Depth Of Field is very small compared to the working distance. This illustrates the difficulty of positioning the scanner parallel to the painting's surface, especially for the entire surface.

### Hardware Setup

In theory, the selected hardware meets the design requirements set in the introduction. Two Canon EOS 5Ds cameras, two TS-E 90mm tilt-shift lenses and the 25 mm extension tubes enable the scanner

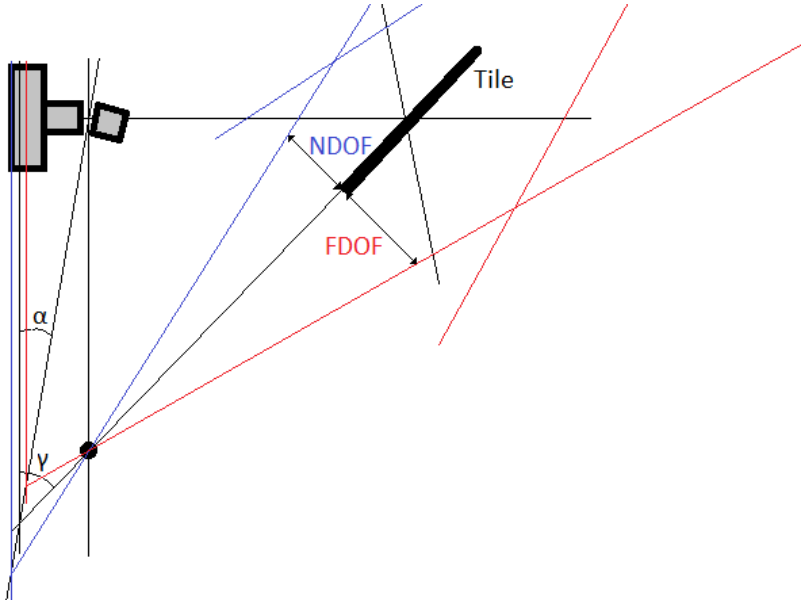



Figure 5.3: Determining the smallest Near and Far DOF for a camera with a tilt-shift lens and extension tubes. The Near and Far DOF enclose a triangle, which enables the computation of the total DOF.

to reach the required spatial resolution and depth precision to register the craquelure. However, as mentioned, the camera angle relative to the painting's normal vector ($\beta$), illustrated in figure 3.1, determines the trade-off between the depth precision and the ability to register the craquelure pattern of paintings. A greater camera angle increases the depth precision, but also increases the chance of occlusions. Moreover, the equations above (A.3.1) and the maximal tilt angle ($\alpha$ =8°) of the camera lens depict the maximal camera angle ($\beta$) at 23°, in which the entire image is still in focus. A camera angle of 23° is close to the examples of 20° used in chapter 2, which determined that a depth of 70 µm could be registered for the finest craquelure. Therefore we decided on a camera angle of 21.5°, because the depth precision is important for the depth estimation of entire surface of the painting and craquelure is just a small part of that. Furthermore, as mentioned in the introduction, there is little known about the dimensions of craquelure, it therefore makes little sense to focus on designing the scanner for deep craquelure that might not often occur.

### xy-stage

The current xy-stage can be controlled electronically in horizontal (x) and vertical (y) direction, enabling precise parallel translation in front of the painting. Furthermore, the xy-stage has the possibility to manually translate in the z-direction and manually rotate about the x-axis, as illustrated in figure 5.4. The rotation about the x-axis at the base is limited to a little over 90 degrees, to avoid the device tipping over into the painting. Furthermore, the xy-stage is unable to align the x axis parallel to the surface of the painting. This can be corrected by translating the device in z direction, as long as the orientation of the tile remains within the DOF. This would require an orientation estimation of the painting relative to the painting for every tile and the z-axis control for every tile change with an accuracy of about 0.1 µm (±5 samples within the DOF [13]).

Ideally, the x-axis and painting are aligned parallel within the DOF of 0.739 mm, when the device is placed in front of the painting. The alignment of the y-axis can be achieved with the rotation about the x-axis. This would require the estimation of the orientation and the alignment of the device to be done only once at the beginning instead of for every tile, reducing the acquisition time significantly. In either case the orientation of the painting's surface has to be determined to position the device parallel to the painting within the 0.739 mm.

## 5.1. Requirements and Objective

The orientation of the painting relative to the device has to be estimated, to determine whether the entire painting is within the Depth Of Field. The orientation is estimated by measuring the distance to

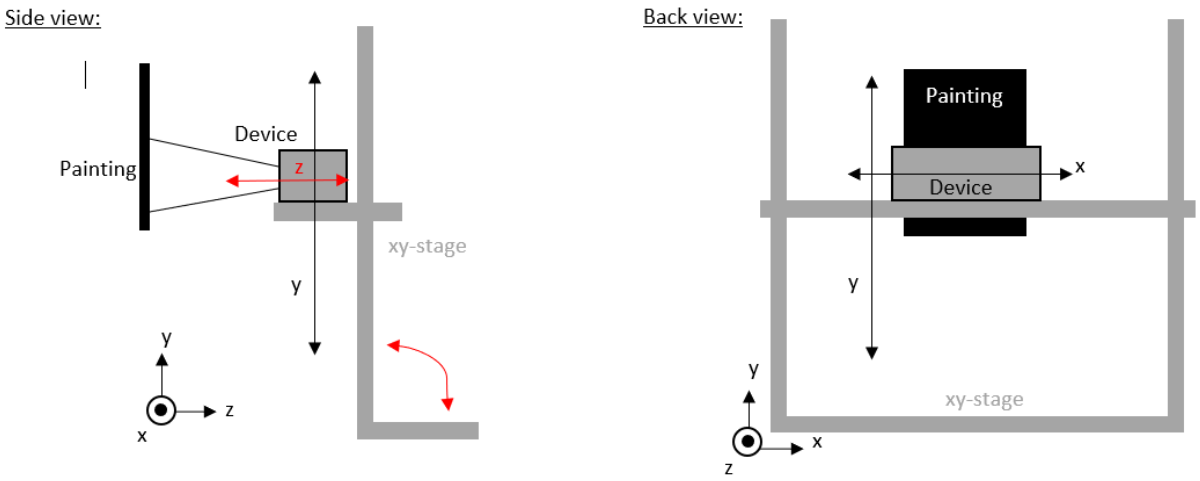



Figure 5.4: A side-view of the xy-stage, which can electronically translate in the x and y direction. The stage can manually rotate about the x-axis at the base and manually translate in the z direction.

the surface for multiple points. In theory, a straight, flat surface in 3D can be defined with a set of three points, but the surface of a painting is not flat nor straight. So to obtain an accurate estimation of the orientation, the distance should be determined with multiple points. With a spatial resolution of 10 mm A.3.2 an estimation of the orientation can be computed. All depth estimation techniques discussed in chapter 2 are able to reach such spatial resolution and thus is the spatial resolution excluded from the requirements.

### Distance estimation

The focus of the methods is the distance estimation, which should at least be able to determine whether each point is within the Depth of Field with enough certainty. Based on a total Depth Of Field of 0.739 mm, the distance estimation method should reach at least a depth precision of 0.1 mm [13].

### Acquisition Time

The acquisition time should be minimised, since it would only increase the total scanning time without contributing to the colour and depth images. Therefore, determining the orientation once at the beginning or even excluded from the scanning time is preferred. To give an idea, the acquisition time of one tile (60x40mm) should maximally take 3 seconds, which sums up to about 4 minutes for 1750 $cm^2$ (A.1).

### Costs

The need for orientation estimation is an unanticipated byproduct of the resolution enhancement and thus initially not calculated in the budget of the device development. However, since the costs of the resolution enhancement is lower than expected, the available budget is roughly estimated at 1000 euros. So the objective is to estimate the distance with a depth precision of at least 0.1 mm, while minimising the acquisition time and costs.

- Depth precision of 0.1 mm
- Acquisition time 4 minutes
- Costs € 1000

## 5.2. Methods

The methods can be classified into distance estimation methods and methods based on the auto-focus of cameras. The distance estimation methods are related to the techniques discussed in chapter 2. The auto-focus based methods do not measure a distance, but optimise the sharpness of the photograph. Normally in auto-focus, the focus is adopted until an optimal is achieved that is related to the sharpness of the photograph, but the tilt-shift lens prevents the camera from using the auto-focus. Therefore, the

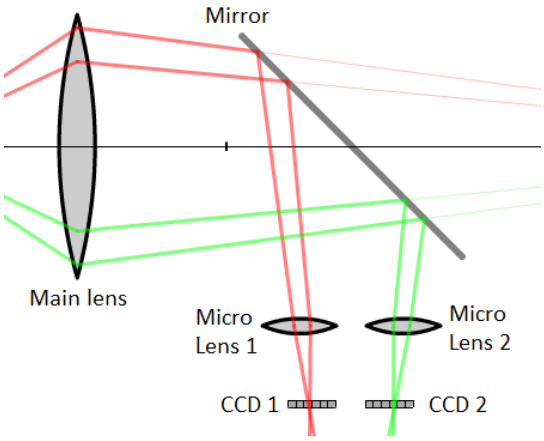



Figure 5.5: Auto-Focus from phase detection: The light enters the camera and is directed to the phase detection sensor. When the light enters the phase detection sensor, the light is redirected with a micro lens to a group of sensors (two sensors per AF point). The (CCD) sensors compare the images, when they are not identical the focus of the lens is changed [168].

focus is replaced with the control of the z-axis to optimise sharpness of the image with the distance of the camera. The methods used in auto-focus are discussed first, followed by the distance estimation methods.

### 5.2.1. Phase Detection

Phase detection is a common auto-focus technique used in expensive Digital Single Lens Reflection cameras (DSLR ), like the Canon EOS 5Ds. The light enters the camera and is directed to the phase detection sensor. When the light enters the phase detection sensor, the light is redirected to a group of sensors (two sensors per AF point) with a micro lens. The images from these sensors are compared and if they do not look identical, the focus of the lens is adjusted [168][169]. Figure 5.5 illustrates the working principle of phase detection.

As mentioned, adjusting the focus of lens is replaced by adjusting the distance to the painting with control of the z-axis. For phase detection auto-focus to work correctly, the distance between the lens mount and the camera sensor, as well as the distance between the lens mount and the phase detection sensor must be identical [170][171], therefore does phase detection for a camera with a tilt-shift lens not work. So phase detection is not viable method and will not be taken into account.

### 5.2.2. Contrast Detection

Most Digital Still Cameras (DSC ) and mobile cameras use contrast detection to get the photograph in-focus. Auto-focus based on contrast detection, similarly to phase detection, changes the focus of the lens until the optimum is found [172][173]. Like with phase detection the translation of the z-axis will replace changing the focus.

Contrast is the difference between opposite colours, often the difference between white and black. Figure 5.6 shows three examples with different contrast levels. The histograms show that a sharper image results in a larger contrast. In other words, a sharper image concentrates the tone distribution towards their true tone value and thus larger peaks [174].

In this method, auto-focus does not involve actual distance measurement, and thus the method is unable to determine the orientation. This creates significant challenges when tracking moving subjects, since a loss of contrast gives no indication of the direction of motion towards or away from the camera. Moreover, it optimises the contrast for the current positioning, but that does not take into account a more ideal orientation, and thus is the technique sensitive to local optima. Furthermore, during the optimisation of the contrast, the estimation generally overshoots before it finds the optimum, increasing the acquisition time [175].

Contrast detection does not measure the distance, so it is difficult to determine the depth precision. The accuracy of contrast detection depends on the measurable difference in contrast or the step size

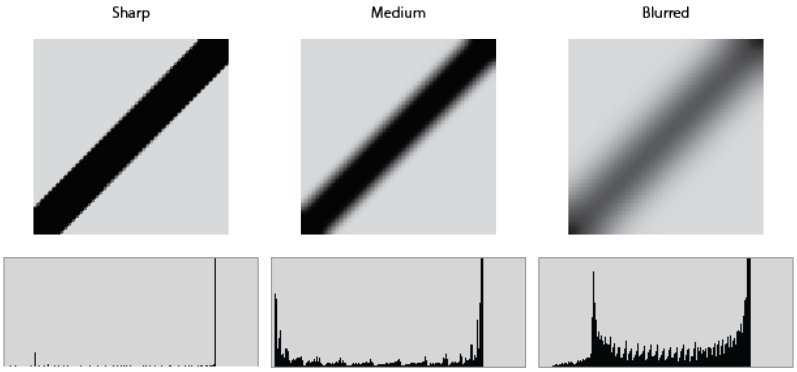



Figure 5.6: Auto-Focus from contrast: The sharpness of a image is related to the contrast of the image. A higher contrast means the difference in tone is higher, which optimal contrast level is achieved when the image is in-focus. The histograms clearly show that a for a sharper image the tone is concentrated more and thus a higher contrast.

of the z-axis. The z-axis translates 1 mm per rotation (screw: ISO 262), motorising the z-axis with a common stepper motor with 200 steps per rotation [176], would have a step size of 0.005 mm. So the step size has enough precision to reach the requirement.

The precision of the contrast is determined by the difference in tone. The Canon EOS 5Ds has a bit-rate of 14-bit, meaning it has 16384 ($2^14$) different tones from black to white for each pixel. It is difficult to relate that to the depth precision, but assuming the contrast will linearly differ from complete unfocused to focused in a spacious 100 mm then the contrast would change every 0.006 mm. So the approximated contrast precision is similar to the translational precision.

The acquisition time of contrast detection depends on the translation speed of the z-axis, the accessibility of camera software, and the amount of tiles. The quickest auto-focus is when the live contrast computations of the camera can directly control the z-axis, instead of waiting for the photograph each time the z-axis is translated, but direct access to the cameras is required. This method requires to focus the image for each tile. The amount of tiles depends on the size of the painting, which is 187 tiles for 1750 $cm^2$ with the ultra-high fringe encoded stereo imaging with a 30% overlap. It is estimated to take 2 seconds to determine the optimal contrast in the ideal situation, which would add about 6 minutes to the scan of the painting A.3.3. However, the contrast detection can be computed with the cameras of the device, requiring no additional costs.

### 5.2.3. Line Laser Triangulation

As discussed in chapter 2, laser triangulation determines the depth when the position and orientation of the light source relative to the observer is known. Multiple distance measurements create a 3D estimation of the painting's surface [31]. Moreover, laser triangulation has the possibility to scan with multiple points in a line, reducing the acquisition time .

In theory, the technique can reach a depth precision of 0.4 µm and an acquisition time of 15 minutes for the entire painting (A.3.4). The costs are related to the required depth precision, which is lower than in than chapter 2. Therefore, the Line laser triangulation is estimated at about 500 euros [40].

### 5.2.4. Laser Time Of Flight

Distance estimation based on Laser Time Of Flight is also discussed in chapter 2 [177]. It is a point measurement technique and thus require a longer acquisition time, and as discussed it is less accurate than line laser triangulation. This method is an economical option to give a quick estimation of the orientation. Moreover, this technique is currently used to determine the distance to the painting for the lower resolution scanner [84].

As mentioned in chapter 2, the laser Time Of Flight can reach a depth precision of 0.1 mm just meeting the requirement, which is not ideal since it is a theoretical estimate. The acquisition time is about 40 times longer than line laser triangulation and thus more than an hour. However, the costs of about €100 makes the method is more economical than laser triangulation and therefore is the technique a good alternative for a quick check.

### 5.2.5. Projector-Camera(s) Triangulation

The Projector-Camera Triangulation determines the distance and orientation with one photograph, when the position and orientation of the cameras relative to the projector are known. Since the spatial resolution is not a limiting requirement, the projected pattern could be structured light or just a checkerboard to determine the depth information [178][179]. The main drawback is that the projector needs to be calibrated as well, while this was eliminated in fringe encoded stereo imaging.

The depth precision can reach 0.05 mm similar to structured light, which reaches the requirement. Assuming the amount of tiles without overlap, it would take a estimated 3 minutes (A.3.5) at the beginning of the scan. The orientation of the device is adopted to the estimated surface of the painting. Moreover, projector-camera triangulation uses the hardware from fringe encoded stereo imaging and therefore no additional costs have to be made.

### 5.2.6. Sparse Stereo Imaging

As discussed in chapter 3, sparse stereo image matching has a high depth precision and computes the distance information quicker compared to other matching algorithms. The main drawback is that sparse matching has issues finding feature to match in non-salient areas, and therefore unable to match the entire image. However, to estimate the orientation of the painting, not the entire image has to be matched. Moreover, a pattern can be projected to amplify the features of the painting to to get a better match. Since this method is based on stereo imaging, the hardware of ultra-high resolution fringe encoded stereo imaging can be used, reducing costs and the implementation time. Furthermore projector calibration is not required for this technique.

As determined in chapter 2, stereo imaging can achieve 0.5 µm in theory, but in chapter 4 is estimated that 3 µm is more realistic for the ultra-high resolution fringe encoded stereo imaging. Similar to projector-camera triangulation only one photograph is required to estimate the orientation of the surface. Therefore, it is estimated to take 3 minutes to scan the entire painting. Moreover, no additional costs are required since all the hardware is already present.

## 5.3. Comparison

The methods were classified in distance estimation and auto-focus techniques. For the auto-focus methods to work, the control of focus is replaced with the control of the z-axis, to obtain the optimal distance to the painting for every tile. These methods do not determine the orientation, but the optimal image sharpness for the current orientation. The distance estimation methods determine the orientation of the painting with multiple distance measurement points.

| method | Depth precision [mm] | Acquisition Time [min] | Costs [€] | Notes |
|---|---|---|---|---|
| Contrast Detection | ±0.05 | 6 | 0 | - No distance estimation<br>- Sensitive to local optima |
| Laser Triangulation | 0.0004* | 15 | 500 | |
| Laser Time Of Flight | 0.1* | >60 | 100 | |
| Projector-Camera Triangulation | 0.05* | 3 | 0 | - Projector calibration |
| Sparse Stereo Imaging | 0.03 | 3 | 0 | |

Table 5.1: An auto-focus method based on contrast and several methods to estimate the distance of the device to the painting's surface, to construct the orientation of the painting relative to the device. * The depth precision is their theoretical potential and will therefore be higher in practise.

The table shows a clear winner among the methods: Sparse Stereo imaging. The method has second best depth precision and lowest estimated acquisition time, without any additional costs.

## 5.4. Validation

The depth precision is determined with the stereo imaging from ultra-high fringe encoded stereo imaging, so the same depth precision will be reached. Therefore, the validation of the depth precision is equal to the validation discussed in chapter 4.

## 5.5. Discussion

The curvature of paintings and/or high difference in paint height, can in certain cases exceed the Depth Of Field. In the case of larger paintings, like 'the Night Watch', the curvature of the canvas is larger than the Depth Of Field In that case the entire painting's surface will never fit within the Depth of Field of the device without changing the z-axis.

In the case of concentrated high difference in paint height, like with the sleeve in 'the Jewish Bride' several millimeters [180], the height difference is larger than the Depth Of Field. This could be solved by positioning the concentrated height difference in the middle of the image, since the Depth Of Field is larger in the middle. However, when the height difference even exceeds that, the device will be unable to give an accurate colour and topography reproduction of that area. This is the main drawback of increasing the resolution of this technique.

The acquisition time can be reduced by an initial global scan of the corners and centre, instead of scanning the entire painting directly. This kind of global scan would give a good initial estimate of the orientation of the painting, since the frame at the edges primarily forms the orientation, and in general, the greatest depth change, due to curvature, is expected in the centre. After the device is re-positioned to the global orientation, a more thorough scan can be conducted to verify the global scan more accurately.

## 5.6. Conclusion

An estimate of the orientation of the painting is required, to position the device such that the painting fits within the Depth Of Field. To determine the orientation of the painting several methods were accessed on the depth precision, acquisition time and costs. Sparse stereo imaging is the best method to determine the orientation, because of high depth precision of 0.03 mm and low acquisition time of about 3 minutes, relative to the other methods. Moreover, the method uses the hardware of the fringe encoded stereo imaging, and therefore no additional costs are required and the implementation should be swift.

The minimum Depth Of Field is estimated to be 0.739 mm, for the proposed ultra-high resolution fringe encoded stereo imaging. The entire painting is preferably placed within the Depth Of Field of the device, to obtain in-focus photographs to determine the topographical information. Positioning the device such that the painting is within the Depth Of Field, is achieved by motorising the manual rotation about the x-axis and the translation of the z-axis of the xy-stage. The parallel alignment of the x-axis to the painting, is preferred as precise as possible to minimise the need for z-axis corrections.

# 6

# Conclusion

The literature study begins with a holistic view of different areas of techniques that can be used in capturing colour and topographical information of the craquelure pattern of the painting. The holistic literature study concludes that fringe encoded stereo imaging holds the most potential for craquelure scanning. Apart from being able to reach an ultra-high spatial resolution and high depth precision, the main advantage is having state-of-the-art cameras and high-tech xy-stage directly available at no initial cost. However, the spatial resolution and depth precision of the current technique reach 25 µm and 10 µm resp., which is insufficient to capture the craquelure pattern of paintings.

Chapter 3 analyses fringe encoded stereo imaging in depth, to determine how to reach the required spatial resolution and depth precision to capture the craquelure pattern.

Analysis of the technique also shows that fringe encoded stereo imaging requires careful calibration of the white balance, the camera orientations and the colour to reproduce the painting accurately. Additionally, the colour accuracy is sensitive to change of environmental lighting conditions, and thus a dark room is preferred to obtain the highest colour accuracy.

Furthermore, the projector should be replaced with a high intensity (>1000 Lumens) and high resolution (HD) projector, while fitting between the cameras. This will reduce the acquisition time and increase the Signal-to-Noise Ratio.

The in-depth analysis finally concludes that the camera and its lens determine the spatial resolution and depth precision of fringe encoded stereo imaging. The software and other hardware can handle the increase in resolution. There are several methods to enhance the resolution of the camera and lens and therefore another chapter will analyse the possibilities to adapt the camera and especially the lens to increase the spatial resolution from 25 µm to lower than 10 µm.

Further literature research indicated, that fringe encoded stereo imaging has multiple camera resolution enhancement methods, from which extension tubes are the most suited solution, from both theoretical and practical points of view. However, increasing the spatial resolution reduces the Depth Of Field of the device. In other words, the entire painting has to be orientated in a thinner parallel area of focus. Initial estimations place the Depth Of Field smaller than a millimetre, which makes it difficult orientate the painting within the Depth Of Field by hand.

Final chapter studies literature on estimating the orientation of the painting, to be able to position the device such that the painting fits within the Depth Of Field. The study indicated several methods for determining the orientation of the painting, from which sparse stereo imaging is the best method for the precision and acquisition time.

This literature study concludes that it is feasible to accurately scan the craquelure pattern of the entire painting with enhanced fringe encoded stereo imaging and accurate positioning.

# A

# Appendix: Computations

In this chapter underlying computations and estimations for the entire literature study were done. With these equations and computation are the estimation reproducible.

## A.1. Chapter 2

The estimation of the time acquisition are computed with the code below, where the a full-frame camera or common microscope can be selected by uncommenting. The tiles are comuted for the device and for an overlap of 30 %. The magnification is 0.3 for the current set-up.

```
%% Master Thesis: x-y microscope scanner for paintings
% Mathijs van Hengstum 04/2017

clear all;  close all; clc

%% Parameters
% painting
painting.x = 0.39;              % [m] 445x390mm (GwtPE) width 4.38
painting.y = 0.445;              % [m] height 3.63
% painting.x = 0.445;              % [m] 445x390mm (GwtPE) width 4.38
% painting.y = 0.39;              % [m] height 3.63
painting.pigment = 5*10^-6; % [0.1 - 5μm] http://www.engineeringtoolbox.com/particle-sizes-d_934.html
pixels_per_pigment = 5;      % pixels^2
n_images = 13;

% Microscope (RH-2000) ( Literature//HIROX-RH-2000-EN.pdf )
% MXB 2016Z  20x - 160x 15.4-2 mm (H)
% MXB 5040RZ 50x - 400x 6.1-0.78 mm (H)
% MXB-10C   140x - 1400x 2.46-0.26 mm (multiple) (H)
% Full frame 50.3Mpx 8688x5792 px (35.88x23.92 mm)

% MICROSCOPE
% microscope.zoom    = [0.1  1];     % min max zoom of lens
% steps = 2;                         % steps in microscope zoom
% microscope.pixel.x  = 1920;       % [pixels] width
% microscope.pixel.y  = 1080;       % [pixels] height
% microscope.HFOV     = 0.002;      % [m] full frame horizontal
% microscope.VFOV     = 0.001125;     % [m] full frame vertical
% microscope.confocal = 1;          % number of depth images
% % microscope.weight   = 1;        % [kg] weight of the microscope
% microscope.FPS      = 1;          % [frame/s] Frame Per Seconds

% FULL_FRAME CAMERA
microscope.zoom    = [0.6  1];     % min max zoom of lens
microscope.pixel.x  = 8688;       % [pixels] width
microscope.pixel.y  = 5792;       % [pixels] height
microscope.HFOV     = 0.036;      % [m] full frame horizontal
microscope.VFOV     = 0.024;      % [m] full frame vertical
microscope.confocal = 1;          % number of depth images
```

```
% microscope.weight   = 1;       % [kg] weight of the microscope
microscope.FPS       = 1;       % [frame/s] Frame Per Seconds

% x-y control/1
overlap              = 0; % [%] overlap of pictures
movement_time        = 240;    % [seconds] average time to move to new position

%% Computations
% microscope.HFOV     = linspace(microscope.FOV(1),microscope.FOV(2),steps);
% microscope.HFOV     = 30.8./microscope.HFOV/1000;   % [m] Horizontal Field of View
% microscope.VFOV     = microscope.HFOV*(microscope.pixel.x/microscope.pixel.y);
% microscope.zoom     = linspace(microscope.zoom(1),microscope.zoom(2),steps);
microscope.HFOV = microscope.HFOV./microscope.zoom;
microscope.VFOV = microscope.VFOV./microscope.zoom;

painting.x = painting.x+2*overlap*microscope.HFOV ;
painting.y = painting.y+2*overlap*microscope.VFOV ;

% Number of pictures
n_x = ceil(painting.x./((1-overlap).*microscope.HFOV));
n_y = ceil(painting.y./((1-overlap).*microscope.VFOV));
n = n_x.*n_y*microscope.confocal;

% Time spend on scanning
Time = (n*movement_time)/3600; % [time] %/microscope.confocal+ n/microscope.FPS)/3600; % [time]

% Length per pixel & required for pigment
pixel_size = microscope.HFOV/microscope.pixel.x;
required = painting.pigment/pixels_per_pigment;


%% Plots
figure(1);subplot(3,1,1)
plot(microscope.zoom,1.2*n*microscope.pixel.x*microscope.pixel.y*n_images/10^9)
ylabel('Total data size [Gb]')
title('Estimated required storage space [Gb]')

figure(1);subplot(3,1,2)
plot(microscope.zoom,Time)
ylabel('Time [Hours]')
title('Hours to Scan')

figure(1); subplot(3,1,3)
plot(microscope.zoom,pixel_size*10^6); hold on
plot(microscope.zoom,required*ones(length(microscope.zoom))*10^6,'r');
legend('Microscope','Required for pigment size')
xlabel('zoom [X]')
ylabel('Length per pixel [µm/pixel]')
title('Length per Pixel')
```

### A.1.1. Line laser estimation

Assuming the line lase triangulation takes 40 measurements per second with about 250 point each measure [36]. The with the spatial resolution of 10 µm there are 44.500x39.000 measure points.

$$(44.500 * 39.000)/(40 * 250)/3600 = 48.21 hours \tag{A.1}$$

The size an photograph of a digital camera, at the required spatial resolution, would ideally take about 50 hours.

### A.1.2. Stereo Imaging

Using a full-frame Canon EOS 5Ds with a pixel pitch of 4.1 µm and the highest magnification lens (5:1) would give a spatial resolution of about 1 µm and depth precision of lower than 0.5 µm. The theory is explained in-depth in chapter 5.

### A.1.3. Depth from Focus

For the estiamation of the Depth of Field the equation below is used, where Aperture Number (N), Circle of confusion (C),the lens-object distance (U), and focal length (f) determine the depth of Field. It is explained more in depth in section 5.

$$DOF_{total} \approx \frac{2NCU^2}{f^2} \tag{A.2}$$

With a small aperture number and long focal length the smallest depth of field is achieved. For a tilt-shift lens it would mean a focal lenght of 135 mm, Circle of confusion of 1 pixel 0.00413 mm, a small aperture of 1.4 and a object lens distance of 270 mm gives 46,3 µm Depth Of Field.

### A.1.4. Structured Light

Assuming an 4K ultra-high resolution projector (4096 x 2160 px) could focus at a short distance and the projected structured light fits with in the current window of 165x110 mm. The structured light has to be covered for the entire window, so the resolution is determined by the height thus the maximum resolution is:

$$110/2160 = 0.05092mm \approx 50m \tag{A.3}$$

## A.2. Chapter 4

### A.2.1. Camera model Comparison

```
%% Pixel pitch vs resolution
% (c) Mathijs van Hengstum 2018
% compares current data of commercial cameras


data = readtable('cameras.xls');
data2 = struct;

for i =1:526
    data2(i).res = cell2mat(data{i,3});
    data2(i).res = str2double(data2(i).res(1:4));
    data2(i).pitch = cell2mat(data{i,4});
    data2(i).pitch = str2double(data2(i).pitch(1:4));
end

figure(1)
for i = 1:526
plot(data2(i).res,data2(i).pitch,'.k');
hold on
end
xlabel('Resolution [Mega pixels]')
ylabel('Pitch [µm]')

surface = 440*390; % [mm]



figure(2)
for i =1:526
    ratio(i) = 1/(surface/(0.7*data2(i).res)*10^2);%/data2(i).pitch;
    plot(i,ratio(i),'.k')
    hold on
end
xlabel('Camera model')
ylabel('Number of Tiles^{-1}')
title('The Inverse inverse of Acquistion tiles for 526 Commercial cameras ')
% ylim([0 0.5*10^7])
text(85+5,ratio(85),'Hasselblad H6D-100c')
text(35+5,ratio(35)*1.05,'Canon EOS 5D Mark IV')
text(130+5,ratio(130)*0.99,'Canon EOS 5DS (R)')
```

### A.2.2. macro lens

Combining the Canon EOS 5Ds with a pixel pitch of 4.13 µm and the 5:1 magnification macro lens gives a spatial resolution of exactly 4.13/5 = 0.826 µm.

### A.2.3. close-up filter

$$Magnification = 4.13m/10m = 0.41 \tag{A.4}$$
$$MP_0 = 0.41/0.3 = 1.377 \tag{A.5}$$
$$MP_0 = (0.25m)\Phi + 1 \tag{A.6}$$
$$m = (MP_0 - 1)/0.25 = 0.377/0.25 = 1.5dpt \approx 2dpt \tag{A.7}$$

## A.3. Chapter 5

### A.3.1. DOF

The Depth Of Field is determined with the equations below.

```
%% Camera Computations Written by Mathijs van Hengstum, Novermber 2017
% The DoF is determined for Canon 5DS with different lenses
% A TS-E 90mm tilt(-shift) lens is enhanced with a extension tube

% Parameters
f       = 90;           % [mm] Focal lenght
Alpha   = 8;            % [degrees] tilt angle
A       = 110+25;       % [mm] Lens rear node to Image plane
N       = 16;           % [] Aperture Number
tile_x  = 60.8;         % [mm] horizontal size tile
% U       = 270;          % [mm] lens-object distance

% Circle of Confusion (as small as possiple, about 1px thus 4,13µm)
CoC = 0.00413;          % [mm] (diameter) Circle of Confusion

% Scheimpflug Principle (tilt-shift lens)
J       = f/sind(Alpha);
Gamma   = atand(sind(Alpha)/(cosd(Alpha)-f/A));

% Near and Far Acceptable Image Plane with CoC
s       = f/CoC/N;
A_near = A+A/(s-1);
A_far  = A-A/(s+1);


% Determine NDOF and FDOF
o_near = A_near/tand(Alpha);
o_far = A_far/tand(Alpha);

% Angles N and F
gamma_near = atand(A_near/(o_near-J));
gamma_far  = atand(A_far/(o_far-J));

% Angle difference
d_angle_near = Gamma - gamma_near;
d_angle_far  = gamma_far - Gamma;

% Sharp focus plane length
SF = J/cosd(Gamma);

% near and far DOF
NDOF = (SF-tile_x/2)*tand(d_angle_near);
FDOF = (SF-tile_x/2)*tand(d_angle_far);
```

### A.3.2. Requirements

The amount of points to describe the painting's surface depends on the change in shape. The curvature of a painting can be approximated with a parabola from the frame. Assuming a parabola is described

with 9 samples the spatial resolution is 390/9 = 43 mm. Assuming more complex surfaces can be described by a 4 times a simple parabola, the spatial resolution is about 10 µm.

### A.3.3. Contrast detection

Based on the Driver.m file 187 tiles are required for 'the girl with a pearl earring. Ideally it would take 2 seconds per tile, so that would result in:

$$187 * 2/60 = 6.233 minutes \tag{A.8}$$

### A.3.4. Line laser estimation

Assuming the line lase triangulation takes 40 measurements per second with a effective spatial resolution of 5 points each measure [36], which depends on the spatial resolution and working distance.

$$(445 * 390)/(40 * 5)/60 = 14.46 minutes \tag{A.9}$$

The size an photograph of a digital camera, at the required spatial resolution, would ideally take about 15 minutes.

### A.3.5. Projector-Camera Triangulation

Without overlap 84 tiles are required to capture the girl with a pearl earring, with 2 seconds for each tile. That results in 84*2/60 = 2.8 minutes.

# Bibliography


[1] J. Dik, K. Janssens, G. Van Der Snickt, L. van der Loeff, K. Rickers, and M. Cotte, *Visualization of a lost painting by vincent van gogh using synchrotron radiation based x-ray fluorescence elemental mapping,* Analytical chemistry **80**, 6436 (2008).

[2] ARTMYN Inc., *Artmyn (re)discover art,* (2017).

[3] T. Zaman, *Development of a topographical imaging device for the near-planar surfaces of paintings,* TU Delft Repository (2013).

[4] H. Europe, *Digital Microscope Hirox RH-2000,* (2000).

[5] S. L. Bucklow, *A stylometric analysis of craquelure,* Computers and the Humanities **31**, 503 (1997).

[6] D. Thompson, *On Growth and Form.* (Cambridge University Press, Cambridge, UK, 1917).

[7] G. Hedley, *Relative humidity and the stress/strain response of canvas paintings: uniaxial measurements of naturally aged samples,* Studies in conservation **33**, 133 (1988).

[8] W. Hogarth, *Science for Arts of the Netherlands,* (2014).

[9] Mauritshuis, *Details: Johannes vermeer, girl with a pearl earring, c. 1665,* (2017), accessed: 2017-12-05.

[10] S. Bucklow, *The description of craquelure patterns,* Studies in Conservation **42**, 129 (1997).

[11] Mauritshuis, *Sample during treatment: girl with the pearl earring,* (1994), unpublished research.

[12] J. H. Stoner and R. Rushfield, *Conservation of easel paintings* (Routledge, 2013).

[13] R. N. Bracewell and R. N. Bracewell, *The Fourier transform and its applications*, Vol. 31999 (McGraw-Hill New York, 1986).

[14] W. J. Dixon and F. J. Massey Jr, *Introduction to statistical analysis* (McGraw-Hill, 1957).

[15] J. Wettlaufer, *Physical principles of medical imaging.* Radiology **200**, 504 (1996).

[16] D. Huang, E. A. Swanson, C. P. Lin, J. S. Schuman, W. G. Stinson, W. Chang, M. R. Hee, T. Flotte, K. Gregory, C. A. Puliafito, *et al.*, *Optical coherence tomography,* Science (New York, NY) **254**, 1178 (1991).

[17] V. K. Kinra and C. Zhu, *Ultrasonic nondestructive evaluation of thin (sub-wavelength) coatings,* The Journal of the Acoustical Society of America **93**, 2454 (1993).

[18] T. Pialucha and P. Cawley, *The detection of thin embedded layers using normal incidence ultrasound,* Ultrasonics **32**, 431 (1994).

[19] Block Engineering, *Michelson interferometer operation,* (2012), accessed: 2017-12-05.

[20] M. Halliwell and P. N. Wells, *Acoustical Imaging*, Vol. 25 (Springer Science & Business Media, 2006).

[21] E. Konkov, *Ultrasonic interferometer for high-accuracy linear measurements,* Measurement Science Review **9**, 187 (2009).

[22] Physik Instrumente, *Vt-80 linear stage datasheet,* (2017).

[23] I. Arias and J. D. Achenbach, *A model for the ultrasonic detection of surface-breaking cracks by the scanning laser source technique,* Wave Motion **39**, 61 (2004).

[24] *How much does an ultrasound machine cost?* (2017), accessed: 2017-11-28.

[25] T. Callewaert, *OCT for the study of varnish layer stratigraphy,* - **31**, 2782034 (2015).

[26] Y. Chen, A. D. Aguirre, P.-L. Hsiung, S.-W. Huang, H. Mashimo, J. M. Schmitt, and J. G. Fujimoto, *Effects of axial resolution improvement on optical coherence tomography (oct) imaging of gastrointestinal tissues,* Optics express **16**, 2469 (2008).

[27] H. Verstraete, *Optimization-based adaptive optics for optical coherence tomography,* (2017).

[28] P. R. Berman, *Atom interferometry* (Academic press, 1997).

[29] T. Calleweart, *Oct scanning the girl with a pearl earring: 222.000 measurements in 10 hours,* (2018), unpublished research.

[30] Ernie Bowling, *Affording oct in your practise,* (2015).

[31] Z. Ji and M.-C. Leu, *Design of optical triangulation devices,* Optics & Laser Technology **21**, 339 (1989).

[32] D. Acosta, O. García, and J. Aponte, *Laser Triangulation for shape acquisition in a 3D Scanner Plus Scanner,* Proceedings - Electronics, Robotics and Automotive Mechanics Conference, CERMA 2006 **2**, 14 (2006).

[33] N. Van Gestel, S. Cuypers, P. Bleys, and J. P. Kruth, *A performance evaluation test for laser line scanners on CMMs,* Optics and Lasers in Engineering **47**, 336 (2009).

[34] V. Lombardo, T. Marzulli, C. Pappalettere, and P. Sforza, *A time-of-scan laser triangulation technique for distance measurements,* Optics and Lasers in Engineering **39**, 247 (2003).

[35] N. Broten, T. Legg, J. Locke, C. McLeish, R. Richards, R. Chisholm, H. Gush, J. Yen, and J. Galt, *Long base line interferometry: a new technique,* Science **156**, 1592 (1967).

[36] S. Paulus, T. Eichert, H. E. Goldbach, and H. Kuhlmann, *Limits of active laser triangulation as an instrument for high precision plant imaging,* Sensors **14**, 2489 (2014).

[37] Anon, *In-line profile measurement,* Wire **49** (1999).

[38] D. Acosta, O. García, and J. Aponte, *Laser triangulation for shape acquisition in a 3d scanner plus scan,* in *Electronics, Robotics and Automotive Mechanics Conference, 2006*, Vol. 2 (IEEE, 2006) pp. 14–19.

[39] Mirco-Epsilon, *Laser scanners for 2d/3d profile measurement,* (2017).

[40] M. Epsilon, *Measurement Product Guide,* (2010).

[41] B. Zitova and J. Flusser, *Image registration methods: a survey,* Image and vision computing **21**, 977 (2003).

[42] H.-C. Huang and Y.-P. Hung, *Panoramic stereo imaging system with automatic disparity warping and seaming,* Graphical Models and Image Processing **60**, 196 (1998).

[43] D. G. Lowe, *Distinctive image features from scale-invariant keypoints,* International journal of computer vision **60**, 91 (2004).

[44] M. DEUBER, S. CAVEGN, and S. NEBIKER, *Dense image matching,* (2014).

[45] R. Jain, R. Kasturi, and B. G. Schunck, *Machine vision*, Vol. 5 (McGraw-Hill New York, 1995).

[46] S. K. Nayar and Y. Nakagawa, *Shape from focus,* IEEE Transactions on Pattern analysis and machine intelligence **16**, 824 (1994).

[47] K. S. Pradeep and A. N. Rajagopalan, *Improving shape from focus using defocus cue,* IEEE Transactions on Image Processing **16**, 1920 (2007).

[48] M. Subbarao and G. Surya, *Depth from defocus: a spatial domain approach,* International Journal of Computer Vision **13**, 271 (1994).

[49] C. Zhou, S. Lin, and S. K. Nayar, *Coded aperture pairs for depth from defocus and defocus deblurring,* International journal of computer vision **93**, 53 (2011).

[50] G. Sirat and D. Psaltis, *Conoscopic holography,* Optics letters **10**, 4 (1985).

[51] James K., *Bifringent crystal,* (2017).

[52] Optimet, *Basics of conoscopic holography,* (2017), accessed: 2017-12-07.

[53] B. B. Wang, *Linear birefringence measurement instrument using two photoelastic modulators,* Optical Engineering **41** (2002).

[54] A. Wilson, *Conoscopic holography measures submicron distances,* (1998).

[55] B. L. Olmstead and A. Shearin, *Extended depth of field imaging system using chromatic aberration,* (2007), uS Patent 7,224,540.

[56] E. Leonhardt, *Product presentation: Nano point scanner,* (2017), unofficial presentation of future products.

[57] S. C. Tucker, W. T. Cathey, and E. R. Dowski, *Extended depth of field and aberration control for inexpensive digital microscope systems,* Optics Express **4**, 467 (1999).

[58] H. Rubinsztein-Dunlop, A. Forbes, M. Berry, M. Dennis, D. L. Andrews, M. Mansuripur, C. Denz, C. Alpmann, P. Banzer, T. Bauer, *et al.*, *Roadmap on structured light,* Journal of Optics **19**, 013001 (2016).

[59] D. Scharstein and R. Szeliski, *High-accuracy stereo depth maps using structured light,* in *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on*, Vol. 1 (IEEE, 2003) pp. I–I.

[60] R. J. Valkenburg and A. M. McIvor, *Accurate 3d measurement using a structured light system,* Image and Vision Computing **16**, 99 (1998).

[61] J. Salvi, S. Fernandez, T. Pribanic, and X. Llado, *A state of the art in structured light patterns for surface profilometry,* Pattern Recognition **43**, 2666 (2010).

[62] 3Dnatives, *3d sanning through structured light projection,* (2017).

[63] Stephen Jukic, *Best 4k projector,* (2017).

[64] C. Wust and D. W. Capson, *Surface profile measurement using color fringe projection,* Machine Vision and Applications **4**, 193 (1991).

[65] P. S. Huang, Q. Hu, F. Jin, and F.-P. Chiang, *Color-encoded digital fringe projection technique for high-speed 3-d surface contouring,* Optical Engineering **38**, 1065 (1999).

[66] S. S. Gorthi and P. Rastogi, *Fringe projection techniques: whither we are?* Optics and lasers in engineering **48**, 133 (2010).

[67] C. Quan, W. Chen, and C. Tay, *Phase-retrieval techniques in fringe-projection profilometry,* Optics and Lasers in Engineering **48**, 235 (2010).

[68] P. Bing, X. Hui-Min, X. Bo-Qin, and D. Fu-Long, *Performance of sub-pixel registration algorithms in digital image correlation,* Measurement Science and Technology **17**, 1615 (2006).

[69] P. Kühmstedt, C. Munckelt, M. Heinze, C. Bräuer-Burchardt, and G. Notni, *3d shape measurement with phase correlation based fringe projection,* in *Optical Measurement Systems for Industrial Inspection V*, Vol. 6616 (International Society for Optics and Photonics, 2007) p. 66160B.

[70] M. K. Johnson, F. Cole, A. Raj, and E. H. Adelson, *Microgeometry capture using an elastomeric sensor,* in *ACM Transactions on Graphics (TOG)*, Vol. 30 (ACM, 2011) p. 46.

[71] K. Ikeuchi, *Determining surface orientations of specular surfaces by using the photometric stereo method,* IEEE Transactions on Pattern Analysis and Machine Intelligence , 661 (1981).

[72] R. Basri, D. Jacobs, and I. Kemelmacher, *Photometric stereo with general, unknown lighting,* International Journal of Computer Vision **72**, 239 (2007).

[73] R. J. Woodham, *Photometric method for determining surface orientation from multiple images,* Optical engineering **19**, 191139 (1980).

[74] P. Tan, S. Lin, and L. Quan, *Resolution-enhanced photometric stereo,* Computer Vision–ECCV 2006 , 58 (2006).

[75] B. R. Masters, *Handbook of biological confocal microscopy,* Journal of biomedical optics **13**, 029902 (2008).

[76] D. E. Caldwell, D. R. Korber, and J. R. Lawrence, *Confocal laser microscopy and digital image analysis in microbial ecology,* in *Advances in microbial ecology* (Springer, 1992) pp. 1–67.

[77] OpenStax, *College Physics* (OpenStax OpenStax College Physics, 2012) Chap. 27.

[78] R. H. Webb, *Confocal optical microscopy,* Reports on Progress in Physics **59**, 427 (1999), arXiv:arXiv:1011.1669v3 .

[79] E. L. H. E. specialist), *Hirox rh-2000: estimation price for rijksmuseum,* (2017), unpublished research.

[80] M. Mudg, C. Schroer, G. Earl, K. Martinez, H. Pagi, C. Toler-Franklin, S. Rusinkiewicz, G. Palma, M. Wachowiak, M. Ashey, *et al.*, *Principles and practices of robust, photography-based digital imaging techniques for museums,* (2010).

[81] A. A. Management, *Aerial photogrammetry,* (2016).

[82] PhotoModeler, *Factors affecting accuracy in photogrammetry,* (2017), accessed: 2017-12-08.

[83] W. S. Elkhuizen, B. A. Lenseigne, T. Baar, W. Verhofstad, E. Tempelman, J. M. Geraedts, and J. Dik, *Reproducing oil paint gloss in print for the purpose of creating reproductions of old masters,* in *Measuring, Modeling, and Reproducing Material Appearance 2015*, Vol. 9398 (International Society for Optics and Photonics, 2015) p. 93980W.

[84] W. S. Elkhuizen, T. Zaman, W. Verhofstad, P. P. Jonker, J. Dik, and J. M. Geraedts, *Topographical scanning and reproduction of near-planar surfaces of paintings,* in *Measuring, Modeling, and Reproducing Material Appearance*, Vol. 9018 (International Society for Optics and Photonics, 2014) p. 901809.

[85] T. Zaman, P. Jonker, B. Lenseigne, and J. Dik, *Simultaneous capture of the color and topography of paintings using fringe encoded stereo vision,* Heritage Science **2**, 23 (2014).

[86] D. Clark, *Notes on the resolution and other details of the human eye,* web (2005).

[87] E. O. Inc., *Edmund optics,* web (2018).

[88] E. Hsu, T. Mertens, S. Paris, S. Avidan, and F. Durand, *ACM Transactions on Graphics (TOG)*, Vol. 27 (ACM, 2008) p. 70.

[89] K. Hirakawa and T. W. Parks, *Chromatic adaptation and white-balance problem,* in *Image Processing, 2005. ICIP 2005. IEEE International Conference on*, Vol. 3 (IEEE, 2005) pp. III–984.

[90] S. McHugh, *Cambridge in colour,* web (2005-2018).

[91] Spencer Cox, *What is lens diffraction,* (2018).

[92] S. T. Mchugh, *Understanding Photography - Core Concepts and techniques with digital image capture* (Cambridge in Colour, 2016).

[93] E. Nocerino, F. Menna, F. Remondino, J.-A. Beraldin, L. Cournoyer, and G. Reain, *Experiments on calibrating tilt-shift lenses for close-range photogrammetry.* International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences **41** (2016).

[94] O. Albers, A. Poesch, and E. Reithmeier, *Flexible calibration and measurement strategy for a multi-sensor fringe projection unit,* Optics express **23**, 29592 (2015).

[95] F. Remondino and C. Fraser, *Digital camera calibration methods: considerations and comparisons,* International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences **36**, 266 (2006).

[96] Z. Zhang, *A flexible new technique for camera calibration,* IEEE Transactions on pattern analysis and machine intelligence **22**, 1330 (2000).

[97] C. S. Fraser, *Automatic camera calibration in close range photogrammetry,* Photogrammetric Engineering & Remote Sensing **79**, 381 (2013).

[98] R. Chen, J. Xu, H. Chen, J. Su, Z. Zhang, and K. Chen, *Accurate calibration method for camera and projector in fringe patterns measurement system,* Applied Optics **55**, 4293 (2016).

[99] W. Zhang, W. Li, L. Yu, H. Luo, H. Zhao, and H. Xia, *Sub-pixel projector calibration method for fringe projection profilometry,* Optics Express **25**, 19158 (2017).

[100] D. Moreno and G. Taubin, *Simple, accurate, and robust projector-camera calibration,* in *3D Imaging, Modeling, Processing, Visualization and Transmission (3DIMPVT), 2012 Second International Conference on* (IEEE, 2012) pp. 464–471.

[101] M. Files and K. Griggs, *Technical Guidelines for Digitizing Cultural Heritage Materials : Creation of Raster Image,* Image Rochester NY , 101 (2010).

[102] K. D. Gennetten, *Rgb to cmyk conversion using 3d barycentric interpolation,* in *Device-Independent Color Imaging and Imaging Systems Integration*, Vol. 1909 (International Society for Optics and Photonics, 1993) pp. 116–127.

[103] Y. Chen, R. S. Berns, L. A. Taplin, and F. H. Imai, *Color and Imaging Conference*, Vol. 2003 (Society for Imaging Science and Technology, 2003) pp. 277–281.

[104] G. Marcu, *Color and Imaging Conference*, Vol. 1992 (Society for Imaging Science and Technology, 1998) pp. 159–162.

[105] M. Anderson, R. Motta, S. Chandrasekar, and M. Stokes, *Color and imaging conference*, Vol. 1996 (Society for Imaging Science and Technology, 1996) pp. 238–245.

[106] G. Wyszecki and W. S. Stiles, *Color science*, Vol. 8 (Wiley New York, 1982).

[107] L. Rui-juan, *Study on color space conversion model from rgb to ciexyz [j],* Packaging Engineering **3**, 029 (2009).

[108] B. Goldluecke, *Variational analysis,* in *Computer Vision: A Reference Guide*, edited by K. Ikeuchi (Springer US, Boston, MA, 2014) pp. 839–840.

[109] A. Ford and A. Roberts, *Colour space conversions,* Westminster University, London **1998**, 1 (1998).

[110] I. C. Consortium *et al.*, *Image technology colour management-architecture, profile format, and data structure,* Specification ICC. 1: 2004-10 (Profile version 4.2. 0.0) (2004).

[111] R. S. Berns *et al.*, *Billmeyer and Saltzman's principles of color technology* (Wiley New York, 2000).

[112] C. A. Arikan, A. Brunton, T. M. Tanksale, and P. Urban, *Color-managed 3d printing with highly translucent printing materials,* in *Measuring, Modeling, and Reproducing Material Appearance 2015*, Vol. 9398 (International Society for Optics and Photonics, 2015) p. 93980S.

[113] C. Parraman, P. Walters, B. Reid, and D. Huson, *Specifying colour and maintaining colour accuracy for 3d printing,* in *Three-Dimensional Image Capture and Applications 2008*, Vol. 6805 (International Society for Optics and Photonics, 2008) p. 68050L.

[114] G. Sharma and R. Bala, *Digital color imaging handbook* (CRC press, 2002).

[115] W.-C. Cheng, T. Keay, N. O'Flaherty, J. Wang, A. Ivansky, M. A. Gavrielides, B. D. Gallas, and A. Badano, *Assessing color reproducibility of whole-slide imaging scanners,* in *Medical Imaging 2013: Digital Pathology*, Vol. 8676 (International Society for Optics and Photonics, 2013) p. 86760O.

[116] D. Scharstein and R. Szeliski, *A taxonomy and evaluation of dense two-frame stereo correspondence algorithms,* International journal of computer vision **47**, 7 (2002).

[117] T. Kanade and M. Okutomi, *A stereo matching algorithm with an adaptive window: Theory and experiment,* IEEE transactions on pattern analysis and machine intelligence **16**, 920 (1994).

[118] O. Miksik and K. Mikolajczyk, *Evaluation of local detectors and descriptors for fast feature matching,* in *Pattern Recognition (ICPR), 2012 21st International Conference on* (IEEE, 2012) pp. 2681–2684.

[119] J. L.Alvarez, R.Deriche and J.Weickert, *Dense disparity map estimation respecting image discontinuities,* (2000).

[120] Dinesh Nair, *A guide to stereovision and 3d imaging,* (2012).

[121] J. Dijk, P. Verbeek, S. Promotiecommissie, N. O. Voor, *et al.*, *c,* (2004).

[122] T. Hoang, B. Pan, D. Nguyen, and Z. Wang, *Generic gamma correction for accuracy enhancement in fringe-projection profilometry,* Optics letters **35**, 1992 (2010).

[123] H. Guo, H. He, and M. Chen, *Gamma correction for digital fringe projection profilometry,* Applied optics **43**, 2906 (2004).

[124] Y. Surrel, *Design of algorithms for phase measurements by the use of phase stepping,* Applied optics **35**, 51 (1996).

[125] P. S. Huang, Q. J. Hu, and F.-P. Chiang, *Double three-step phase-shifting algorithm,* Applied optics **41**, 4503 (2002).

[126] C. Quan, W. Chen, and C. J. Tay, *Phase-retrieval techniques in fringe-projection profilometry,* Optics and Lasers in Engineering **48**, 235 (2010).

[127] M. Takeda and K. Mutoh, *Fourier transform profilometry for the automatic measurement of 3-d object shapes,* Applied optics **22**, 3977 (1983).

[128] X. Su and W. Chen, *Fourier transform profilometry:: a review,* Optics and lasers in Engineering **35**, 263 (2001).

[129] S. Yunhui and R. Qiuqi, *Continuous wavelet transforms,* in *Signal Processing, 2004. Proceedings. ICSP'04. 2004 7th International Conference on*, Vol. 1 (IEEE, 2004) pp. 207–210.

[130] V. Srinivasan, H.-C. Liu, and M. Halioua, *Automated phase-measuring profilometry of 3-d diffuse objects,* Applied optics **23**, 3105 (1984).

[131] J. Tian, X. Peng, and X. Zhao, *A generalized temporal phase unwrapping algorithm for three-dimensional profilometry,* Optics and Lasers in Engineering **46**, 336 (2008).

[132] X. Peng, Z. Yang, and H. Niu, *Multi-resolution reconstruction of 3-d image with modified temporal unwrapping algorithm,* Optics communications **224**, 35 (2003).

[133] C. Polhemus, *Two-wavelength interferometry,* Applied Optics **12**, 2071 (1973).

[134] R. Dändliker, R. Thalmann, and D. Prongué, *Two-wavelength laser interferometry using superheterodyne detection,* Optics letters **13**, 339 (1988).

[135] M. Takeda, Q. Gu, M. Kinoshita, H. Takai, and Y. Takahashi, *Frequency-multiplex fourier-transform profilometry: a single-shot three-dimensional shape measurement of objects with large height discontinuities and/or surface isolations,* Applied optics **36**, 5347 (1997).

[136] C. E. Towers, D. P. Towers, and J. D. Jones, *Time efficient chinese remainder theorem algorithm for full-field fringe phase analysis in multi-wavelength interferometry,* Optics Express **12**, 1136 (2004).

[137] C. Zuo, L. Huang, M. Zhang, Q. Chen, and A. Asundi, *Temporal phase unwrapping algorithms for fringe projection profilometry: A comparative review,* Optics and Lasers in Engineering **85**, 84 (2016).

[138] C. Quan, C. Tay, and L. Chen, *A study on carrier-removal techniques in fringe projection profilometry,* Optics & Laser Technology **39**, 1155 (2007).

[139] Nederlandse Stichting voor Verlichtingskunde, *Verlichting in musea en expositieruimten* (Rijksdienst voor het Cultureel Erfgoed, 2008).

[140] I. C. Cuthill, *Color perception,* Bird coloration **1**, 3 (2006).

[141] F. Birren, *Principles of color: a review of past traditions and modern theories of color harmony* (Schiffer Publishing, 1969).

[142] Justin Lam, *Measuring the spectram characteristics of a light thereapy lamp,* (2018).

[143] P. E. Debevec and J. Malik, *Recovering high dynamic range radiance maps from photographs,* in *Proceedings of the 24th annual conference on Computer graphics and interactive techniques* (ACM Press/Addison-Wesley Publishing Co., 1997) pp. 369–378.

[144] L. Wareham, *Tilt and shift lenses,* (2005).

[145] E. Nocerino, F. Menna, F. Remondino, J.-A. Beraldin, L. Cournoyer, and G. Reain, *Experiments on calibrating tilt-shift lenses for close-range photogrammetry.* International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences **41** (2016).

[146] S. Khot, N. P. Yelve, R. Tomar, S. Desai, and S. Vittal, *Active vibration control of cantilever beam by using pid based output feedback controller,* Journal of Vibration and Control **18**, 366 (2012).

[147] M. Trindade, A. Benjeddou, and R. Ohayon, *Piezoelectric active vibration control of damped sandwich beams,* Journal of Sound and Vibration **246**, 653 (2001).

[148] S. K. Nayar, X.-S. Fang, and T. Boult, *Removal of specularities using color and polarization,* in *Computer Vision and Pattern Recognition, 1993. Proceedings CVPR'93., 1993 IEEE Computer Society Conference on* (IEEE, 1993) pp. 583–590.

[149] A. Ludwig, *The definition of cross polarization,* IEEE Transactions on Antennas and Propagation **21**, 116 (1973).

[150] M. Riechert, *All cameras,* web (2018).

[151] S. McHugh, *Cambridge in colour - macro lens,* web (2005-2018).

[152] NA, *Macro photography- table of macro photography lenses,* (2014).

[153] CameraLand, *Canon mp-e 65/2.8 1-5x macro,* (2017).

[154] B. Carnathan, *Canon ts-e 50mm f/2.8l tilt-shift macro lens,* web (2017).

[155] N. Mansurov, *What is a teleconverter,* web (2018).

[156] Canon, *Specifications ts-e 90mm,* web (2018).

[157] S. Ringsmuth, *Macro photography on a budget: An introduction to close-up filter,* web (2016).

[158] S. McHugh, *Cambridge in colour - extension tube,* web (2005-2018).

[159] S. S. G. E. Mert Keçeli, Joanne Wen, *Sem image of the week: Only their hairdresser knows for sure.* Midwood Science (2011).

[160] China Printing Company, *The thickness of printing paper list,* (2018).

[161] R. image change, *Gezondheid en veiligheid,* (2018), unpublished research.

[162] W. Streefland, *Measuring and evaluating ink film thickness data, getting technical,* (2006).

[163] Bruker Inc., *Contourgt-k 3d optical microscope datasheet,* (2018).

[164] T. T. Essers, *Manual fine art scanner,* (2017), unpublished research.

[165] M. Levoy, *Lectures on digital photography, lecture 3,* Online Google Lectures (2016).

[166] H. M. Merklinger, *ADDENDUM to FOCUSING the VIEW CAMERA,* (1995).

[167] S. McHugh, *Cambridge in colour - tilt-shift lenses,* web (2005-2018).

[168] M. Levoy, *Lectures on digital photography, lecture 6,* Online Google Lectures (2016).

[169] I. P. Pappas, M. R. Popovic, T. Keller, V. Dietz, and M. Morari, *A reliable gait phase detection system,* IEEE Transactions on neural systems and rehabilitation engineering **9**, 113 (2001).

[170] Nasim Mansurov, *How phase detection autofocus works,* (2018).

[171] A. S. Dhodapkar and J. E. Smith, *Comparing program phase detection techniques,* in *Proceedings of the 36th annual IEEE/ACM International Symposium on Microarchitecture* (IEEE Computer Society, 2003) p. 217.

[172] F. C. Groen, I. T. Young, and G. Ligthart, *A comparison of different focus functions for use in autofocus algorithms,* Cytometry Part A **6**, 81 (1985).

[173] M.-M. Cheng, N. J. Mitra, X. Huang, P. H. Torr, and S.-M. Hu, *Global contrast based salient region detection,* IEEE Transactions on Pattern Analysis and Machine Intelligence **37**, 569 (2015).

[174] D. Pelli, J. Robson, *et al.*, *The design of a new letter chart for measuring contrast sensitivity,* in *Clinical Vision Sciences* (Citeseer, 1988).

[175] G. Konig and P. Schornstadt, *A method and apparatus of contrast-dependent sharp focussing,* (1972), uS Patent 3,691,922.

[176] Kim Lewis, *How to determine the rpm on stepper motors,* (2017).

[177] R. A. Jarvis, *A laser time-of-flight range scanner for robotic vision,* IEEE Transactions on Pattern Analysis and Machine Intelligence , 505 (1983).

[178] T. Stahs and F. Wahl, *Fast and robust range data acquisition in a low-cost environment,* in *Close-Range Photogrammetry Meets Machine Vision*, Vol. 1395 (International Society for Optics and Photonics, 1990) p. 13951R.

[179] O. Bimber, D. Iwai, G. Wetzstein, and A. Grundhöfer, *Computer Graphics Forum*, Vol. 27 (Wiley Online Library, 2008) pp. 2219–2245.

[180] Museum Het Rembrandthuis, *The jewish bride,* (2017).

# Glossary

**CIE** Commission International de L'Eclairage.

**CIELAB** CIE L*a*b* colour space.

**CIELUV** CIE L*u*v* colour space.

**CMYK** Cyan Magenta Yellow blacK.

**DPI** Dots Per Inch.

**DSC** Digital Still Camera.

**DSLR** Digital Single Lens Reflection.

**HSV** Hue Saturation Value.

**ICC** International Colour Consortium.

**ISO** International Organisation for Standardisation (quantification of the light sensitivity).

**NPS** Nano Point Scanner.

**OCT** Optical Coherence Tomography.

**SNR** Signal to Noise Ratio.

**TOF** Time Of Flight.

# III

# Appendix: Instruction Manual

# Manual FAST Scanner

The FAST scanner focuses on ultra-high resolution colour and topography of the paintings, to digitally visualise the craquelure pattern of painting. This manual and some of the software were based or copied from the Fine Art Scanner, which focuses at scanning the visual appearances (color, texture and gloss) of a painting and reproducing it with 3D printing. The FAST scanner does not scan gloss, but there might be some software titles in this manual that still contain gloss, which is not part of the actual software.

This document explains in detail the work flow of the FAST scanning process. In seven chapter the entire scanning for capturing a painting is explained. This manual includes an introduction to the different processes, system requirements, the assembly of the scanner, acquiring data and image processing.

Written by *Tessa Essers* and *Mathijs van Hengstum*

TU Delft

# Table of contents

# 1. Introduction

This manual explains in detail the work flow of the fine art scanner. It provides information about the software and the different actions executed by the user and the programs Matlab and Arduino. The total scanning process has been divided in five sub processes to create a clear overview for the scanning process. In this chapter you can find information about what these sub processes are and its related actions.

## 1.1. Sub processes

The total process for capturing a painting can be divided in five sub processes and these have to be executed in a specific order, see figure 1. The start of the scanning process depends on the status of the scanner. For example, the scanner is already build up and calibrated, then the system can be started directly and the calibration process can be skipped.

In the sub processes the following actions are covered:

- Assembly scanner: the actions to assemble and start the scanner.
- Calibration: calibrate the position of the cameras and the color.
- Capturing: the actual scanning of the painting.
- Image processing: the process for creating color, height and gloss maps and the stitching of these maps with each other.
- Disassembly scanner: the actions to disassemble the scanner and finish the scan.

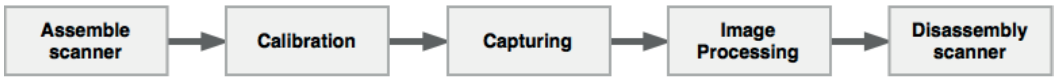



***figure 1.*** *The order in which the sub processes of the scanning process need to be executed.*

## 1.2. Actions

Each sub-process can be divided in three categories: the user actions, Matlab actions and Arduino actions. These categories do not have to be carried out in a specific order, but the order of the categories can be shifted within each sub process.

The actions can be described as the following:

- User actions: these are the actions that have to be executed manually by the user.
- Matlab: these are the actions executed by the program Matlab.
- Arduino: these are the actions executed by the program Arduino.

## 1.3. Total scanning process

An overview of the scanning process can be found in figure 3. On the horizontal axis, the sub processes are shown and on the vertical axis you can find the actions. In this document the sub processes have each their own chapter and will be explained further in detail.

The overview of figure 3 shows a good structure of each process with its corresponding

### 1.4. Capturing process in museum

sub processes. Scanning in this order is not convenient in time and use when you want to scan a painting in a museum. Due to time limitations you do not want to spend lot of time on the image processing. Thus, all image processing that can be conducted after capturing the painting is better.

The total scanning process gets a different order (see figure 2) when the before mentioned case is applied to the scanning process in a museum. All the sub processes of the calibration process need image processing to acquire the data. However, not all the image processing parts are needed before capturing the painting, the illumination and color calibration are only needed for the processing of the painting images. Additionally, you must know what kind of camera settings are used in the capturing process since the calibration objects must be captured with exactly the same settings. Thus, the sub processes of the calibration are split in two groups, the white balance and camera calibration are executed before the capturing sequence and the illumination and color calibration after the capturing sequence.

Another point that has changed from figure 3 is that the 'Image processing' will be conducted after the disassembly of the scanner. The processing is not a step which has to be conducted in a museum, but can be done afterwards.

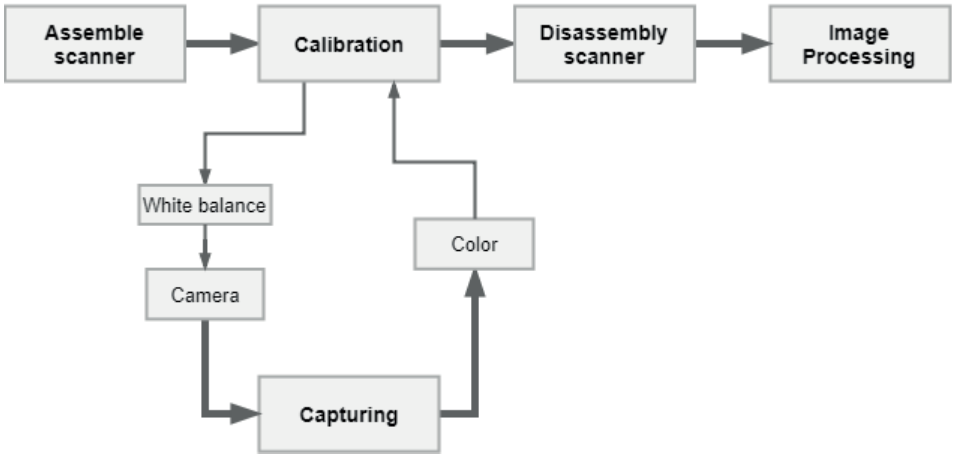
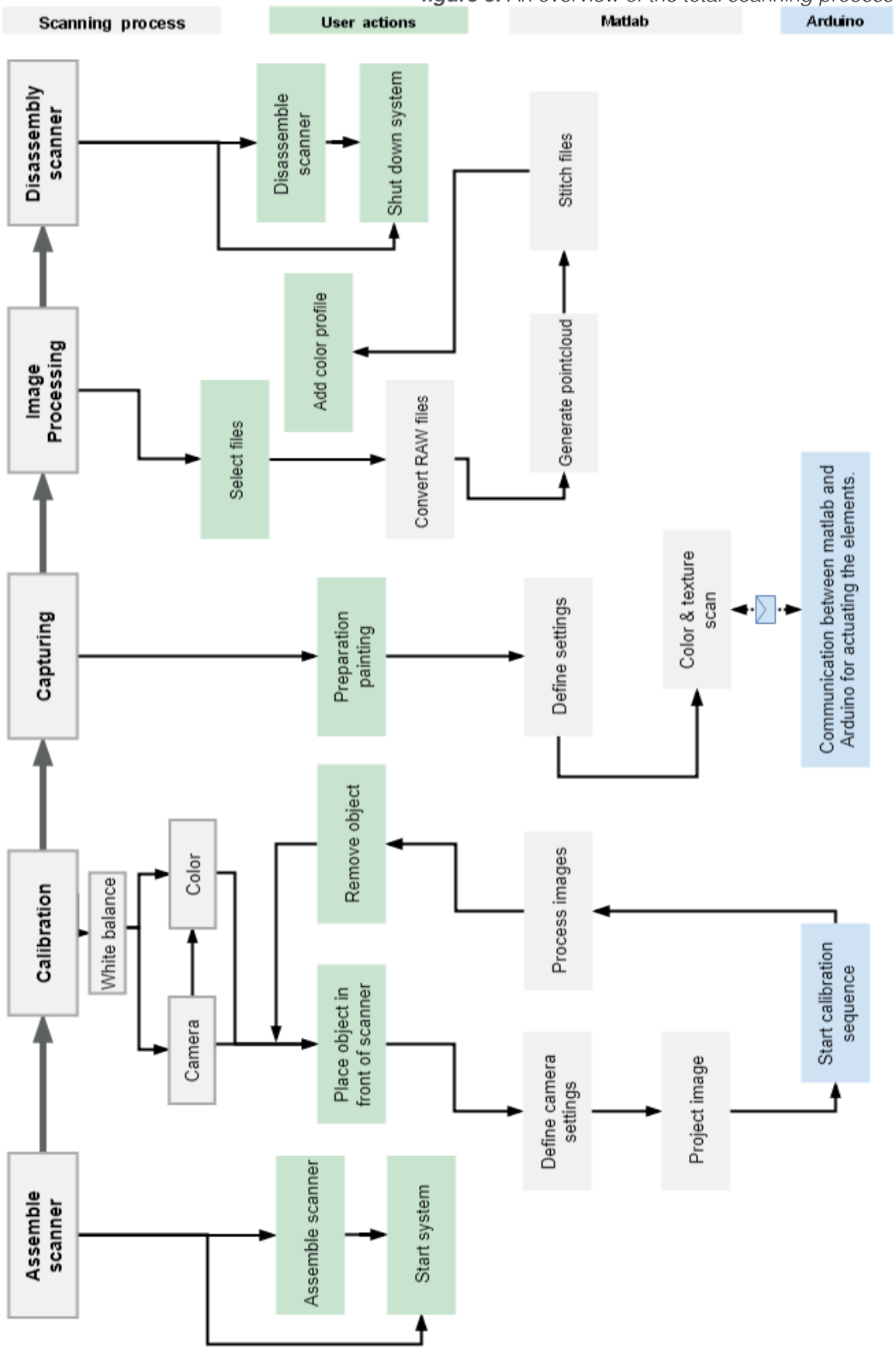



***figure 2.*** *An overview of the scanning process for museums.*

***figure 3.*** *An overview of the total scanning process*



# 2. System requirements

For the software of the fine art scanner several programs and toolboxes are required . Our laptop contains windows 7, if you have a newer version of windows we cannot guarantee that all the programs and toolboxes are compatible.
The following programs and toolboxes are required to actuate the different elements of the scanner:

Matlab 2017a. The calibration and capturing process use Matlab 2014b, however the image processing include some functions which are only available in Matlab 2017. This version of Matlab is needed for the stitching of the images.

Arduino. Two microcontroller boards of Arduino are used for actuating the different elements of the scanner and moving the scanner in horizontal and vertical position. Matlab sends messages to Arduino and Arduino again know what kind of action must be executed per message. Arduino is only needed when you want to communicate directly with the microcontrollers.

Gphoto2 and camera drivers. Matlab uses the program gphoto to trigger the cameras and transfer the images from the cameras to the computer. You have to install libusb-win32. In case you want to use a different type of camera then the current system, you have to download the latest version of libusb-win32 on http://sourceforge.net. Then install the following folders and software.

- set CAMLIBS=camlibs
- set IOLIBS=iolibs
- gphoto2.exe

Pschytoolbox-3. This toolbox is needed for to project images with the project. You can download it from the following website: http://psychtoolbox.org. Read the system requirements for this installation, since GStreamer 1.x is required for the pschytoolbox.

dcraw-9.16-ms-64-bit This program is necessary for the conversion from .NEF to .tif. With matlab will be communicated with the program. You can find this in the script convert2tiff.m

Argyll The color calibration is conducted partly with matlab and the final creation of the icc profile is created with the program Argyll. Download the executable for your computer system on argyllcms.com.

vlfeat0.9.18 This program is needed to find the features for stitching.

# 3. Elements FAST

The FAST exists of several elements. This chapter provides a list of the different elements illustrated with photos. In the next chapters the names of these elements will be used to explain the software and system. This chapter can be used as a reference when you do not know what one of the elements is. Every element is numbered and in the pictures these numbers are connected with the corresponding element.

1. Device box
2. Camera 1 (left)
3. Camera 2 (right)
4. Projector
5. Arduino board
6. Electric box XY-stage
7. On/off switch XY-stage X-axis
8. X-axis
9. Y-axis
10. Z-axis
11. Rotation about x-axis belt
12. Camera calibration checkerboard
13. Grey-charts for WB
14. Micro Colour target
15. Projector Cable
16. Camera USB Cables (Left and Right)
17. Arduino USB Cable
18. XY-stage control cable
19. Extra usb and Dvi cable

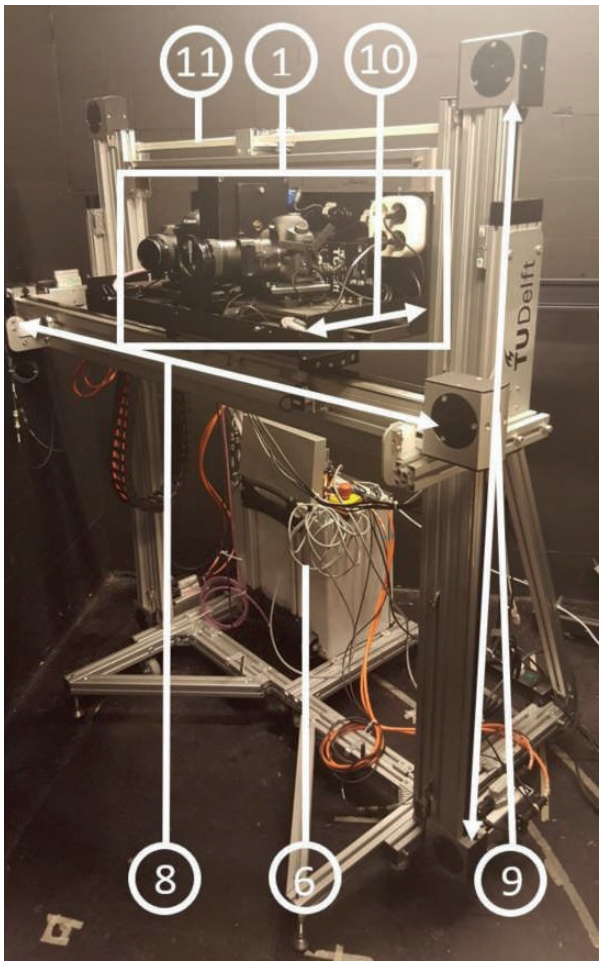



***figure 5.*** *FAST scanner*

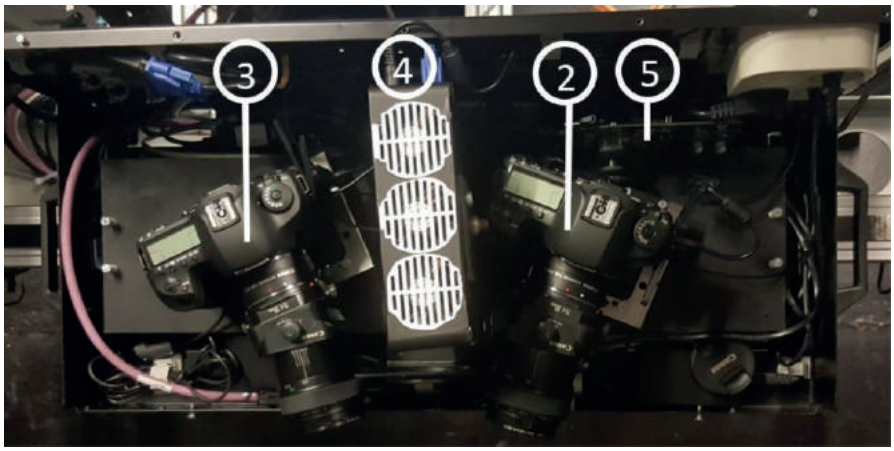



***figure 4.*** *Top view of platform*

***figure 6.*** *Electronics box*

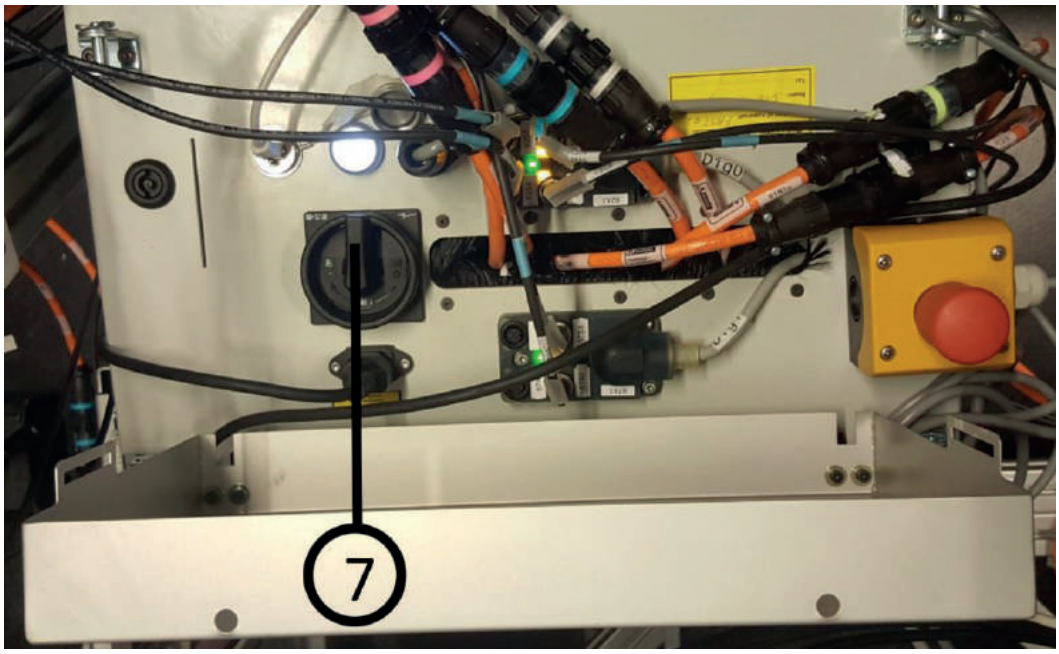



***figure 7.*** *Calibration objects*

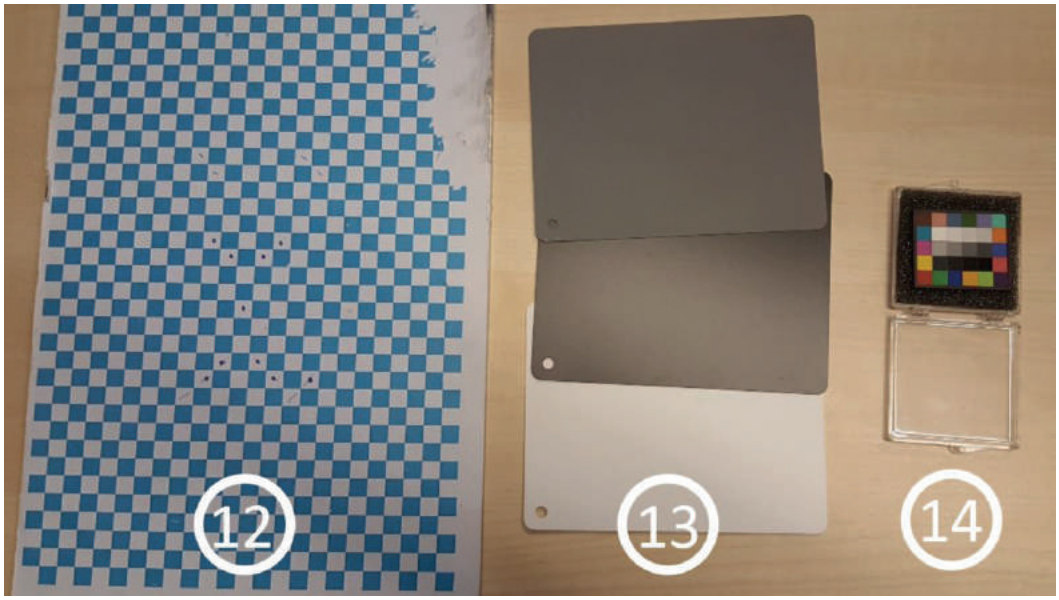
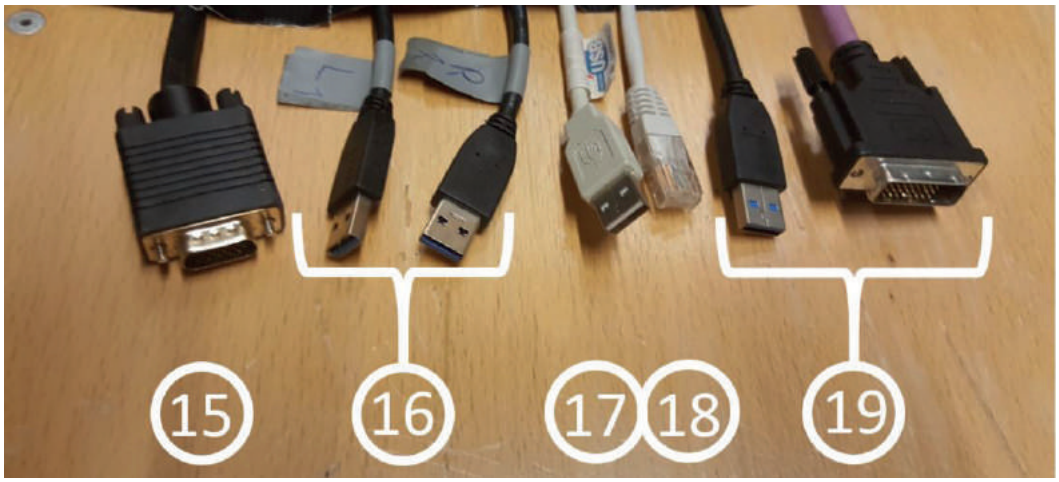



***figure 8.*** *Cables of the scanner*



# 4. Assembly / disassembly

The assembly is the start of the scanning process and is only needed when the scanner is moved to a different place, for example when an artifact is scanned in a museum or at the university when it is disassembled and placed in a different room. Only user actions are needed for this sub process, Matlab and Arduino are not yet used.

In figure 9 the subparts of the assembly process are shown. The user actions can be divided in two sections: the actual assembly of scanner and the start of the system. When the scanner is already set-up in position only the start of the system need to be checked.

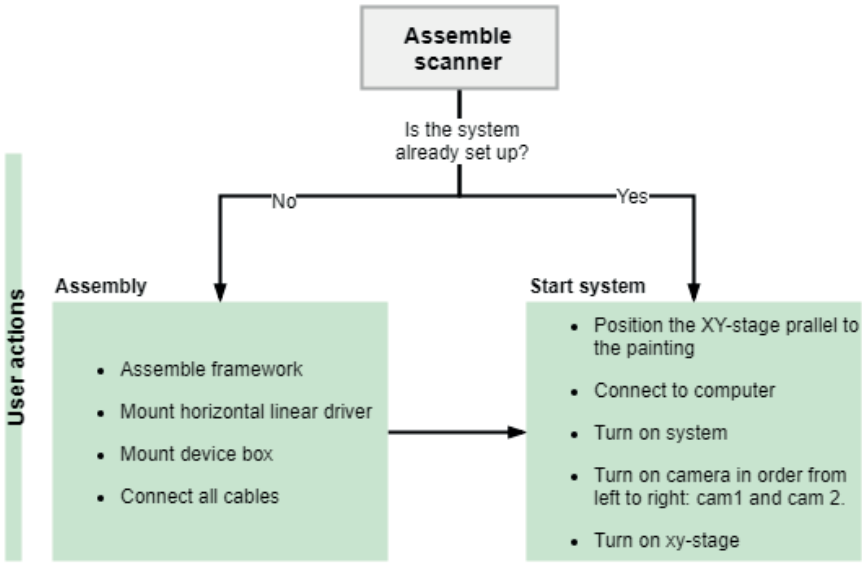



***figure 10.*** *A zoomed in image of the user actions within the assembly process.*

## 4.1. Assembly

The scanner only needs to be assembled when it has been transported to a new location and must be placed in dark room, without any stray light of the environment light. The order for the assembly is described below. This order will start at the moment the scanner arrives at a new location. The assembly exists of three steps:

1. The FAST is driven to the location and unloaded. It will be placed on a trolley to move it around at location.
2. The 3D FAST is assembled according to the 'Operation Manual FAST Scanner', but shortly summarized:
   a. Connect the base of both vertical linear drivers
   b. Connect the 3 upper rods between the vertical linear drivers
   c. Mount the Rotation about x-axis belt
   d. Mount side holders of vertical linear drivers
   e. Mount safety racks
   f. Fix the horizontal linear driver

g. Install electric cable guide
h. Mount device box
i. Connect all cables according their marking/numbering
j. Adjust cameras and projector

### 4.2. Start system

For the start of the scanner three tasks need to be checked. The first one is to check if the scanner is connected to the computer. The system cannot be actuated when the USB ports of the device and serial port of the XY stage are not connected with the computer. Thus, check if all the USB cables of the cameras and the Arduino box is attached to the computer. Turning on the system is an obvious step; the scanner need to have power to work. The switch on the electric box has to be turned and after 5 seconds the flashing button needs to be pushed to turn on the XY-stage. Make sure the firewall is turned off for the serial port and the FastMotionDriver.exe is active to control the XY-stage. Moreover, all the elements of the platform are connected to a power supply unit and this unit is often turned off too. Thus, take care that this unit is also turned on. Besides make sure that the projector and cameras are turned on. The last and most important step for the start of the system is that the cameras need to be turned on in a specific order from left to right. Thus, first camera 1 and then camera 2. When this does not happen in right order the number that the camera receives in Matlab will be wrong and thus the images of the wrong cameras will be used for the calibration and image processing. This can result to a crash of the system.

### 4.3. Disassembly

The disassembly of the scanner is the end of the scanning process and is exactly the same as the assembly but then in the opposite order. First the system needs to be turned off before you disassemble the framework. When you want to scan another time the painting at another moment then you do not have to disassemble the framework, but just turn off the power.

# 5. Calibration

With the calibration process the position of the cameras and the color of the images will be calibrated. For each part, another object is photographed; the user places this object on the easel. An overview of the calibration process can be found in figure 10. The loop for the user actions, Matlab and Arduino repeats for each sub process of camera, color and gloss calibration.

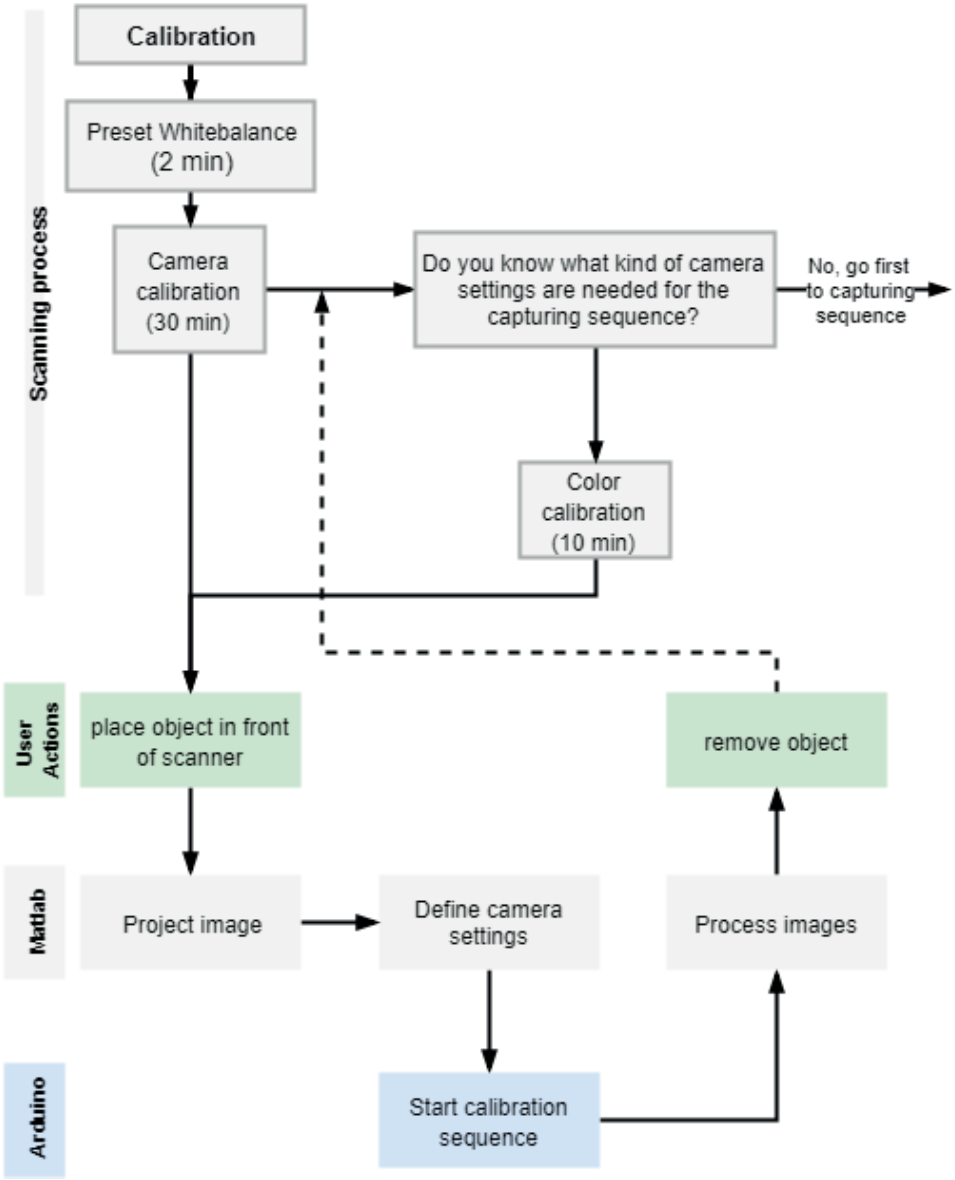



***figure 11.*** *An overview of the calibration process*

When the *'start system'* has not been executed before starting with the calibration, then redo this step of the assembly and start the whole system. It is especially important how you have to turn on the cameras.

## 5.1. Graphic User Interface (GUI) for Calibration processes

The different sub processes of the calibration process can be controlled with a graphic user interface of Matlab. With this user interface you do not have to open the functions for all the sub processes separately. This GUI can be found in the folder 'paintscan3D / Matlab_projector' under the name 'gui_calibration.m'.

Before opening the gui the global_path must be linked to the right folder where the captured images are placed in. This line can be found in the beginning of the script (around line 74) and need to be written as the following:

```
global_path='O:\nameFolder\';
```

This folder must contain a main folder named 'Calibrations'. If this global_path does not contain this folder, a calibration folder will automatically be created when you first start the GUI. When you run the script the GUI will pop up (figure 11).

In the left side of the GUI in the panel 'preparation calibration' the settings for the cameras, projector and platform can be defined. On the right side the separate processes for the camera, LED illumination and color calibration are explained.

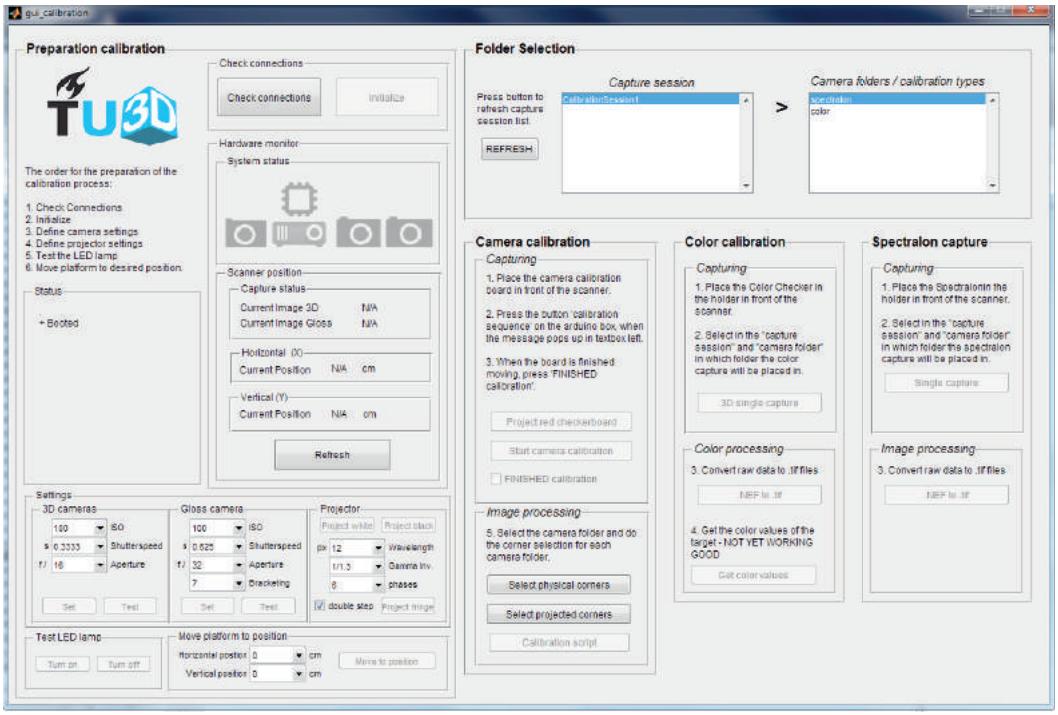



***figure 12.*** *The graphic user interface for the calibration process.*

## 5.2. Preparation calibration process

As preparation for the calibration processes you create a connection between the scanner and Matlab for communication and actuating. It is important that you execute the step 'check connections' and 'initialize' first. The order of the other steps does not matter.

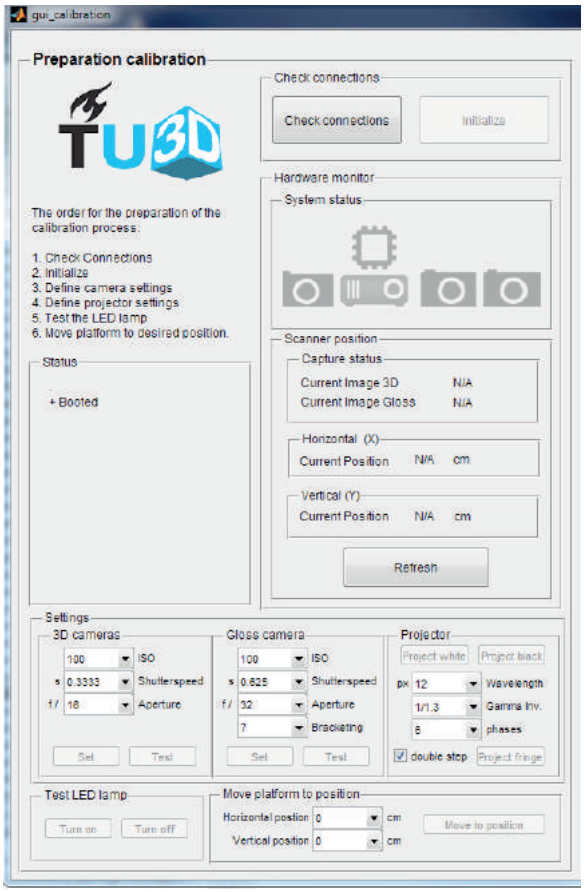



***figure 13.*** *Preparation calibration panel*

### 5.2.1. Check connections / initialize

When you click on the button 'check connections', Matlab will search for the COM ports of the Arduinos and initializes the cameras and projector. Each element will turn blue when they are found. If all appear blue the 'check connections' has been executed successfully.

Often a problem occurs with the search for the COM ports for Arduino. The following message must appear in the command window of Matlab:

```
Checking System Status.. Please
Wait..

portlist =
'COM1'
'COM3'
```

An error message is shown when only one COM port appears. You have to restart matlab or unplug and replug the Arduino USB cable in the computer. The next step is to check another time the connections. It can happen that the Matlab script still cannot find the COM port, then repeat this process till Matlab can find them. Another option is to restart the computer. Try to keep patience when it does not work immediately.
**Note!!** When you unplug the Arduino cables you have to restart the cameras in the specific order.

The 'initialize' button can be pressed when the check connections has been completed successfully. This button makes the communication portal available between the cameras and Matlab. It needs to be pressed before defining all the separate settings.

### 5.2.2. Camera settings

For the cameras we have to define three settings: aperture, ISO and shutter speed. With the current cameras we cannot define the aperture with Matlab, but this is done manually on the camera. Just place the same aperture in this field as defined on the camera. The aperture is f16 for cam 1 & 2. The ISO is always at the lowest number (100) to avoid noise and the shutter speed can be changed with Matlab. You can test the shutter speed if you are unsure which shutter speed is the best for the object that you want to scan. First you need to select the 'set' button for defining the camera settings before making a test image. Under the settings for the cameras a test button is placed to test the shutter speed of the cameras.

The scanner will take one image of the painting when you press this button. This have to be done separately for the 3D cameras and the gloss camera. This test image will be taken in the position where the platform of the scanner is positioned at that moment; most often this shall be in the zero position. The platform can be moved with step 5.2.4.

### 5.2.3. Projector settings

With the projector settings we define the fringe pattern that is used to make the height map and next to that a black or white image can be projected. The black projection is used for the gloss capturing in combination with the LED light source. For the calibration processes we do not need the fringe projection but only the black and white image. In case you need the fringe projection then the option is there.

### 5.2.4. Movement platform

With this button you can move the platform to the desired position in front of the painting. You decide how much the platform must move in horizontal and vertical position in meters. **NOTE!!** in the GUI it is written as cm but for this scanner it is meters (m).

### 5.2.5. Status bar

In the status window the settings for the system are visible. In here you can find an overview of all the settings that you have filled in. When some of the settings will change you see the change in this window.

## 5.3. White balance

The white balance must be set before you capture any object placed in front of the scanner. Otherwise, if you miss this step, the colors of the images can be different than in real-life. Next to that, the white balance of the cameras are predefined per light source. When we do not adjust the light source for each situation, the images can have a strange color balance and can have a pink or blue appearance.

The white balance (WB) is different per light source. This means that the projector and LED light source have a different predefined WB in the camera. Camera 1 and camera 2 only need a custom white balance for the white projection on the grey chart.

A white reference target (the grey chart) must be placed in front of the scanner in such a way that the complete image is filled with the target (or most of the image) . Then turn on the light source; the intensity of the projector is the same as with the color (see GUI for calibration process). When the illuminates the grey chart, the white balance can finally set manually on the cameras. Please refer to the manual of the camera[1] of how to set the white balance.

## 5.4. Camera calibration

With this process, we calibrate the positions of the three cameras in respect to each other and the painting. This process takes around two hours if the corner clicking is executed perfectly. The camera calibration is needed for the creation of the color, height and gloss map: the images of the left and right camera are matched with each other and the result is a height and color map of the surface and the image of the gloss map is projected on top of the color and height map. In figure 13 you can find the part of the camera calibration in the GUI. Every button represents another action.

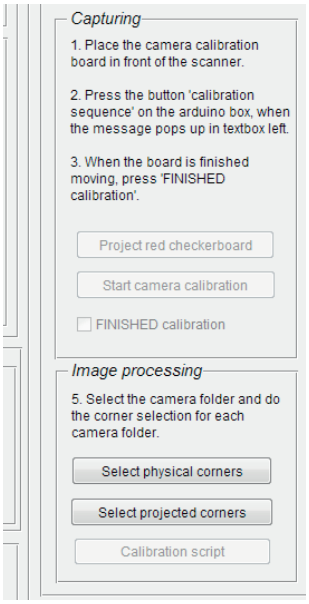



***figure 14.*** *The actions for the camera calibration in the right order.*

1 http://gdlp01.c-wss.com/gds/9/0300018669/01/eos5ds-5dsr-im-en.pdf

The order of the steps is described in the GUI. Two actions are conducted manually, 1) pressing the button 'calibration sequence' on the Arduino box and 2) check the 'FINISHED calibration' checkbox. In appendix C the complete matlab scripts of these buttons can be found

### 5.4.1. Position camera calibration board

A checkerboard is placed in front of the scanner for the camera calibration. First you need to attach an arm in front of the scanner; place this one between the triangular holders. Then the calibration board can be clicked in front of the scanner, check if the distance between the board and the laser distance meters is 441 mm (±1mm).

### 5.4.2. Define camera and projector settings

The marked 7x3 squares of the checkerboard must be well visible on the photos of all three cameras. This to know for sure that the cameras capture an image of the same part of the painting. We pasted on the checkerboard some markers to mask this area.

The next step is the camera settings. Cam 1 and 2 must have an aperture of F16 and a shutter speed of 0,6 second. The shutter speed can change depending on the environment light; the more environment light, the shorter the shutter speed. The shutter speed of the cameras can be defined in the GUI, see step 5.2.2.

The last step for the camera settings is the position of the polarization filter of cam 1 and cam 2. Rotate the polarization filter to filter away the white dots in the blue checkerboard, as shown in figure 14. Only rotate the polarization filter, do not rotate the lens of the camera.

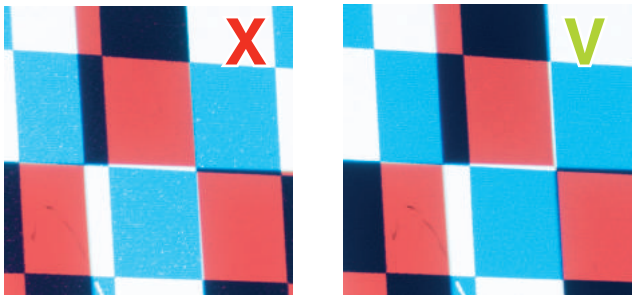



***figure 15.*** *The differences between the images when you rotate the polarization filter of cam 1 and cam 2. The right image shows the result what you want to see.*

### 5.4.3. Start calibration sequence

You have to restart the cameras in the right order (cam1, cam2) when you unplugged the USB cables of the cameras in step 5.4.2. Press 'start calibration sequence' to start the camera calibration.

Fill in the number of images you want for calibration in the GUI. We suggest at least 20 for an accurate camera calibration. Next start the camera calibration sequence in the GUI. Change the orientation of the camera calibration checkerboard after each photograph until all images are captured.

### 5.4.4. Corner clicking

For the corner selection it does not matter if you first select the physical corners or the projected corners. The corner selection needs to be executed for cam1 and cam2. Before you select the button 'select physical corners' you need to select the desired camera folder. The most important part of this step is that you have to click the exact same corners for all camera folders. If this is not the case the calibration is completely wrong. We use the markers on the calibration board as help.

First the script must find the images. In the command window of Matlab you have to give values to search for the images. First it asks for the image format, .tif (t). Next the window size for the corner selection have to be defined, wintx = 30 and winty = 30. This window size is so small to have an accurate result for the corner selection; otherwise the wrong pixels for the corners are automatically selected.

Then the corner selection can start. In each image, you need to select exactly the same area. Take into account that when you mismatch one corner that you have to redo the whole process. When you select the outer corners of the first image some other questions appear. First you have to provide the number of squares in the x and y direction. Count the number of the white and black squares in the specific direction. This question can appear more often in the command window, when you slightly click on a different spot for the corners. Next to that you have to fill in the size of one square in x and y direction. This is 7.692 mm (of our calibration board). Once all the corners of the images are selected a value appears for the pixel error. This value has to be approximately 1; with our set-up we have often a value between 2 and 1. When this value is higher you have to redo the total corner selection for the set you just selected. The process described in the two previous paragraphs repeats itself for the next corner clicking.

### 5.4.5. Remove calibration board

At the end of the camera calibration the calibration board can be removed. Depending on the next process the following object can be placed in position.

## 5.5. Color calibration

For the color calibration another object, the color board, need to be placed in front of the scanner. This object is used to calibrate the colors how they are seen under the specific lightning conditions. The result of this color calibration is a color (.icc) profile with which we can change the color map of the painting, so we can reconstruct the same colors of the painting on the 3D printed replica.

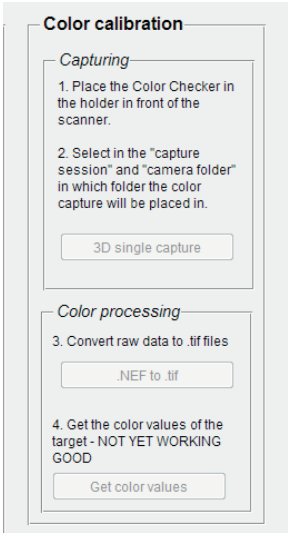



***figure 16.*** *Calibration GUI for the color calibration*

### 5.5.1. Position color target

First position the color target, artist paint target, with 175 mm (±1mm) in front of the scanner.

### 5.5.2. Define camera settings

Adjust the camera settings of cam 1 and 2 in the GUI, aperture F16, shutter speed same as capturing process.

### 5.5.3. Capture color target

Select in the list in which folder the images of the capture can be placed and then capture the color target. This can easily be done using the button '3D single capture' in the GUI.

### 5.5.4. .CR2 to .tif

.tif files are needed for the eventual process of getting the color values. Press button '.NEF to .tif' to convert the images.

### 5.5.5. Get color values

8Only one image is needed for the color values. In this script the first .tif image of the selected folder will be used to retrieve the color values. With the button 'Get color values' the script can be started. The color reference target has 4 rows and 6 columns with different colors and each square must be manually selected. An message appears which square you need to select in the right order. The output of this script is an excel file containing R, G and B values that are scaled from 0-65536 (16bit), 0-255 (8-bit) and 0-100. The

matlabscript for the color extraction can be seen below. The ICC profile is automatically generated in the color folder.

```
1. 1. for i=1:30
2. 2. ColourList = {'1' '2' '3' '4' '5' '6' '7' '8' '9' '10' '11' …
3. 3. ... '12' '13' '14' '15' '16' '17' '18' '19' '20' '21' '22' ...
4. 4. ... '23' '24' '25' '26' '27' '28' '29' '30'};
5.      message1 = sprintf(['Zoom in to area ' ColourList{i}]);
6.      message2 = sprintf(['Select colour area of ' ColourList{i}'...
7.      ...and double click']);
8.      uiwait(msgbox(message2, message1));
9.
10.     Crop=imcrop(C);
11.     RobustMean16bit_R= mean2(Crop(:,:,1));
12.     RobustMean16bit_G= mean2(Crop(:,:,2));
13.     RobustMean16bit_B= mean2(Crop(:,:,3));
14.     RobustMean8bit_R= (mean2(Crop(:,:,1))/2^16)*255;
15.     RobustMean8bit_G= (mean2(Crop(:,:,2))/2^16)*255;
16.     RobustMean8bit_B= (mean2(Crop(:,:,3))/2^16)*255;
17.     RobustMean0_100_R= (mean2(Crop(:,:,1))/2^16)*100;
18.     RobustMean0_100_G= (mean2(Crop(:,:,2))/2^16)*100;
19.     RobustMean0_100_B= (mean2(Crop(:,:,3))/2^16)*100;
20.     Values(i).Name = ColourList{i};
21.     Values(i).R_16 = RobustMean16bit_R;
22.     Values(i).G_16 = RobustMean16bit_G;
23.     Values(i).B_16 = RobustMean16bit_B;
24.     Values(i).R_8 = RobustMean8bit_R;
25.     Values(i).G_8 = RobustMean8bit_G;
26.     Values(i).B_8 = RobustMean8bit_B;
27.     Values(i).R_100 = RobustMean0_100_R;
28.     Values(i).G_100 = RobustMean0_100_G;
29.     Values(i).B_100 = RobustMean0_100_B;
30. end
31.
32. message1 = sprintf('get colorValues');
33. message2 = sprintf('Finished selection');
34. uiwait(msgbox(message2, message1));
35.
36. save([capt_path_now '\colorValues.mat'], 'Values');
37. ValuesTable=struct2table(Values);
38. writetable(ValuesTable, [capt_path_now '\colorValues.xls']);
```

### 5.5.6. Assign color profile

Navigate to the folder /Library/ColorSync/Profiles and copy the profile here. Open the color file of your 3D scan in adobe Photoshop and go to edit > assign profile and select new color profile. 'Save as…' file, make sure to tick the box 'embed color profile' here; this should mention the name of your profile here.

# 6. Capturing

The capturing process is the actual process of capturing the painting. The previous two processes were just for the preparation of the scanner. Within the capturing process we will capture the color and surface texture of a painting. In figure 16 you can find an overview of the capturing process. The first two steps, ‘preparation painting’ and ‘define settings’, are still the preparation for the scanning process. The other steps is the actual capturing of the painting, in these steps Matlab and Arduino communicate with each other to actuate the different elements of the scanner.

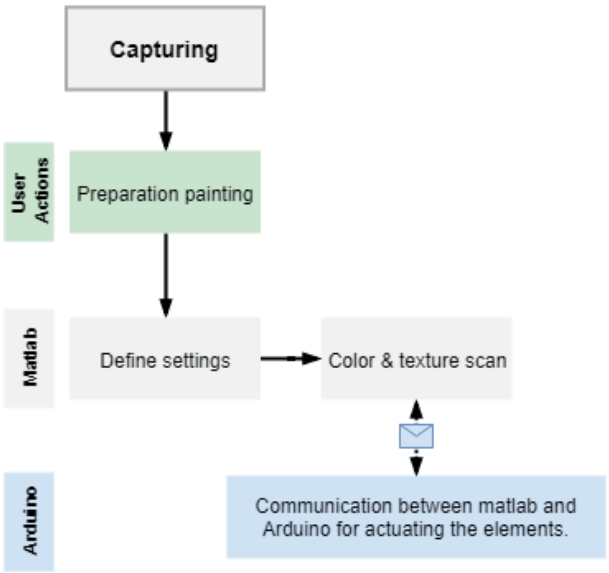



***figure 17.*** *An overview of the capturing process*

## 6.1. Preparation painting

The most important part of the capturing process is that the painting is positioned in front of the scanner with the right distance between scanner and painting. In total four actions are needed.

### 6.1.1. Position painting

Place the painting on the easel. Be sure that the painting is placed straight on the easel, thus that you level it.

### 6.1.2. Measure dimensions

The dimensions of the painting are needed to define the amount of steps the platform has to move in horizontal and vertical position. This will be defined with the matlab graphic user interface, see 6.3.2. Measure the horizontal dimensions, from left to right edge, and vertical dimensions, from bottom to top edge.

### 6.1.3. Position easel

With the calibration process the system has been calibrated with 175 mm (±1mm) of the scanned object. Thus, the painting has to stand in the exact same position.

### 6.1.4. Blur projector

The projector is turned a bit out of focus to blur the edges of the fringe pattern; otherwise we can see the pixels of the projector on the painting. The projector will be used a second screen when nothing will be projected by matlab. The background shows the TU Delft logo, when the logo is focused it consist out of one clear line (A in figure 17), when you slightly blur it the middle of the torch will change into a 'double' line' (B in figure 17).

**Note!!** Only change the settings of the projector and the polarization lens, do not rotate the

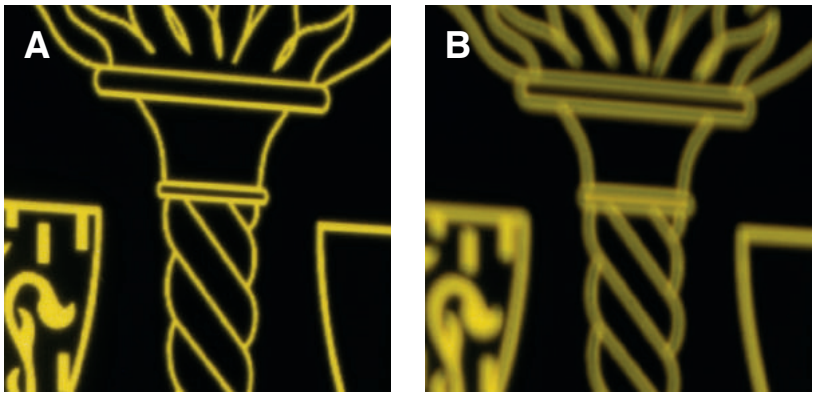



***figure 18.*** *Projector in focus (A) and projector out of focus (B)*

camera lens and the other elements!

### 6.1.5. Delete images cameras

Delete all images from the camera to know that only the images of what you want to capture are available on the computer.

## 6.2. Graphic User Interface (GUI) capturing process

With the GUI you define the settings for the capturing process and start the capturing sequence. Besides a start with the image processing can be made. The Matlab script 'gui_3Dscanner_gloss_new_processing.m' is used for the capturing process, it can be found in 'paintscan3D / Matlab_Projector'. In the appendices, you can find parts of the script, due to the length the total script is not added in this document.

First you need to change the global_path in the script; this is the folder where the images will be placed during the capturing process and where the calibration data can be found. This line is placed in the beginning of the script (around line 78) and has to be written as follow:

```
global_path='O:\scan_painting\';
```

The line between the apostrophe must be adjusted to your directory. This folder needs to contain four sub-folders: 'Calibrations', 'Captures', and 'Tests'. Matlab creates these folders when it cannot find these folders in the global_path. This only happens when you run the script. Within these sub-folders more sub-sub-folders can be placed to divide your capture sessions.
For the 'Calibrations' folder it is important that it contains the latest camera calibration folder, this can have 'yyyymmdd' name. Otherwise you cannot find the right calibration file in the calibrations list. When the script is run, the Graphic User Interface (GUI) for the capturing process will pop up (figure 18).

**Note!!** These sub-sub-folders only appear in the capture sessions and calibration list when you restart the GUI.

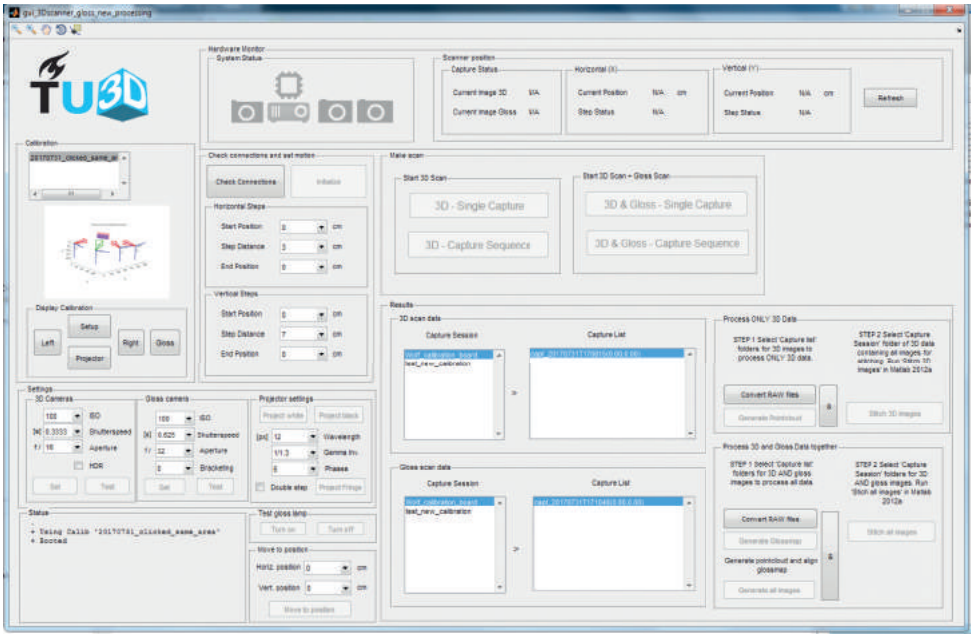


***figure 19.*** *The graphic user interface for the capturing process.*

On the left side of the user interface the settings for the cameras, projector and step size can be defined. Next to that a test capture of a part of the painting can be made to analyze if the cameras have the right shutter speed. On the right side the capturing processes can be started. You have to choose if you want to make a single capture or if you want to run the total capture sequence for the 3D. In the bottom right, the created images can be processed. More information about the image processing will be provided in chapter 7.

## 6.3. Define settings

The part of the GUI to define the settings is partly similar as the GUI for the calibration. The buttons which are the same are explained but you have to refer back to chapter 5.2. The settings need to be defined in a specific order. It is important that you first select the calibration data. Then check the connections and initialize cameras before defining the camera settings and projector settings. Otherwise problems will occur further in the process.

### 6.3.1. Calibration files

First you need to select the data of the calibration you just made to load the position of the cameras. Select the good one! When you click on the buttons left, right and gloss you get the calibration image for cam1 (left), cam2 (right) and cam3 (gloss). This image appears on the place that shows in figure 19 the positions of the cameras.

### 6.3.2. Horizontal / vertical steps

In the user actions, you have measured the dimensions of the painting. Thus, we know where the painting is placed in horizontal direction and the distance it has to cover, the same for the vertical steps. The step size is defined in cm. Take in mind that there is enough overlap for the stitching of the images. With the predefined settings this overlap is good, for the horizontal steps this is 6 cm and for the vertical steps this is 4 cm.

### 6.3.3. Projector settings

With the projector settings we define the fringe pattern that is used to make the height map of the paint visible. The fringe is created by the phases, wavelength and gamma. The phases are set on 6 for the best result. It defines the amount of times that the fringe pattern changes during the capturing process. The wavelength defines the number of pixels between the fringes and is set on 12. This is needed for the calculation. Gamma can stay on the predefined setting and the double step can stay unchecked.

**Note!!** A warning can appear when you select 6 phases. This warning can be ignored.

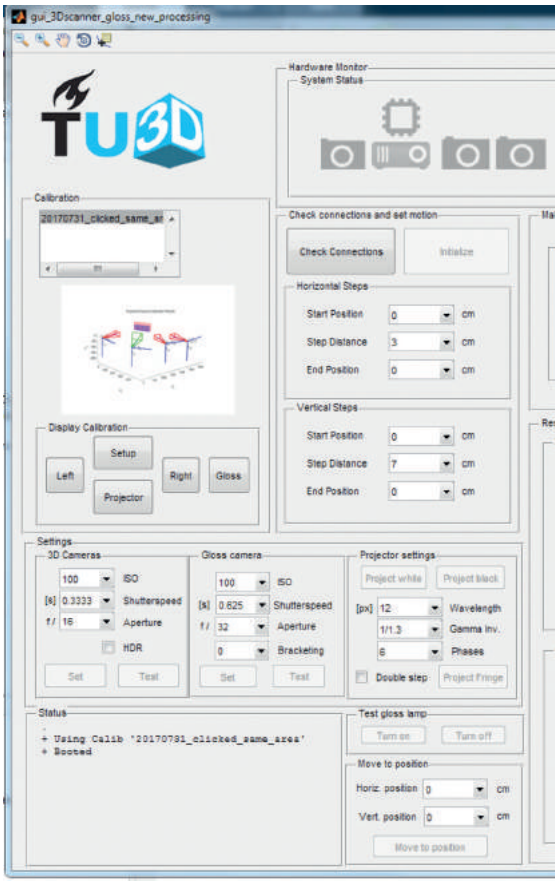



***figure 20.*** *Define settings capturing process*

### 6.4. Capturing sequence

Once all the settings are defined, the system is ready for the capturing process. For the capturing sequence you have to choose what you want to scan. Do you want to make a single capture of the color & texture (3D)? Or do you want to capture the whole painting? In the top right of the GUI the position of the scanner is shown (top part of figure 18), so you know the current horizontal and vertical position.

In the GUI the capturing sequence can be started by pressing the buttons in figure 20. Within this sequence matlab communicates with the arduino Uno and Mega to trigger the cameras of the scanner.

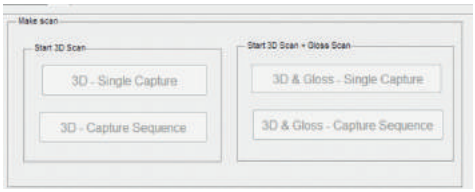



***figure 21.*** *Start capturing sequence*

In figure 21 the sequence for scanning the color and surface texture of the whole painting is shown. When you want to have a single capture, matlab goes through the loop once. Then it makes 12 pictures for the 3D scan; six time the fringes in horizontal direction and six times in vertical direction for both cameras.

The FAST system has two camera triggering methods. Depending on the available scanning time the capture can be triggered with Gphoto2 or with the Arduino.

The gphoto2 method takes about 10 times more time (390 seconds per tile), but the file naming and positioning in the folder are automatically done. Therefore, this method is implemented as standard.

Triggering the cameras with an Arduino takes about 40 seconds per tile, but the images are saved on the CF card in the cameras. The images have to downloaded, named, and placed in the correct folder.

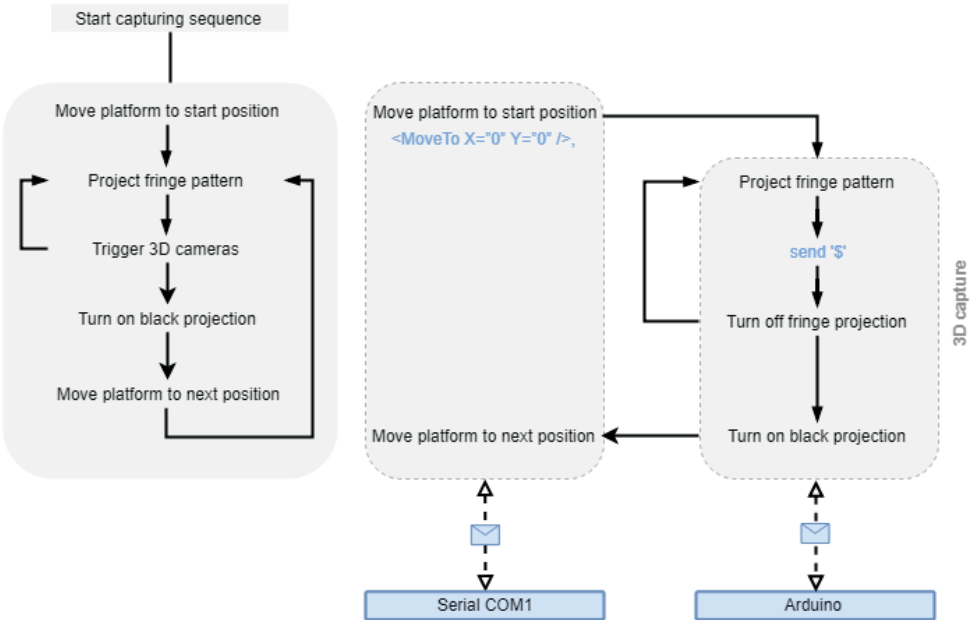



***figure 22.*** *Overview of the communication between Matlab and Arduino.*

# 7. Image Processing

Image processing is the process from the raw image files towards the 3D printing files. The biggest part of the image processing can be executed in within the GUI of the capturing process. However, some final steps to finalize the 3Dprinting files need to be executed with Adobe Photoshop. In figure 22 an overview can be found of the processing.

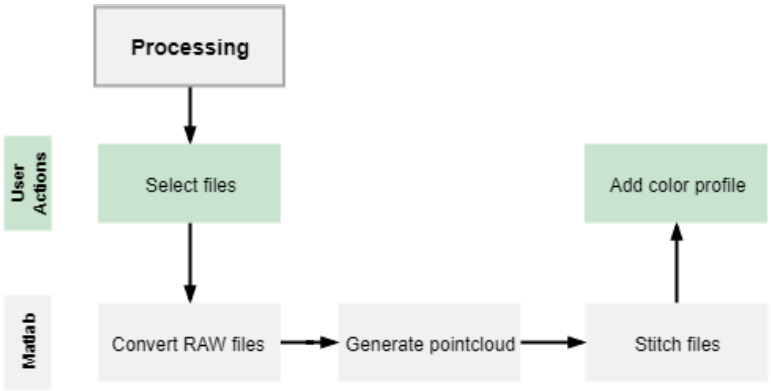



***figure 23.*** *An overview of the Image processing*

## 7.1. GUI processing

You can find the scripts for the image processing in the GUI of the capturing process. . The scripts for image processing are already integrated in the GUI and only the buttons need to be pressed. In the right bottom corner of the GUI you can find the process part (figure 23).

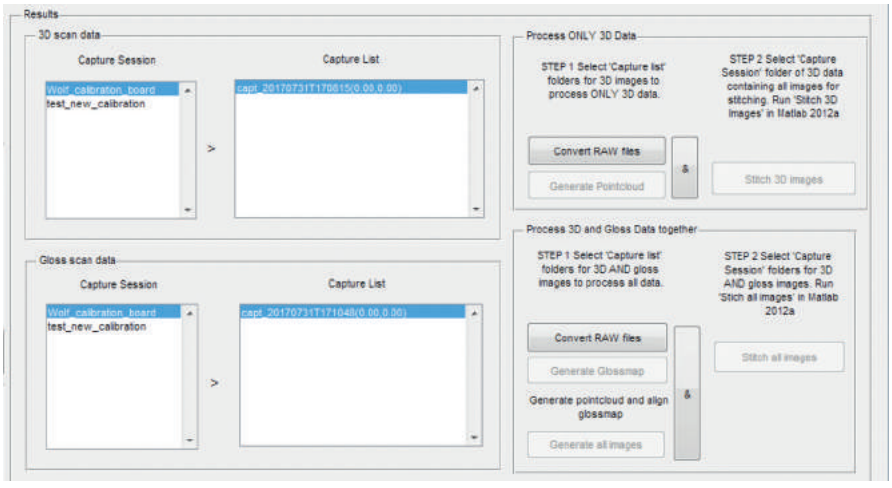



***figure 24.*** *Buttons for image processing.*

### 7.1.1. Select files

First you need to select the folder(s) that have to be processed. In the capture session you can select the session which you want to process and in the capture list the specific rows or columns. With shift and control you can select multiple files at once.

**Note!!** You can select multiple folders within one capture session, but not multiple capture sessions. Take into account that you select the amount of folder for the 3D data as for the gloss data.

### 7.1.2. Convert RAW files

The cameras make .CR2 images during the capturing process. For the image processing .tif images are needed. Thus first these .CR2 files are converted to .tif files. This can be achieved with the button 'Convert raw files'. For this conversion the matlab script 'convert2tif.m' is called for the 3D scan data.

### 7.1.3. Generate pointcloud

The button 'Generate Pointcloud' available in 3D scan data section will generate a color and height map of the images of cam 1 and 2. For each capturing position a new color and height map is generated. This buttons calls the script 'slProcess5'.
The images of cam 1 and 2 are combined with each other to create the height map. The script searches for features on the fringe pattern so it can stitch the image of cam 1 and 2 perfect with each other. For this phase it is important that the camera calibration is executed very well. The height maps contains the wrong dimensions if the distances of the capturing process and calibration process are different.

The first height map that the program creates is used as reference for the rest of the script. This means that when the first pointcloud is created in an angle that the other maps are also placed in an angle. This angle has a reason that features and information is found in the background of the image (the area of the easel and the edges of the painting). The easiest solution is to capture the painting with a black easel and black background in a dark room to avoid any detail in the background of the images. Otherwise you have to make the edges black (the parts that are not the painting) in Adobe Photoshop.

**Note!!** If you scan an object with very features it sometimes happens that the height map is generated with lot of artifacts (holes in the image). For example while scanning a gloss card, you have to add an extra strip with lot of details in color and patterns.

### 7.1.4. & button

If you press the '&' button all the buttons that are placed at the left of the '&' button will be executed. On this way you do not have to press separately on the buttons.

### 7.1.5. Stitch images

The buttons 'Stitch 3D images' and 'Stitch all images' stitches the different maps with each other to have one tiff image of the painting. It executes the scripts 'StitchImages.m' for only the 3D data and 'StitchImages_gloss.m' for the 3D data and gloss data. Both scripts have the following steps. First it sorts the capture folders in rows. These folders are copied to the specific row folder and a list is created of all capture folder names. Secondly the adaption of the XYZ2planes is executed. The script loops over code identical xyz2planes and extract the RGB and Z files. These files are sorted and alphabetically ordered folders. The next step is to merge the stitched images is one tiff with 'fastMerge'. This is executed per row first. Then in the next two steps these rows are stitched together by the script 'stitchIt4.m' and merged again in one image with 'fastMerge'.