

Martynas Dabravalskis

**Near-field Photometric stereo for quality inspection of highly specular
surfaces**

MSci Hons Computer Science

(With Industrial Experience)

17/06/2022

Name: Martynas Dabravalskis

Student ID: 34850546

Project Title: Near-field Photometric stereo for quality inspection of highly specular surfaces

Module: MSci Hons Computer Science (with Industrial Experience) Date: 17/06/2022

I certify that the material contained in this dissertation is my own work and does not contain unreferenced or unacknowledged material. I also warrant that the above statement applies to the implementation of the project and all associated documentation. Regarding the electronically submitted version of this submitted work, I consent to this being stored electronically and copied for assessment purposes, including the Department's use of plagiarism detection systems in order to check the integrity of the assessed work. I agree to my dissertation being placed in the public domain, with my name explicitly included as the author of the work.

The link to the code of this project is: <https://github.com/MartisD/Year4Project>

Date: 17/06/2022

Signed: 

Abstract

Various errors can occur during the embossing, cutting and creasing steps of the folding carton manufacturing process. The expected 5% yearly growth of the folding carton market, combined with the growing customer needs, makes good quality control a necessity. This paper aims to improve the quality control of folding carton manufacturing by exploring photometric stereo techniques of surface reconstruction for easier inspection. Normal maps, obtained from three different photometric stereo systems, as well as four different normal map acquisition techniques, are compared and analysed. Additionally, a method for embossed foil extraction from multiple images of a surface taken under different light conditions is developed and evaluated. Based on the results, recommendations for the future development of such systems are provided.

Contents

1.	Introduction.....	6
1.1	Research goals and contribution	7
2.	Background	8
2.1	Project context	8
2.2	Folding carton boxes	8
2.2.1	Manufacturing process	8
2.3	Digital Measurement.....	10
2.4	Photometric stereo	10
2.4.1	Normal map acquisition	11
2.4.2	Diffuse-specular separation.....	12
3	Design.....	13
3.1	Methodology	13
3.2	Hardware.....	13
3.3	Software	13
4	Implementation	14
4.1	Hardware.....	14
4.2	Software	16
4.2.1	Normal map generation.....	16
4.2.2	Foil detection and extraction.....	16
5.	Testing and Evaluation.....	22
5.1	Test setup	22
5.2	Data set.....	22
5.3	Results.....	22
5.3.1	Time of execution	23
5.3.2	Normal map quality comparison.....	23
5.3.3	Foil extraction	26
5.4	Evaluation	27
5.4.1	Normal maps	27
5.4.2	Foil Extraction	28
5.5	Recommendations.....	29
6.	Conclusions.....	30
6.1	Review of Aims	30
6.2	Future work.....	30
6.3	Challenges.....	31

6.4	Skills learned.....	31
6.5	Final remarks	32
References.....		33

1. Introduction

The total value of the folding carton market is expected to reach over \$160 billion by 2026 (Harrod, 2021). This creates a need for fast and effective quality control, as it is essential for reducing costs, and liability risks, maintaining or improving market position and increasing customer loyalty (Plex Team, 2021).

One of the most error-prone steps of folding carton manufacturing is embossing, cutting and creasing. Incorrectly creased cartons may have to be discarded as the box could become impossible to fold with a folding machine, and incorrectly embossed surfaces can lead to potential delamination of the surface. These errors are hard and labour-intensive to inspect and detect, as they can be very small and barely visible to the naked eye. As such, a method for inspecting such surfaces could prove very beneficial.

Photometric stereo techniques could be used for an easier inspection of these elevated surfaces. Photometric stereo uses multiple images taken from the same viewpoint under different lighting directions to recreate the surface of an object. A normal map can then be used to inspect these surfaces. A depth map can also be created for a more detailed inspection.

The basic Lambertian method of computing photometric is based on least-square regression (L2). It functions under the assumption that surfaces are Lambertian (Ideal diffusive reflection) and have a uniform albedo. While this assumption may be valid for some surfaces, most objects are not ideally Lambertian. As such, the accuracy of the basic photometric stereo algorithm decreases as the surfaces become more reflective (exhibit more Specular reflection) or contain bright specular spots on an image (Woodham, 1992).

This is especially relevant in the packaging industry. While the inner side of packaging usually exhibits diffuse reflection, as it is usually composed of plain cardboard, the painted side of the folding carton can exhibit significant specular reflectance. While fold creases can be inspected by performing photometric stereo on the inside surface, the embossed elements of packaging are usually on the outside surface, which is painted with paint that is not ideally matte. Additionally, Glossy foil finishes are becoming increasingly more popular (POST PRESS MAGAZINE, 2020). These finishes are glossy and non-Lambertian. Furthermore, most packaging surfaces are mixed, with some parts Lambertian and some not, as well as having different colours, leading to the surfaces not having uniform albedo.

As the original assumptions are unrealistic for most surfaces and cause errors in a surface normal recovery, several different normal map acquisition methods have been developed that try to mitigate errors caused by specularities. Methods, such as Robust Principal Component Analysis (RPCA) or Sparse Bayesian learning (SBL) can be used to recover the normal maps of these surfaces more accurately.

Another important factor is the quality of photometric stereo images, which depends on the device used to capture these images. For example, ring light lighting generally produces better

results as more images can be taken from different fixed light angles (Fan, et al., 2022). However, this should lead to longer computational times, which is important for a quick inspection of surfaces. Alternatively, as few as 3 or 4 lighting sources could be used to perform photometric stereo. The most common method is providing lighting from 4 different sides at a 90° angle, as these systems are easy to design.

The imaging device can also make a difference. A regular camera might experience lens distortion which can impact the quality of the recovered surface and make precise measurements difficult. A scanning system that had been worked upon during my placement might prove to be more suited to the packaging industry. A scanned image may experience less distortion and could be better integrated into an assembly line of a manufacturing facility.

This paper implements and conducts a comparison between different photometric stereo methods mentioned above on different folding carton surfaces. Three different photometric-stereo-capable devices will also be compared (A 4-light camera system, a ring light camera system, and a 4-light scanner system). Conclusions will be drawn from the results and recommendations for future development will be provided. Additionally, a novel method of extracting embossed foil is developed and evaluated.

1.1 Research goals and contribution

- Review of related work
- Design of a scanner-based system
- Design of a ring-light-based system
- Comparison of normal map quality between 3 systems.
- Analysis of the systems and recommendations/methods for improvement.
- A method for finding specular/reflective areas of images using multiple lighting angles and a fixed camera position.
- A proprietary data set of images that can be used to develop algorithms.

2. Background

2.1 Project context

I have been working with Hardy & Ellis Inventions LTD for 2 months to assess the feasibility and prototype a scalable, fast, low-cost flatbed scanner that can be combined with multiple light sources to achieve a photometric stereo effect. The main market of the company is developing, producing and maintaining Quality Control machines in the packaging industry. The insights gained from my work have been useful for the company and contributed to a future product line.

This project continues the work conducted during my placement by evaluating the effectiveness of the prototype scanner device compared to other photometric stereo-capable systems as well as further exploring the use of photometric stereo for quality control.

2.2 Folding carton boxes

Cartons are small to medium-sized boxes made from cardboard. They comprise a significant proportion of the packaging found in the retail sector. Folding cartons are a cost-effective way of packaging. Their manufacturing process makes it easy to print on the surface of the packaging, which provides information and visual impact. Additionally, it can make the product more attractive by allowing for simple embossing of foil or other eye-catching material on the surface. Unlike rigid cartons, folding cartons are manufactured in sheets and efficiently delivered to the packer, where they are erected, filled and closed (Kirwan, 2008).

Due to the ease of printing, applying varnish, printing, and foil embossing on the sheets, they are used to package a variety of products, such as food, pharmaceuticals, perfumes, cosmetics, toiletries, toys, games, etc.

2.2.1 Manufacturing process

The manufacturing process can be divided into 5 main steps:

- **Printing** – Offset lithography, flexography and gravure are used to print text, illustrations and descriptive diagrams on the box.
- **Flatbed dies** – Pre-set cutting knives and creasing rules are used to accurately cut and create creases in carton sheets. The creases are used for easier folding of the carton.
- **Hot Foil stamping and embossing** - Surface printing and decoration. Uses a heated die, containing the design, from a special film. The colour may be either a pigment or plain(silver) or coloured aluminium foil.
- **Glueing** - Sealing carton side seams.
- **Folding** – Folding the carton sheets into boxes.

This project seeks to improve the quality control operations in the highlighted sections above, which are cutting and creating creases, hot foil stamping and embossing. As the paper sheets are fed to the dying and embossing machines via a conveyor belt, a photometric device, such as a scanning camera, could be placed at the end of these machines for quality inspection.

2.2.1.1 Flatbed dies

The cutting and creasing of a folding carton are performed by cutting and creasing rules, which are pre-set on a special frame (See Figure 1). Throughout the creasing process, the sheet experiences large tensile and compressive stresses, which can create tears and may lead to delamination of the carton. Additionally, both the cuts and the creases can be inaccurate due to either the incorrect placement of the sheet or errors induced by stress. Small tears and creases can be hard to detect, so surface texture reconstruction via photometric stereo could allow quicker detection of these errors.

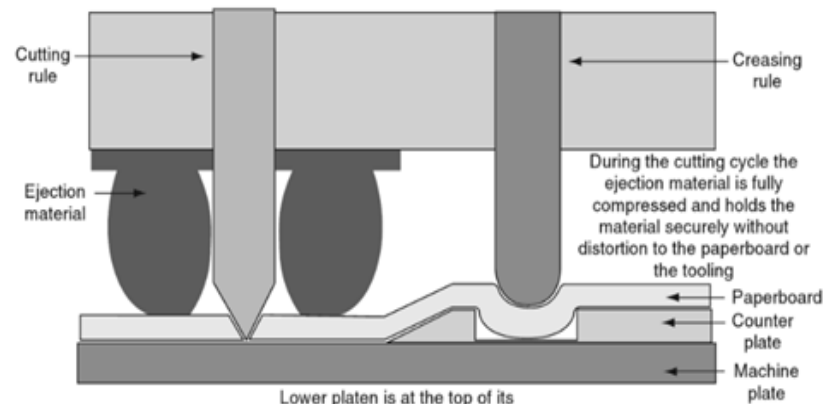


Figure 1 Cutting and Creasing (Kirwan, 2008)

2.2.1.2 Hot foil stamping and embossing

This process uses a heated die which embosses pre-made foil onto a surface (See Figure 2). This process can be prone to errors. The foil is usually made out of 5 layers. These layers may come misaligned from the factory, or the layers can move and separate from the heat and pressure of the die. Imprecise cuts can also cause misalignment, and the heated die may also damage or warp the previously applied paint on the paperboard. An example of misaligned foil placement caused by an error can be seen in Figure 3.

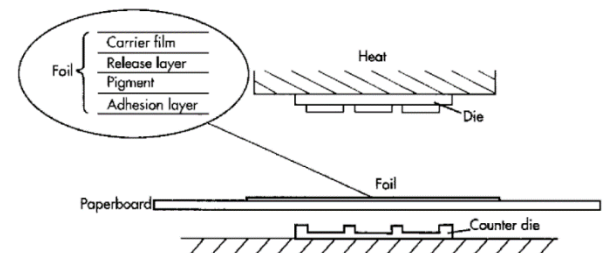


Figure 2 Embossing via a heated die (Kirwan, 2008)

The misalignment can be hard to see and inspect visually, especially if the misalignment is small. However, as the layers are stacked, such misalignment could be detected by looking at the difference in depth on a reconstructed surface.



Figure 3 Misaligned embossed foil

2.3 Digital Measurement

There are multiple ways to digitally measure the surface of an object and acquire its shape. These techniques can be divided into Active and Passive shape acquisition techniques (See Figure 4).

The Active methods focus on direct physical contact with the surface, such as using a probe, or an interaction of a surface with an energy signal, such as using time-of-flight sensors to measure the time it takes for a wave to bounce off the surface.

Passive methods acquire shape from their optical appearance under non-tightly focused illumination.

Usually, these rely on some form of stereo vision, where the object is observed from different angles.

Binocular and Photometric stereo are two basic stereo vision techniques. Stereoscopic imaging (Binocular vision) uses multiple images taken from different viewpoints under the same lighting conditions to recover depth maps (Yamashita, et al., 2004). While this method is simple, it usually produces poor-quality results. It also requires multiple cameras, which can be expensive.

Photometric Stereo uses multiple images (3 or more) taken from the same viewpoint under different lighting conditions to recover depth maps. This method can be slower as more images need to be taken, but it is capable of producing highly detailed models even in smooth untextured regions of the surface (Hao, et al., 2011).

Hybrid methods can be used, which combine multiple techniques, such as Binocular Photometric Stereo (Hao, et al., 2011) or Active Stereo Vision Using Structured Light (Jung, et al., 2017) to recover texture information.

Continuing the work done during my placement, this project continues to focus on the photometric stereo techniques.

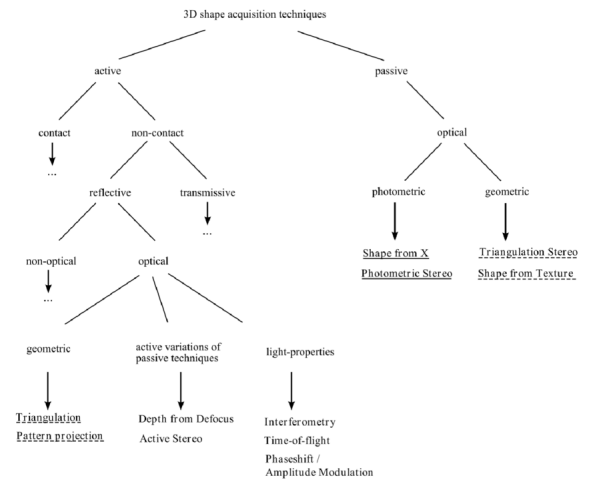


Figure 4 3D shape acquisition techniques (Herbort & Wöhler, 2011)

2.4 Photometric stereo

Photometric stereo is a technique used to estimate the surface normals of an object from a set of images taken from the same viewpoint under different light directions (See Figure 5). A surface normal is a vector that is perpendicular to a given surface at a given point. It is used to recover a surface in photometric stereo by determining the angle between the surface and the incident light.

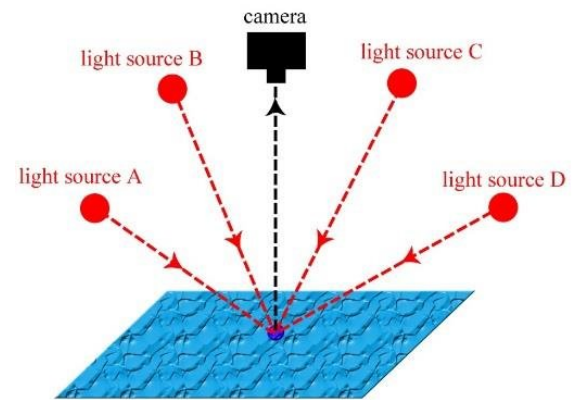


Figure 5 Photometric stereo (Smith, et al., 2018)

One of the challenges with photometric stereo is the difficulty of obtaining the accurate position of the light source. While some methods, such as Singular Value Decomposition, can be used to estimate the light positions, they only give approximations and are not entirely accurate. Additionally, the shape and surface texture of the object can also affect the accuracy of the reconstruction. For example, photometric stereo is typically limited to objects with smooth surfaces and performs poorly with surfaces that contain a lot of specularities.

The number of images taken also impacts the quality of reconstructions. Some more advanced photometric stereo methods only work with a large number of images and perform very poorly if it is presented with a small dataset. However, using many images can be very computationally expensive.

There is no clear answer on how to properly calibrate a photometric stereo system as calibration can vary based on the software and hardware setup used for photometric stereo. However, a typical approach to calibration is to use an object with known dimensions and perform multiple observations under light sources hitting the object from different angles, using software to process the image and calculate the surface normal.

2.4.1 Normal map acquisition

A surface normal is a vector that is perpendicular to a given surface at a given point (See Figure 6). It is used to recover a surface in photometric stereo by allowing for the determination of the angle between the surface and the incident light.

There are several methods of extracting normal from a set of images.

The most basic and fastest method is Least-square regression (L2). This method assumes Lambertian reflectance and regards specular high-lights and cast shadows as outliers.

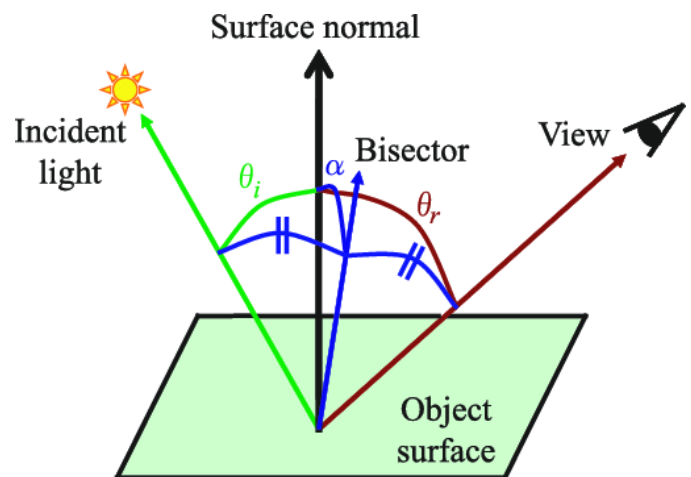


Figure 6 Surface normal (Miyazaki, et al., 2003)

L1 residual minimization (L1) uses L1 regularization to calculate and obtain the normal map. However, L1 regression is computationally expensive.

The more advanced method includes Sparse Bayesian learning (SBL) and Robust Principal Component Analysis (RPCA).

SBL decomposes observations into a linear diffuse component and a sparse, non-diffuse component for capturing shadows, specularities and other corruptions. One drawback of this method is that it is computationally expensive.

RPCA uses convex optimization techniques to find the correct low-rank matrix (recover surface normal) by simultaneously fixing its missing and erroneous entries. Unlike most other methods, this method is much faster (Ikehata, et al., 2012).

2.4.2 Diffuse-specular separation

Diffuse and specular surfaces of a carton can be extracted via diffuse-specular separation. There are a number of methods that achieve this. One simple method is using polarized light. Polarizer filters can be added so that only mostly diffuse reflections are captured by the camera. Then, the specular reflections can be extracted by subtracting the diffuse data from observations made without polarization, which include both diffuse and specular reflections (Müller, 1995). However, this method is not perfect, as polarization does not completely remove the specularities.

Advanced illumination, such as gradient spherical illumination can be used to perform diffuse specular separation as well (Christos, et al., 2018). While these methods can produce excellent results, they usually require heavy calibration and precise lighting.

My method of separating the specular parts of the image works under the assumption that an overhead light hitting an edge of an embossed specular surface is visible as a bright reflection on the camera. These bright spots can then be captured to recover the highly specular areas.

3 Design

3.1 Methodology

An iterative prototyping approach was used to design and construct two photometric stereo-capable prototype systems.

This approach was also used for software design, as the software needed to be adjusted iteratively depending on the results.

3.2 Hardware

Two different photometric stereo capable systems are designed.

The first one is a scanner-based system. Flatbed scanners use line imaging sensors, which scan objects line-by-line and can produce very detailed high-quality digital images at a low cost. Line sensing technology could be used in the quality control part of packaging manufacturing, where a Line scan camera could scan the packages while they are moving through the assembly line. As such, this system can be scalable and fast. The system contains 4 directional light sources (LED strips) placed at a 90° angle to perform photometric stereo. A scanning camera may also be used to accurately measure the dimensions of objects across the picture, as they experience less lens distortion than regular cameras.

The second system uses an industrial camera with a ring light, mounted on a fish-eye lens. The ring light consists of 24 individually addressable LEDs. A ring light may provide better quality normal maps and easier foil extraction, as it allows for a large number of images to be taken under different lighting conditions.

A third and already implemented design will also be used. This design uses an industrial camera with 4 directional light sources placed at a 90° angle. This design uses micro louvre film to control the distribution of light to optimize the lighting for photometric stereo.

3.3 Software

A fast and accurate method for detecting and extracting embossed foil from multiple surface images needs to be designed. The method should work with at least one of the photometric stereo-capable systems and produce results with minimal noise. These extracted images of foil can then be overlayed on a normal map for inspection of foil placement.

Software needs to be written to produce normal maps via the 4 different methods described in the background section.

An Arduino software should be written to control the prototype devices.

Python is used for computer vision, as it provides accessible and easy-to-use tools for data/image manipulation. The OpenCV library is extensively used for image manipulation.

4 Implementation

4.1 Hardware

The ring light device was constructed by mounting a commercial 24 LED ring light onto a fish-eye lens of the camera. The camera and the lens were put into a 20x20x20cm box to prevent interference from outside light. A simple Arduino program was written to control the lighting system via serial for the simple acquisition of images under different lighting conditions



Figure 7 Ring light

A commercial CCD scanner was used for the scanner-based system. A few modifications to the scanner needed to be made to make it capable of taking photometric stereo images. Firstly, the scanning window, on which an object is usually placed, was removed to prevent lensing. Secondly, the backlight of the scanner needed to be turned off. This could not be done in software, as most processing is done inside the scanner and is not controllable via computer. The backlight also needs to be lit up for the scanner to start scanning, as most scanners have a mandatory calibration sequence, which uses light, reflected from a white strip lit by a backlight to calibrate the scanner. Without this calibration sequence, the scanner will refuse to scan. As such, a switch was installed on the power line going to the backlight of the scanner to manually turn off the backlight when the printer starts scanning (See Figures 8 and 9).

A square frame containing 4 LED strips on each side was constructed out of foam board (See Figure 10). The LED strips contain a diffusion filter to more uniformly distribute the outgoing light. The scanned object is then placed below the square frame containing the light strips, and the scanner is placed 1cm above the surface upside down. Due to the limited focal depth of the

scanner, the scanned surface must be as close to the scanning head as possible. Testing was done to ensure that a 1cm distance small enough to produce focused scans.

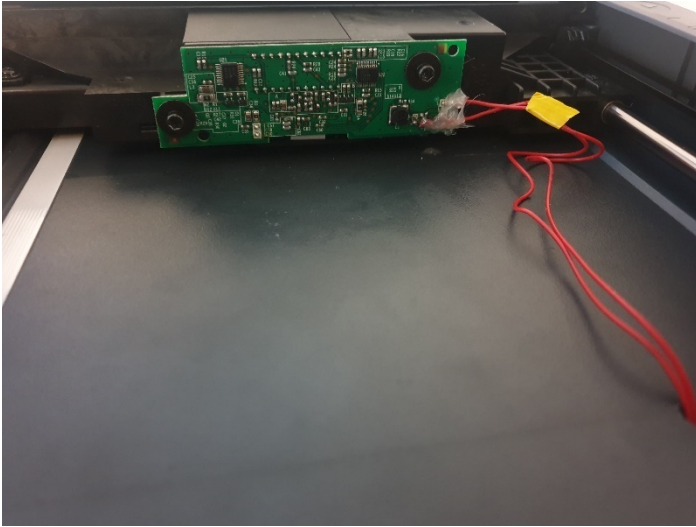


Figure 8 Scanner backlight modification

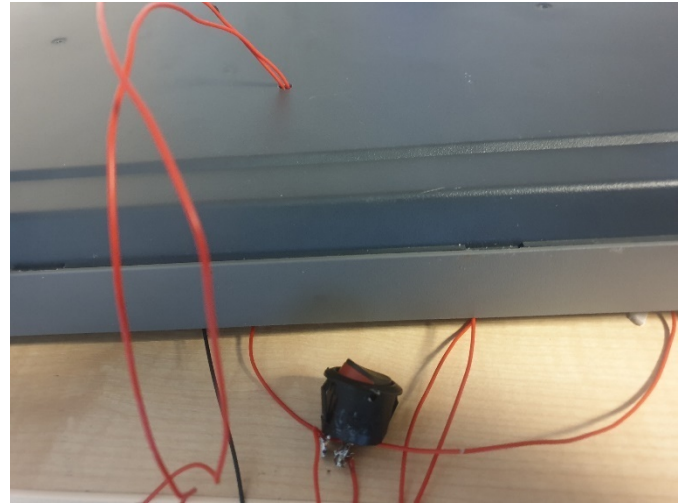


Figure 9 Modified scanner



Figure 10 LED light strip frame

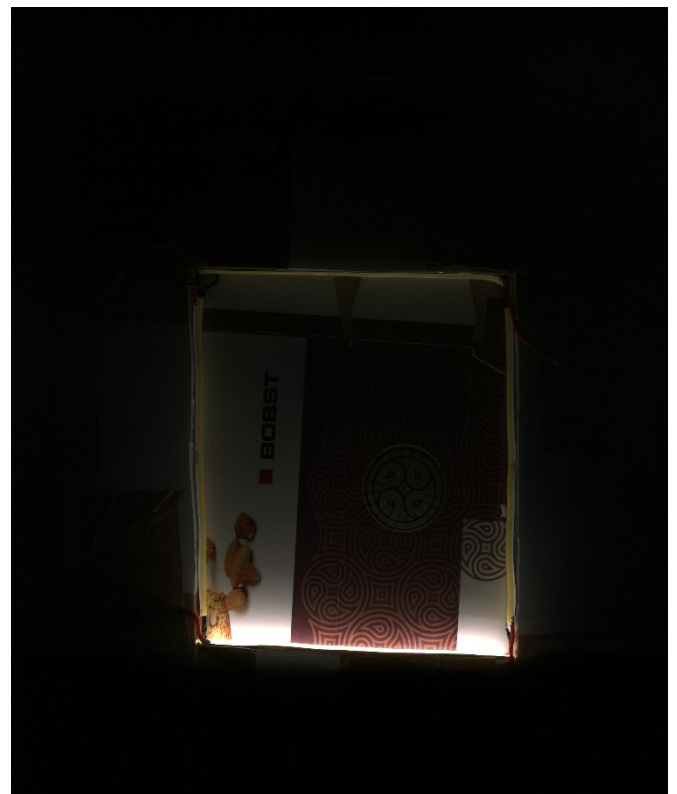


Figure 11 An image with the backlight turned off

4.2 Software

4.2.1 Normal map generation

Python software was implemented to obtain the normal maps from the images.

Publicly available Python implementations of Least-square, L1 minimization, Sparse Bayesian learning, and Robust Principal component analysis normal maps were used (Ikehata, et al., 2012)

First, the edges of the images are centre cropped to increase computation speed. All images are cropped to 1500x1500 pixels from their original size of 5496x3672 for the industrial camera, and 800x800 pixels from 2040x3672 for the scanned images. The scanned images are cropped to a smaller size, as the prototype is limited to a maximum area of the scanned surface of 800x800 pixels at 200dpi.

The developed software loads and pre-processes the images via converting them to grey-scale, normalizing, and passing them through a high-pass filter. An image matrix is created from the images, which is then passed to each of the methods mentioned above. The retrieved normals are then post-processed and converted to images. Sobel edge detection is used to produce another set of images for easier inspection of quality.

Instead of manually measuring the light positions, Singular value decomposition (SVD) was used to estimate light directions for each image.

4.2.2 Foil detection and extraction

The method for extracting reflective foil from a surface assumes that an overhead light hitting an edge of an embossed specular surface is visible as a bright reflection on the camera (See Figure 12). These bright spots can be extracted from each image and combined to recreate the shape of the foil.

However, bright spots also form on the painted background surface. This is because light from a light source comes in a beam, the intensity of which is highest in the centre. As such, the parts of the background that are directly below the ring light LEDs will show as bright spots. This is also made worse by the fact that the paint finish of the surface is not fully diffuse and reflects some of the light specularly. As such, parts of the image that are close to the overhead light source need to be removed, as they contain a large amount of noise. Bright (White) diffuse background also

needs to be removed.

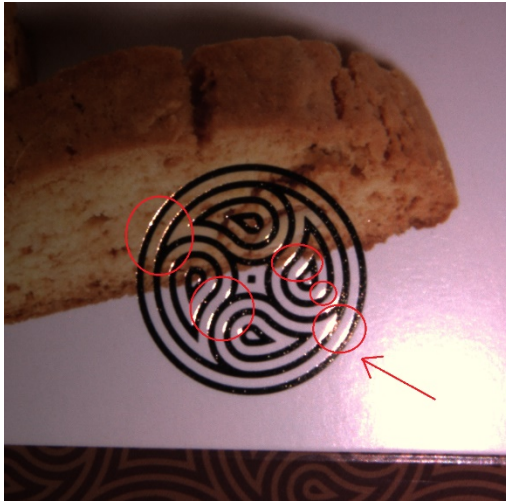


Figure 12 Bright spots formed on the foil

The developed method can be divided into 3 parts:

- Extracting highlights
- Removing noise
- Subtracting background and combining captured foil from each image

4.2.2.1 Highlight extraction

First, the image is converted into grayscale. Then, it is normalized (0-255) to account for different light source intensities. Binary thresholding is used to extract the brightest spots of the image. 99% of the image is removed, but the number can be adjusted as needed. Setting the percentile lower will lead to more features being extracted but will also lead to more noise in the picture. Code snippet:

```
# Turn to grayscale
image= cv2.cvtColor(image, cv2.COLOR_RGB2GRAY)
# Normalization
image = normalize(image,0,255)
image = image.astype(np.uint8)
image = np.array(image)
image[image< np.percentile(images, 99)] = 0
image[image!=0] = 255
```

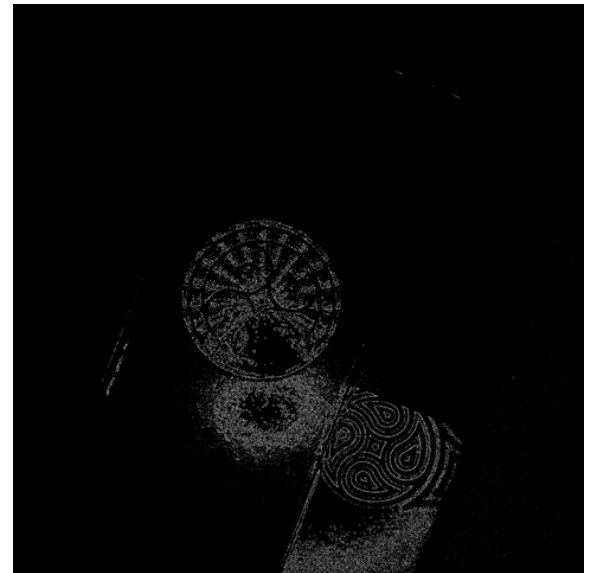


Figure 13 Image with extracted highlights

4.2.2.2 Removing noise

Surface highlights created from a perpendicular beam of light coming from a sensor usually form circular ‘blobs’ that can be detected via blob detection. 95% binary thresholding is used to extract the 5% of brightest spots. A combination of a Gaussian high-pass, erosion, dilation filters and 90% inverted Binary thresholding is then used to extract these noisy highlights (See Figure 14). The coordinates of these noisy highlights are then detected via OpenCVs built-in blob detector. Circularity filtering is used for more accurate detection.

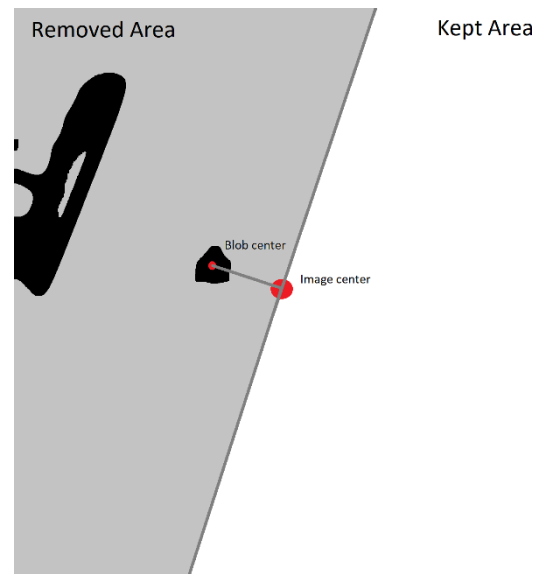


Figure 14 Calculating the area to remove

```
def getLights(images):
    lights = []
    blobimgs = []
    for image in images:
        image[image < np.percentile(image, 95)] = 0
        image[image != 0] = 255
        light = np.copy(image)
        light = highpass(light, 15)

        # erosion followed by dilation
        kernel = np.ones((15,15),np.uint8)
        light = cv2.morphologyEx(light, cv2.MORPH_OPEN, kernel)
        retval, threshold = cv2.threshold(light, np.percentile(light, 90), 255,
                                          cv2.THRESH_BINARY_INV)
        blobimgs.append(threshold)

    for image in blobimgs:
        tempImage = np.copy(image)
        for im1 in blobimgs:
            if not np.array_equal(im1, image):
                tempImage[tempImage == im1] = 255

        tempImage = cv2.morphologyEx(tempImage, cv2.MORPH_OPEN, kernel)
        # Blob detection (for detecting specular highlights from light)
        detector = cv2.SimpleBlobDetector()
        params = cv2.SimpleBlobDetector_Params()
        params.filterByCircularity = True
        params.minCircularity = 0.01
        # Create a detector with the parameters
        detector = cv2.SimpleBlobDetector_create(params)
        lights.append(detector.detect(tempImage))
    return lights
```

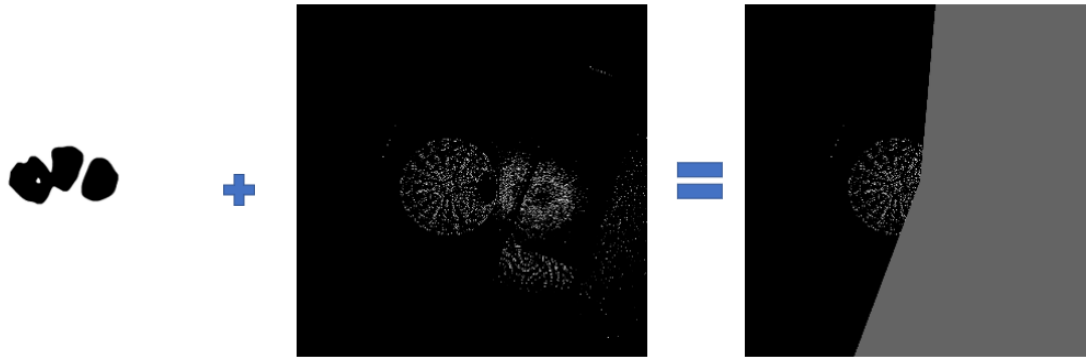


Figure 15 Noise removal

These coordinates are then used to draw a line between the centre of the image and the centre of the blob. A line perpendicular line going through the centre point to the edges of the image is then drawn. The start and end coordinates of the line are used to create and draw a rectangle that covers the area of the image separated by the line. The black rectangle is drawn onto the image obtained in the previous step, as such removing the noise created by the light source.

The noise removal function can be seen below:

```
def removeBlobs(image, lightpoints):
    # Get center point
    cx = int(image.shape[0]/2)
    cy = int(image.shape[1]/2)

    # Add filled circles on specular highlights so that we can ignore those spots
    for lp in lightpoints:
        # Get coordinates of highlight center
        x,y = lp.pt
        x, y = int(x), int(y)

        # Length of line to be drawn across the screen, the length has to be long enough for the line to cut the image
        length = image.shape[0] + image.shape[1]

        # Calculate the coordinates of line going through the center point that is perpendicular to the line going from
        # the center point to the highlight point
        lx = cx - x
        ly = cy - y

        mag = np.sqrt(lx*lx + ly*ly)
        lx = lx/mag
        ly = ly/mag
        temp = lx
        lx = -ly
        ly = temp

        r1x = int(cx + lx * length)
        r1y = int(cy + ly * length)
        r2x = int(cx + lx * -length)
        r2y = int(cy + ly * -length)

        # Create coordinates for a rectangle that will be drawn (To fill the area cut by the cut line)
```

```

r1xo, r1yo, r2xo, r2yo = r1x, r1y, r2x, r2y
if((r1y+r2y)/2 < y):
    r1y += 2*cy
    r2y += 2*cy
else:
    r1y -= 2*cy
    r2y -= 2*cy
if((r1x+r2x)/2 < x):
    r1x += 2*cx
    r2x += 2*cx
else:
    r1x -= 2*cx
    r2x -= 2*cx

# Turn points into contours, draw the contours
points = [[r1xo,r1yo],[r2xo,r2yo], [r2x,r2y], [r1x,r1y]]
points = np.array(points).reshape((-1,1,2)).astype(np.int32)
cv2.drawContours(image,[points],0,0,-1)

```

4.2.2.3 *Subtracting background and combining captured foil from each image*

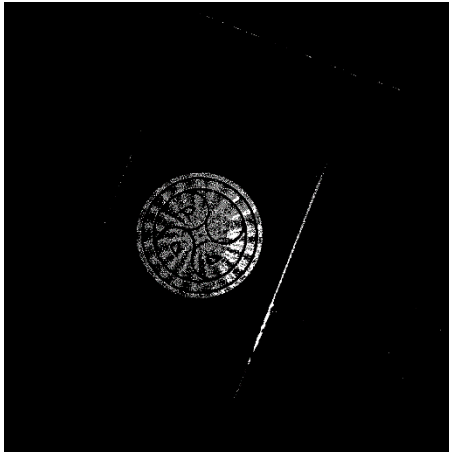


Figure 16 Extracted foil

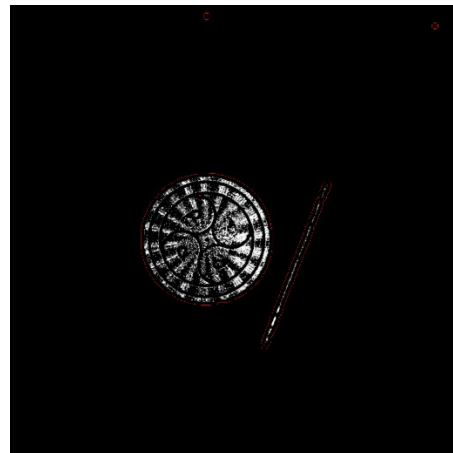


Figure 17 Extracted foil with generated outlines

The remaining background is subtracted from each image obtained in step 2 by removing the parts of the image which are identical across all images obtained in step 1. The images are then combined to form a complete image of the foil (See Figure 16):

```

finalImg = np.zeros(images[0].shape)
bg = np.zeros(images[0].shape)
bg[images[0] == images[1]] = 255
# Get difference between images
for image, lightpoint in zip(images, lightpoints):
    bg[image != bg] = 0
    removeBlobs(image, lightpoint)
    finalImg = finalImg + image

finalimg[finalImg == bg] = 0

```

Contours can be drawn over the extracted areas by passing the final image through a Gaussian low-pass, dilation filters and thresholding(See Figure 17).

5. Testing and Evaluation

5.1 Test setup

A data set of images of folded carton sheets containing foil was obtained via 3 different photometric stereo systems.

Normal maps obtained from 3 different photometric stereo systems are obtained and compared. Only The maps obtained via the Least Squares normal map acquisition method were compared, due to the much higher computational complexity of the other methods (L1, SBL, RPCA).

Normal maps are compared visually. Images obtained via Sobel Edge detection are also compared for easier visual comparison.

Normal maps obtained via 4 different normal map acquisition methods are compared. This comparison uses Images obtained via the 4-light camera setup.

To evaluate the impact of the number of input images on the algorithms, 2 of the fastest normal map acquisition methods are compared (Least-squares and RPCA). This comparison uses 24 images taken via the ring-light setup.

The foil extracted via the developed foil extraction method is tested for computational speed and accuracy by comparing the results generated from 3 different folding carton image datasets obtained from the ring light setup.

5.2 Data set

A large data set of a total of 312 images was collected.

A total of 4 datasets of 5 images each were collected via the 4-light camera setup. All 4 datasets of images are of different types of packaging and contain 4 directional light photos and 1 image with general lighting.

A total of 12 datasets of 24 images each were collected via the ring-light setup. All contain 24 directional light photos. Some of the datasets contain an additional image with general lighting. The images collected were of 4 distinct folding cartons.

A total of 2 datasets of 5 images each were collected via the scanner setup. All 2 datasets of images are of different types of packaging and contain 4 directional light photos and 1 image with general lighting.

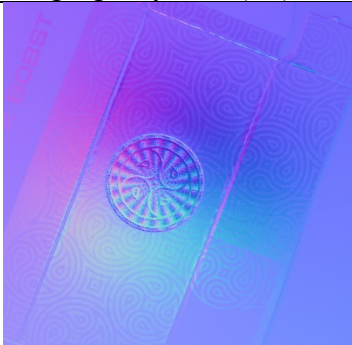
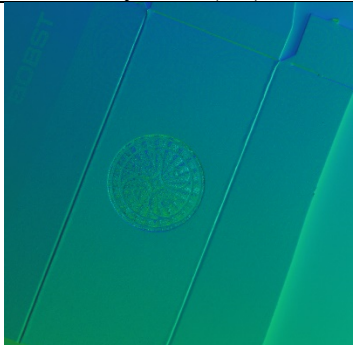
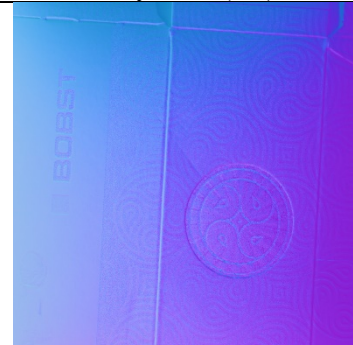


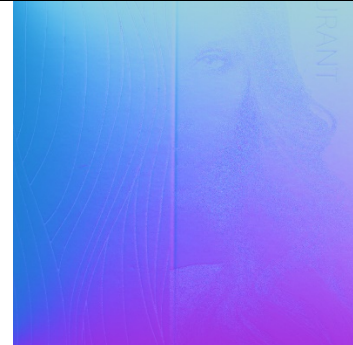


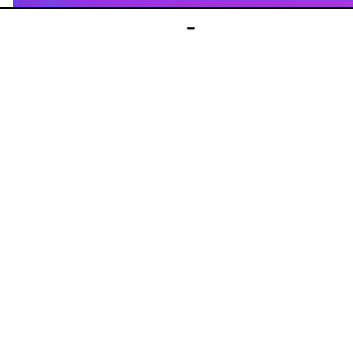
Additionally, each dataset contains an animated GIF, created by combining the photometric stereo images.

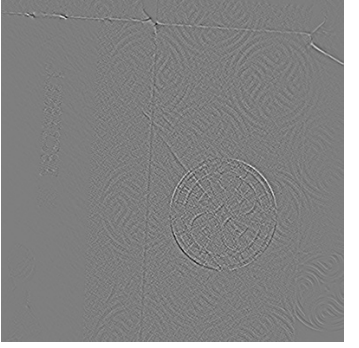
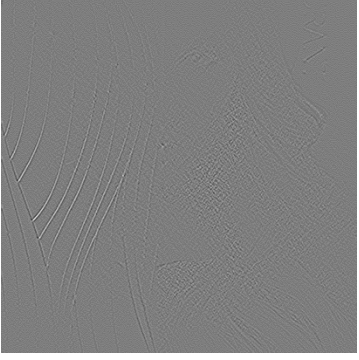
5.3 Results

5.3.1 Time of execution

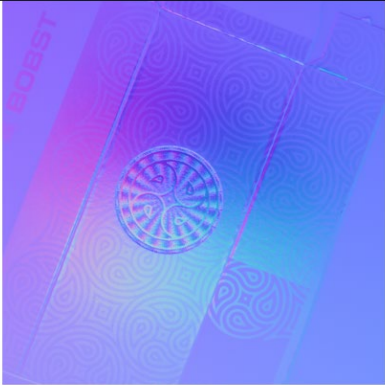

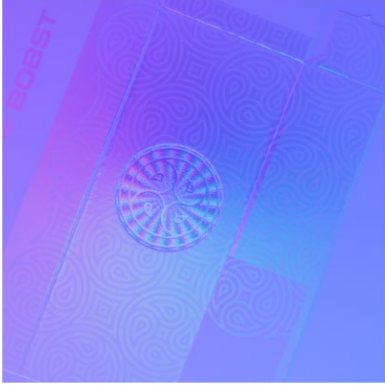

Method	Execution time (s)
L1 (multithread)	1840
SBL (multithread)	over 3600, abandoned
L2	0.5 (4 images), 4 (24 images)
RPCA	25 (4images) 280 (24 images)

5.3.2 Normal map quality comparison

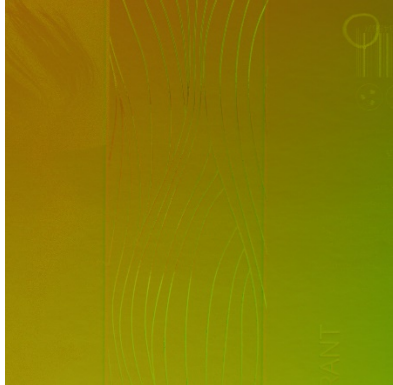
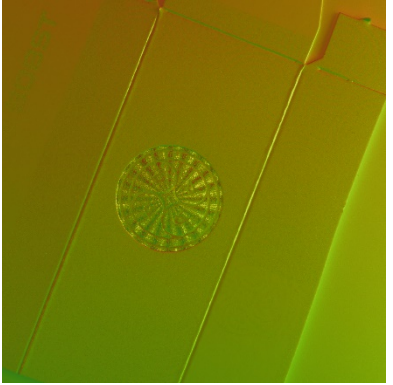


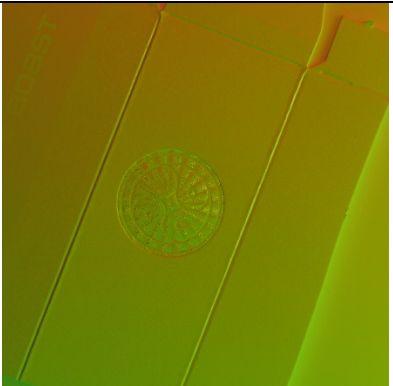

Ring light system (L2)	Camera system (L2)	Scanner system (L2)
		
		
		

Sobel edge detection performed on normal maps		
Ring light system	Camera system	Scanner system
		
		
		

Normal map acquisition method comparison (ring light system)

L2		
RPCA		


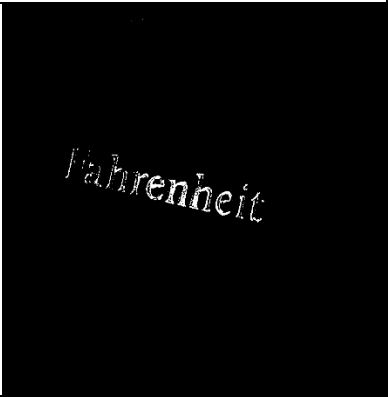



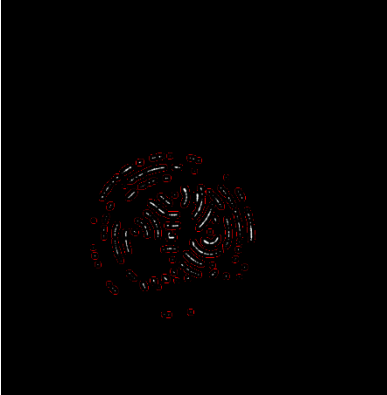

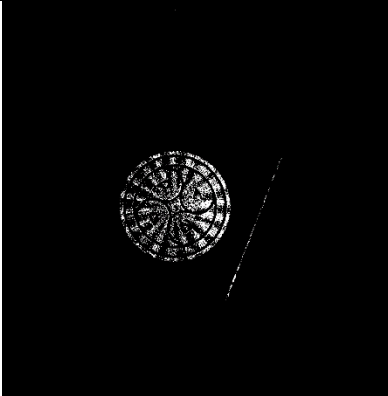

Normal map acquisition method comparison (4-light camera system)

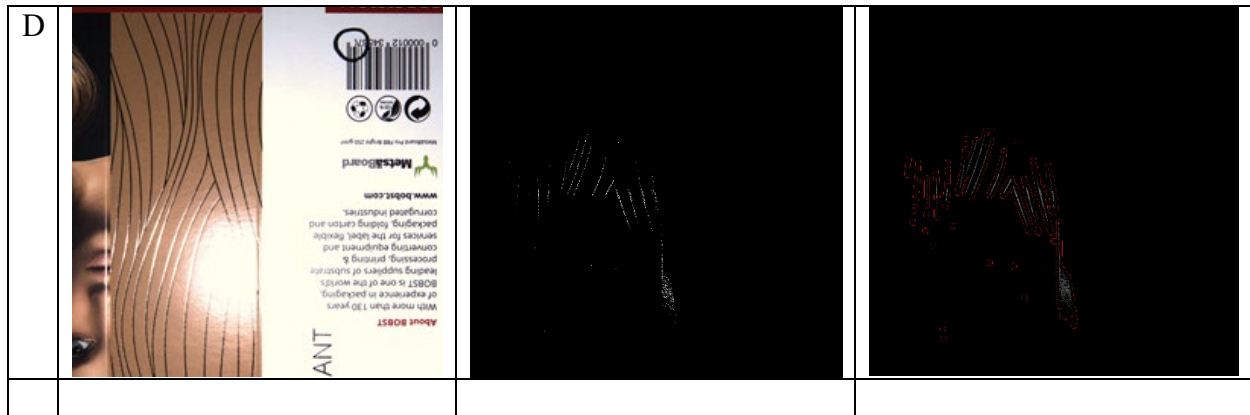
L1			
L2			

RPCA	-	-	-
------	---	---	---

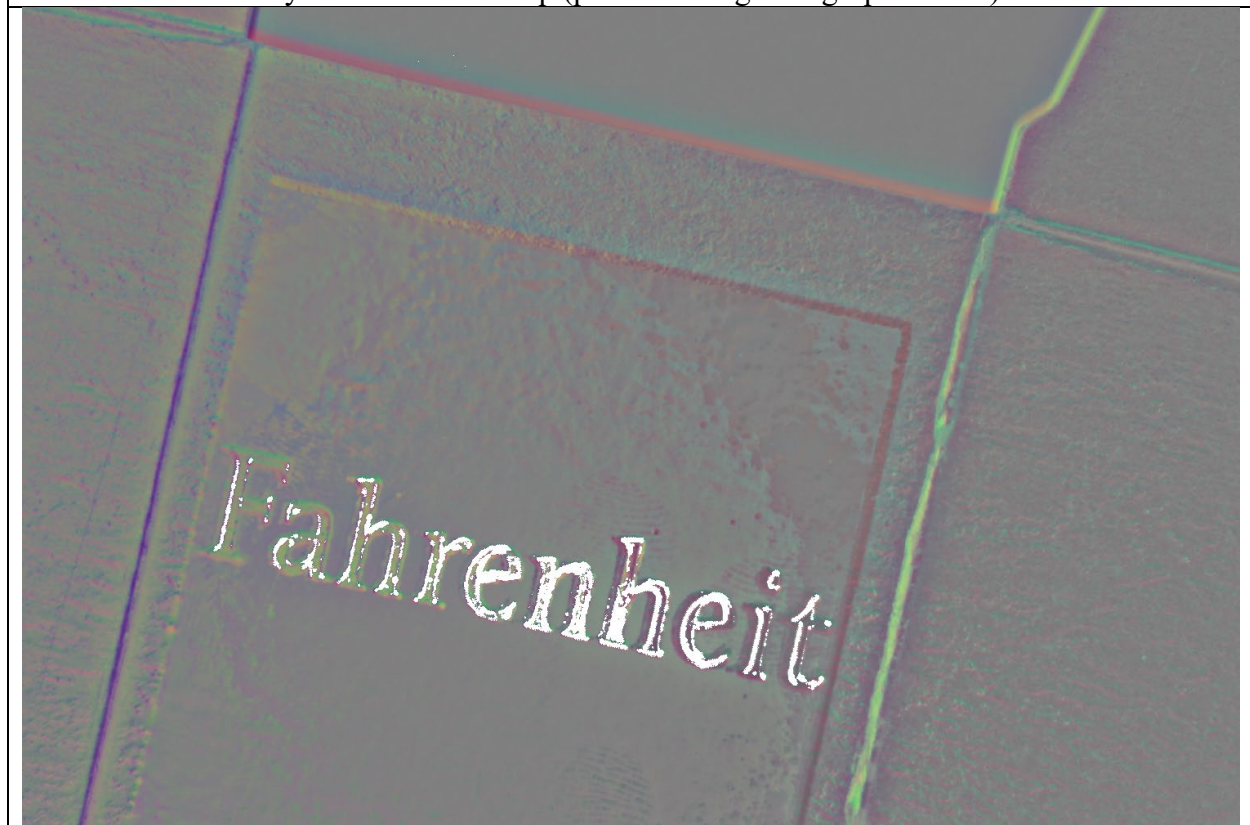
5.3.3 Foil extraction

Memory usage	2x total image size
Time of execution	5s

	Original images	Extracted foil	Extracted Foil with Contours
A			
B			
C			



Extracted foil overlaid on a normal map (passed through a high-pass filter)



5.4 Evaluation

5.4.1 Normal maps

The comparison of L2 normal maps generated from images from 3 different systems shows that the 4 light camera system produced the best results across all samples. The background noise is minimal and the embossed areas and fold grooves are clearly visible. The normal maps generated from the images from the ring light system also clearly show the elevated areas. However, the normal map also contains a large number of errors, as it contains the printed textures on the

carton, which are not elevated or embossed. The scanner system produces the worst results, as the normal maps were undetailed and also contained a lot of errors.

The difference in normal map quality of the 3 systems can mostly be explained by the lighting setup. The 4-light camera system contains micro louvre film which is used to control the distribution of light. The scanner system has no such film. Additionally, the led strips in the scanner system were not perfectly aligned and straight, which further decreased the system's performance.

There is a clear difference in normal map quality between L2 and RPCA, when tested with a ring light system. The Robust Principal Component Analysis clearly reduced the level of noise in the normal map, which is seen when looking at the specular highlights produced from the ring light beam on the L2 example. However, RPCA could not produce a normal map when using the 4-light camera system, which indicates that it only works well with a large number of images.

There is only a small difference in the quality of normal maps acquired by L1 and L2 methods. The L1 normal map contained more detail (foil area) and had less noise. However, L1 was 140 times slower than L2, which is a large amount of time for such a small quality improvement.

The time of execution table clearly shows that only L2 and RPCA could reasonably be used for fast quality inspection. The L1 and SBL methods were incredibly slow, even when using multithreading. SBL tests were abandoned as the method could not produce a normal map from 4 1500x1500px images in under an hour.

Overall, it is clear that L2 is still the best choice for fast normal map generation, as the quality of the normal maps was more dependent on the lighting conditions than the used generation method.

5.4.2 Foil Extraction

The foil extraction method can clearly separate and extract embossed foil from the test images. However, it can be seen that in some cases (Especially B and D), only a small amount of foil is extracted. The D sample performed by far the worst in this case. This is due to the fact that only a small part of the foil surface was inside the boundaries of the ring light. The image extracted from sample C contained more noise than the other samples, but the extracted foil was much more detailed. As such, we can clearly see that there is an inverse relationship between the amount of extracted foil and the noise. This is dependent on the lighting conditions as well as the texture. As such, while this method does seem to work quite well for each test sample, the parameters have to be adjusted for each texture individually to achieve the best results.

The example of extracted foil overlayed on a normal map clearly displays the foil misalignment. This indicates that the method can be used for inspection of such errors.

The method was also tried with other photometric stereo systems, such as the 4-light camera system. However, the small number of pictures and the lighting position lead to very poor and inconsistent results.

The memory usage of the method is around ~2 times the total size of images, which is acceptable. The method takes about 5 seconds to compute, which can be too much in real-world scenarios, where, in the case of an automated inspection system, almost real-time execution is required. However, the time can be significantly reduced, as most calculations were written in Python, which is an interpreted language.

5.5 Recommendations

The quality of light is by far the most important factor when comparing the quality of produced normal maps. As such, it is recommended to include filters to ensure linear directional light diffusion. The micro louvre filters used in the 4-light camera setup significantly increased the image quality produced by the device.

It is recommended to avoid capturing the images with high exposure settings, as this can lead to noise in the form of oversaturated highlights. It should be aimed to capture images with only as much exposure so that the whole object can be lit up.

When performing modifications on a scanning device, it is recommended to not remove the sensor assembly from the optical assembly, as this can cause errors to appear in the image and the placement of the sensor assembly is very hard to accurately adjust. Newer scanners use CIS (Contact image sensors) instead of CCD sensors. This may make it easier to perform modifications on, as the lighting used is LED instead of a fluorescent tube. However, CIS scanners usually have very poor depth of field, so the image might be out of focus with as little as 1mm of distance.

It is recommended to use basic L2 for normal map generation as long as the quality is sufficient, as it is much faster than the other methods.

6. Conclusions

6.1 Review of Aims

Most of the set-out goals were successfully achieved. A method of extracting embossed foil from a surface by using photometric stereo was created and tested. A lot of insight was gained into the possible further development of a photometric stereo-based quality control system for the box manufacturing industry. A proprietary data set of images was created that can be used to develop algorithms. An additional contribution was made by providing animated gifs which provide an easy way of inspecting the photometric stereo images in the datasets.

However, all methods of normal map extraction could be effectively compared due to their time complexity which could have provided more insights into the efficacy of the methods.

Review of related work	Achieved
Design of a scanner-based system	Achieved
Design of a ring-light-based system	Achieved
Comparison of normal map quality between 3 systems.	Achieved
Comparison of normal map quality between 4 methods of normal map acquisition.	Partially Achieved
Analysis of the systems and recommendations/methods for improvement.	Achieved
A method for finding specular/reflective areas of images using multiple lighting angles and a fixed camera position.	Achieved
A proprietary data set of images that can be used to develop algorithms.	Achieved
A human-readable method of viewing photometric stereo by combining multiple images into a GIF	Achieved, not a planned contribution

6.2 Future work

The main area of possible future work is continuing the development of embossed foil extraction techniques (diffuse-specular separation).

Additional normal map acquisition methods could be compared. Methods, such as the usage of gradient illumination or polarized light could be analysed.

An application that combines foil detection and normal maps to identify errors in foil placement and provide accurate measurements of the errors would be very beneficial to the packaging industry. This could be done by using a scanner-based system and knowing the physical dimensions of the scanned image. The tool could be used to automatically measure the misalignment in millimetres, or the tool could allow for an inspector to manually select the desired areas to measure

Performing quantitative comparison without a ground truth example can be very challenging. As such, it would be very beneficial if quantitative evaluation methods for photometric stereo could be developed without knowing the ground truth. This could be done by measuring metrics such as the deviation of normals on an elevated surface from the normals on surfaces of different elevations. Deviation of normals on the same particular elevated surface could also be used to measure the reliability and accuracy of the normal map.

6.3 Challenges

A couple of hard challenges had to be overcome in this project. The main challenge was the shortage of time, which was caused by the difficulty of getting laboratory access to work with hardware as well as the exceptional circumstances I had to face, where I was forced to abandon work for a few weeks.

Working with fragile and precise electronics was also a big challenge. One of the scanners had to be replaced after shorting a connection. Additionally, the sensor assembly of the scanner is very hard to recalibrate once it was taken off the optical assembly (which was needed to perform the necessary modifications).

Lastly, some photometric stereo methods could not be compared effectively due to their time complexity. Some of the methods took a very long time to produce results and as such had to be abandoned due to a shortage of time. Some data for the scanner system was missing as I did not have access to all of the carton sheets.

6.4 Skills learned

I developed a number of soft and technical skills throughout the project. Most of these skills gained were new and complex

The technical skills I improved upon were:

- Photometric stereo techniques.
- Computer Vision techniques
- Python skills.

I also improved upon several soft skills, which include:

- Project planning
- Working with resource and time constraints
- Working with flexible R&D projects.

6.5 Final remarks

Overall, I believe the project was a success. While I faced very hard challenges that heavily affected my work quality, I believe have completed most of my goals set out at the beginning of the project to a satisfactory degree. While I was not able to gather all comparison data or perform a quantitative analysis, I believe that the data I have gathered and the analysis I performed has provided some insight into the area. I believe the method I developed can also be improved and potentially used in real-world scenarios. I think the project could have gone a lot better and the report would have been of higher quality if not for the sudden event in my personal life that caused me to abandon work for a significant period of time, considering the overall length of the project. However, I am proud of myself as I always stayed determined to complete this project. I would like to thank the company I worked with during my placement, as they provided extensive help throughout this project.

References

- Christos, K., Stefanos, Z. & Abhijeet, G., 2018. *Diffuse-Specular Separation using Binary Spherical Gradient*. Netherlands, s.n.
- Fan, H. et al., 2022. Near-field photometric stereo using a ring-light imaging device. *Signal Processing: Image Communication*, Volume 107.
- Hao, D., Dan, B. G. & Steven, M. S., 2011. *Binocular Photometric Stereo*. University of Dundee, s.n.
- Harrod, S., 2021. *The Future of Folding Cartons to 2026*, s.l.: s.n.
- Herbort, S. & Wöhler, C., 2011. An introduction to image-based 3D surface reconstruction and a survey of photometric stereo methods. *3D Research*, Volume 2, pp. 1-17.
- Ikehata, S., Wipf, D., Matsushita, Y. & Aizawa, K., 2012. *Robust photometric stereo using sparse regression*. s.l., s.n.
- Jung, K. et al., 2017. A Photometric Stereo Using Re-Projected Images for Active Stereo Vision System. *Applied Sciences*, 7(10), p. 1058.
- Kirwan, M. J., 2008. *Paper and Paperboard Packaging Technology*. s.l.:John Wiley & Sons.
- Miyazaki, D., Tan, R., Hara, K. & Ikeuchi, K., 2003. *Polarization-based inverse rendering from a single view*. s.l., s.n.
- Müller, V., 1995. *Polarization-Based Separation of Diffuse and Specular Surface-Reflection*. Berlin, Springer.
- Plex Team, 2021. *THE IMPORTANCE OF QUALITY CONTROL IN MANUFACTURING*. [Online]
Available at: <https://www.plex.com/blog/importance-quality-control-manufacturing>
[Accessed 2022].
- POST PRESS MAGAZINE, 2020. *Embellishments on Folding Cartons: Continued Growth and Opportunities*. [Online]
Available at: <https://postpressmag.com/articles/2020/embellishments-on-folding-cartons-continued-growth-and-opportunities/>
- Smith, L., Zhang, W. & Smith, M., 2018. *2D and 3D Face Analysis for Ticketless Rail Travel*. s.l., s.n.
- Woodham, R., 1992. Photometric Method for Determining Surface Orientation from Multiple Images. *Optical Engineering*, Volume 19.
- Woodham, R. J., 1978. *Photometric Stereo*, Vancouver: s.n.
- Yamashita, A., Kawarago, A., Kaneko, T. & Miura, K., 2004. *Shape reconstruction and image restoration for non-flat surfaces of documents with a stereo vision system*. Cambridge, UK, s.n.

