SURVEY OF VARIOUS 3D IMAGING METHODOLOGIES

by

Chua Chern Fei

A Master's Report Submitted to the Faculty of the

COLLEGE OF OPTICAL SCIENCES

In Partial Fulfillment of the Requirements

For the Degree of

MASTER OF SCIENCE

In the Graduate College

THE UNIVERSITY OF ARIZONA

2017

STATEMENT BY AUTHOR

The thesis titled *Survey of Various 3D Imaging Methodologies* prepared by *Chua Chern Fei* has been submitted in partial fulfillment of requirements for a master's degree at the University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this thesis are allowable without special permission, provided that an accurate acknowledgement of the source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: *Chua Chern Fei*

APPROVAL BY THESIS DIRECTOR

This thesis has been approved on the date shown below:

<u>03 May 2017</u>

Jim Schwiegerling   Date

Professor of Optical Science

ACKNOWLEDGEMENTS

I would like to thank the faculty and staff of the College of Optical Sciences, it has been a pleasant and fruitful learning journey for me. I would like to thank the chair of my Master's Committee, Professor Jim Schwiegerling, as well as the members of the Master's Committee, Professor John Koshel and Professor Rongguang Liang. I have learnt a lot through all of you throughout my graduate studies.

I would also like to thank Cindy Roberston for her help in the distance program. Thanks to Mark Rodriguez for helping me on the administrative matter. Special thanks to Professor John Koshel who has been very helpful in sorting out some of my concerns along the way.

Lastly, I would like to show my gratitude towards my boss, Wong Soon Wei, for his financial support of my master's program.

Finally, thanks to my wife for supporting me from every aspect in my life.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLE

Introduction

3D imaging can be considered as the extension of 2D imaging with additional depth information. It is becoming more common owing to the ever increasing computation power. There is increasing demand in 3D imaging solutions due to the advancement of robotics and automation, as well as the improvement of manufacturing process. The applications of 3D imaging include reverse engineering, range sensing, verification of fabricated parts, surface measurement, industrial inspection, biomedical imaging, etc.

There exist various types of 3D imaging methodologies. In this report, we will cover:

- Stereo vision
- Laser scanning
- Structured light
- Depth from focus
- White light interferometry
- Optical coherence tomography

Among the list above, stereo vision, laser scanning, and structured light technologies rely on the principle of triangulation. White light interferometry and optical coherence tomography are interferometric type instrumentations. In this report, we will discuss their working principles, applications, limitations, as well as some case studies.

Chapter 1: Stereo Vision

Stereo vision is considered as the simplest form of 3D imaging, it is the basic principle how human eyes perceive depth. Just like human vision, stereo vision 3D imaging system uses a pair of cameras to extract the depth information of a scene. By comparing the two digital images from the two cameras, using principle of triangulation, one can extract the depth information.

1.1 Imaging as projection

The 3D stereo vision can be modeled as a pair of ideal 2D imaging systems, where each 2D imaging system simply projects arbitrary point in real world (x,y,z) into an flat imaging plane (x',y'), where (x',y') and (x,y) are related by image magnification.

Conventional 2D imaging system creates inverted image, it becomes more convenient if one 'flips' the image 'in front of' the camera lens. In this configuration, it becomes simpler as if the scene is directly projected onto the camera's sensor without inverting image. This is illustrated in figure 1.1.
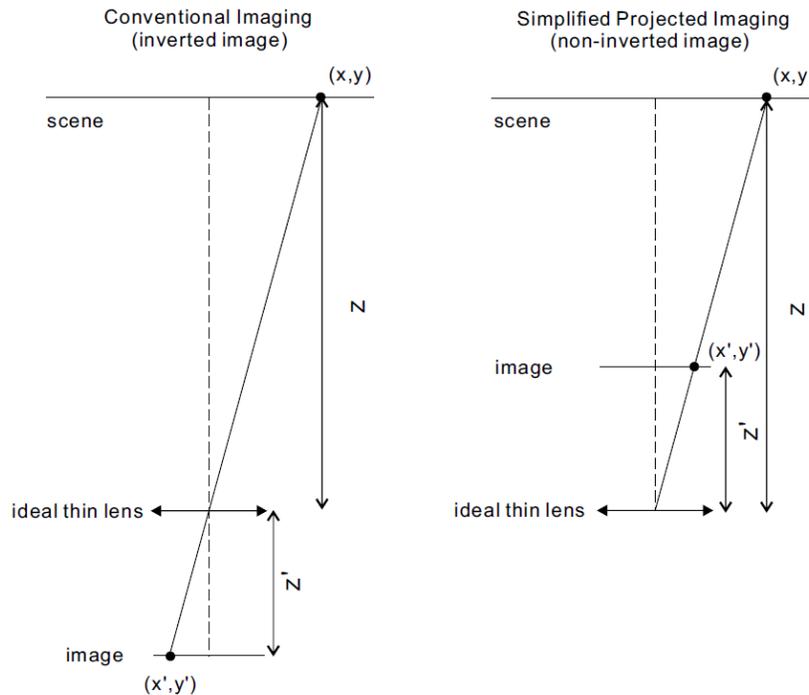
Figure 1.1: (Left) Conventional imaging with inverted image
(Right) Simplified projected imaging where image is projected with positive magnification.

The ideal 2D imaging is described by first order optics, or Gaussian imagery. The scene is imaged by an ideal thin lens, where a perfect image is formed at each camera. Here the 2D imaging system has been converted from conventional imaging (inverted image) to the

projection imaging (non-inverted), the z and z' are defined along the same z axis, hence both are positive quantities. The imaging condition by Gaussian imagery is described by:

$$\frac{1}{z'} + \frac{1}{z} = \frac{1}{f}$$

(1.1)

And the image magnification is defined as:

$$m = \frac{z'}{z}$$

(1.2)

Here the image magnification is positive as well. Lastly, the image location is

$$z' = (1 + m)f$$

(1.3)

1.2 Simplest case: Stereo pairs in parallel form

The simplest form of 3D stereo vision is the case when both left and right cameras are parallel to each other, as shown in figure 1.2.
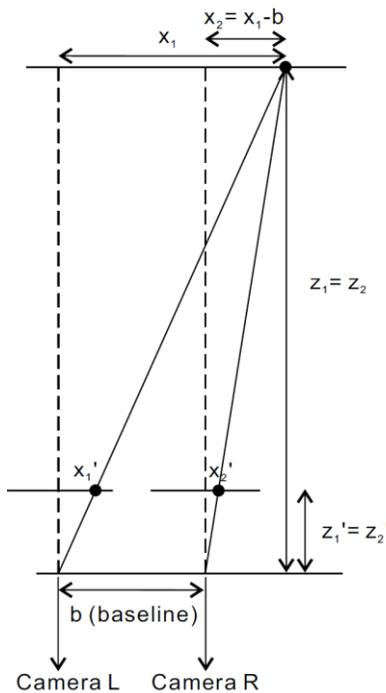
Figure 1.2: Stereo pair with left and right cameras imaging the same scene, with optical axes aligned.

Using similar triangle to both left and right camera, we have:

$$\frac{z_1}{z_1'} = \frac{x_1}{x_1'} \qquad , \qquad \frac{z_2}{z_2'} = \frac{x_2}{x_2'}$$

(1.4)

But, $z_1 = z_2 = z$, $z_1' = z_2' = z'$, and $x_2 = x_1 - b$, where b is defined as the "baseline", it is the distance between the left and right camera. We then have:

$$\frac{z}{z'} = \frac{x_1}{x_1'} = \frac{x_1 - b}{x_2'} \tag{1.5}$$

As we can see, the same object point, is mapped into different locations of the left and right images, such difference is also known as "disparity", which is defined as

$$d = x_2' - x_1' \tag{1.6}$$

Suppose given an object point, we are able to match the left and right images, or in other words, the disparity values can be computed, we then have

$$d = \frac{z'}{z}[x_1 - (x_1 - b)] \tag{1.7}$$

$$z = \frac{bz'}{d} \tag{1.8}$$

For a calibrated imaging system, the focal length and magnification are known, hence z' is known. For a given distance between two cameras (baseline), we can determine the depth information for a given scene by principle of triangulation.

1.3 Depth resolution as a function of baseline

From equation 1.8, it is apparent that the depth is inversely proportional to the disparity. In other words, points of stereo pair images with bigger disparity correspond to object at nearer location, this concept is illustrated in figure 1.3.

Due to the relation of the depth with disparity, it follows that the depth measurement become coarser as the range increases, this effect is illustrated in figure 1.3. In figure 1.3, the image planes of the stereo pair are sampled with 14 pixels, with each pair of adjacent projection lines corresponds to a single pixel field of view.

Figure 1.3 shows the iso-depth planes, these are the planes that have the same depth with equal disparities (in integer number of pixels). The spacing between iso-depth planes, in other words, the depth resolution, depends on the baseline (distance between the stereo pair) and the image magnification.

This effect is illustrated in figure 1.3, when the baseline or/and image magnification is increased, the sampling of the depth at a given distance increases. However, the overlapped field of view to both cameras is reduced as well. Therefore, one should find the best balance between depth resolution and overlapped field of view for a stereo pair system.
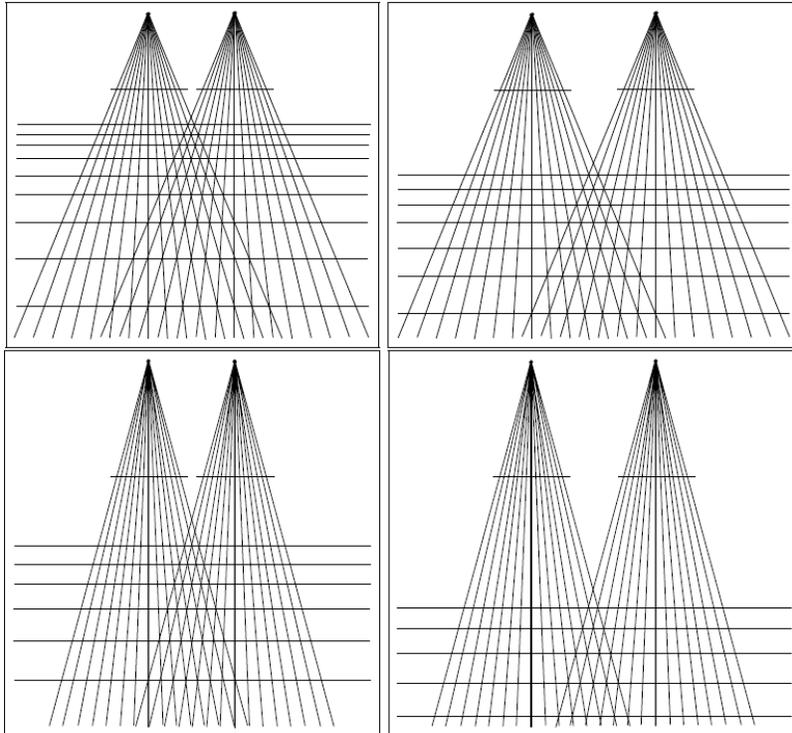
Figure 1.3: Iso-depth planes of stereo pair system as a function of baseline and image magnification, figure courtesy of "3D Reconstruction from Multi Images Part 1: Principles", by Theo Moons et. al. [1]
(Top left): Original configuration (Top right): Increased baseline (Bottom left): Increased image magnification (Bottom Right): Increased both baseline and image magnification. Increasing either baseline or image magnification results in closer iso-depth planes (finer sampling) at a given distance, but the overlapped field of view is reduced.

1.4 Stereo pairs in convergent form

It is also common to have stereo pairs in non-parallel form. In the convergent form of stereo vision, both cameras are placed at a certain angle. This configuration offers advantage over the parallel form, the field of view of the stereo vision is potentially larger, as shown in figure 1.4.

Parallel Configuration : Red

Convergent Configuration : Blue

Field of view of individual camera

Optical axis of individual camera

Field of view of stereo vision (overlapped region)
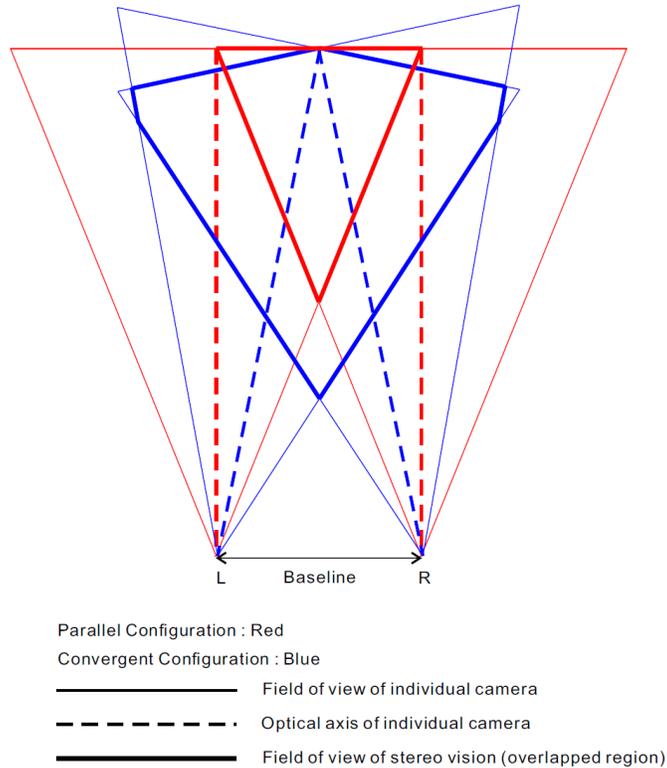
Figure 1.4: Stereo vision in parallel form vs. convergent form.

For stereo vision in convergent form, it becomes more complicated to match the left and right image. Figure 1.5 shows stereo pairs in general form. For the left camera, MO is the line of sight to the camera, hence the images of the points along this line (A, N, Q) are formed at the same location. To the right camera, the images of three these points A, N, Q are formed along a line, which is known as epipolar line.

Similar, for the right camera, MO' is the line of sight to the camera, and the images of the points along this line (M, P, R) are formed at the same location. To the left camera, the images of these three points M, P, R are formed along another epipolar line.

The epipolar lines lead to very important feature in stereo vision, the matching between the left and right images are reduced from a 2D domain (whole image) into a 1D line (epipolar line). That is, for a point in the left image, to find the corresponding point in the right image, one needs to search along the epipolar line. In general, the epipolar line is slanted, it will be easier if one is to transform the epipolar line to be horizontal.
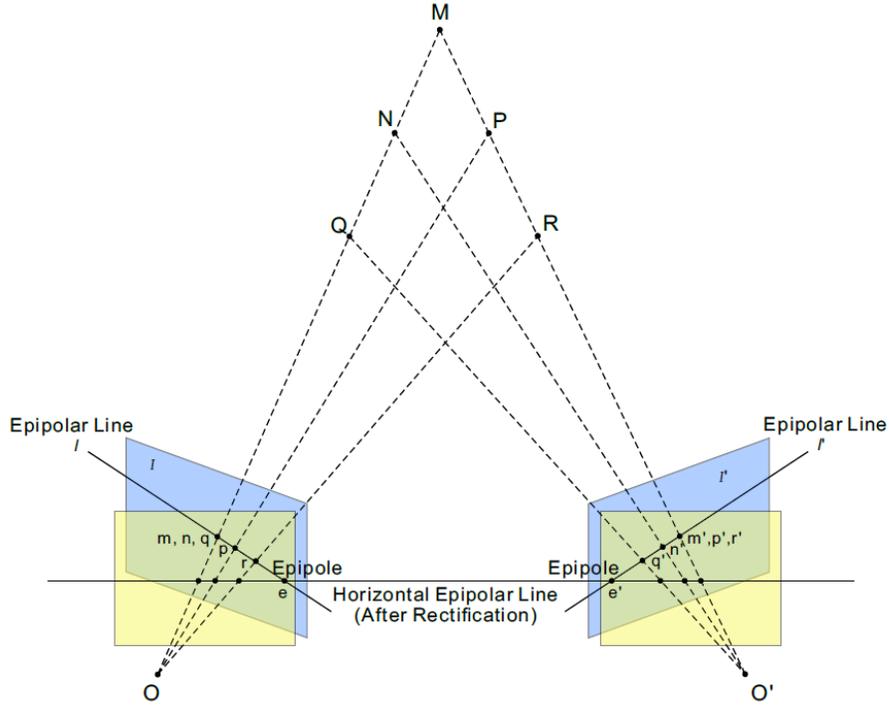
Figure 1.5: Stereo pair in general form and after image rectification

The process to transform the original stereo image pair into the standard form is called image rectification. After the images are rectified, the 2 epipolar lines will be merged into a horizontal line, the image matching between the left and right images will be simply along the horizontal line, this process is illustrated in figure 1.5. After image rectification, the image planes will become co-planar, they will be equivalent to what we will obtain from stereo setup in parallel form.

There exists different methods of rectifying a pair of stereo images, one of the most commonly used method was introduced by Charles Loop and Zhengyou Zhang [2]. It starts by defining a matrix F, which is known as fundamental matrix, which satisfies:

$$\mathbf{m'}^T\mathbf{Fm} = 0 \qquad\qquad (1.9)$$

for all pairs of images corresponding to m and m'. This is also known as epipolar constraint. The fundamental matrix F also satisfies the following:

$$\mathbf{Fe} = \mathbf{0} = \mathbf{F}^T\mathbf{e'} \qquad\qquad (1.10)$$

where e and e' are the epipoles of the left and right images respectively. The image rectification requires the epipoles to be relocated at infinity. To achieve image rectification, H and H' matrices are the homograhies operations that to be applied to images I and I' respectively. Upon rectification, the image points m bar and m' will be:

17

$$\overline{\mathbf{m}} = \mathbf{Hm} \quad , \quad \overline{\mathbf{m}}' = \mathbf{H}'\mathbf{m}' \tag{1.11}$$

After image rectification, it follows from equation 1.9 that:

$$\overline{\mathbf{m}}'^T \overline{\mathbf{F}} \overline{\mathbf{m}} = 0 \tag{1.12}$$

Substituting equation 1.11 into equation 1.12:

$$\mathbf{m}'^T \mathbf{H}'^T \overline{\mathbf{F}} \mathbf{Hm} = 0 \tag{1.13}$$

$$\mathbf{F} = \mathbf{H}'^T \overline{\mathbf{F}} \mathbf{H} \tag{1.14}$$

For a pair of rectified images, the epipoles are at infinity, and $\overline{\mathbf{F}}$ is pre-determined. The homographies H and H' that satisfy equation 1.14 is not unique. The proposed H and H' are constructed to minimize distortion, and they are further decomposed into 3 individual transformations, namely: Projective Transform, Rotation Transform and Shearing Transform.

$$\mathbf{H} = \mathbf{H_s}\mathbf{H_r}\mathbf{H_p} \quad , \quad \mathbf{H}' = \mathbf{H}'_s\mathbf{H}'_r\mathbf{H}'_p \tag{1.15}$$

Figure 1.6 shows the process of image rectification by applying respective transformation in 3 steps.
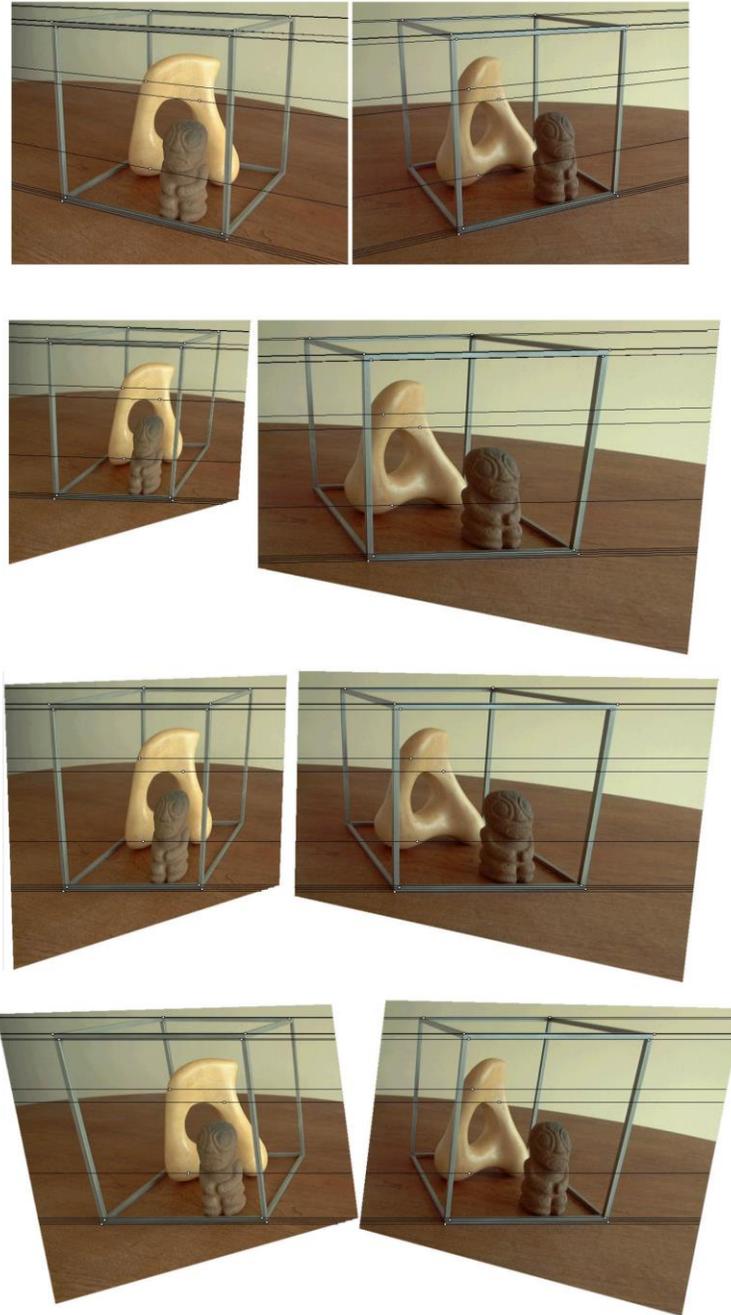
Figure 1.6: Image rectification in steps, from top to bottom:
(1st step) Original image pair overlayed with several epipolar lines.
(2nd step) Image pair after projective transform. Epipolar lines become parallel.
(3rd step) Image pair after rotation transform. Epipolar lines become horizontally aligned.
(4th step) Final image rectification after shearing transform,
with minimized horizontal distortion.

1.5 Correspondence problem

Even when the images are perfectly rectified, finding the corresponding point viewed by one camera in the image of the other camera can be challenging, this is known as the correspondence problem.

After image rectification, a point in the scene will be mapped into the left and right images on the same horizontal line ($y_1=y_2$), but different location ($x_1 \neq x_2$, the difference x_2-x_1 is the disparity d). Therefore, the simplest form of image matching is achieved by finding the disparity value that gives minimum of absolute intensity difference (AD) across the horizontal scanline:

$$AD(d) = \min |I_L(x, y) - I_R(x - d, y)| \qquad (1.16)$$

where *d* is the disparity parameter that scans across a horizontal line. This method attempts to find the best matching by comparing the left and right images pixel by pixel. The computed disparity will be the *d* value(in pixels) that yields the minimum absolute intensity difference.

One might expect this "brute" method fails because in real situation, the noises of the images will lead to noisy disparity mapping. Practically, instead of matching pixel by pixel, we match small patch of the image, which is known as "matching window", as shown in figure 1.7. The size of the matching window affects the result of disparity mapping. In general, smaller window size leads to more detailed mapping, but the result tends to be noisy. On the other hand, larger window size produces less detailed mapping, but fewer isolated errors.
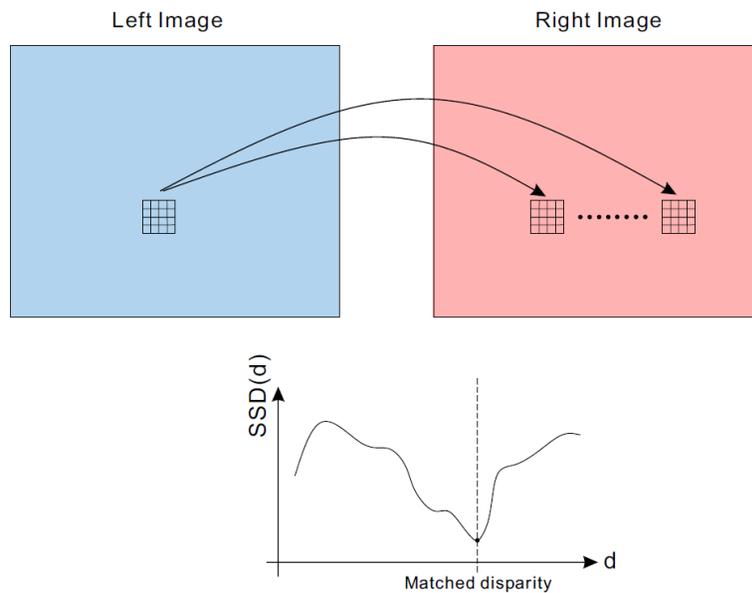
Figure 1.7: (Top) Matching the left and right images for a small window.
(Bottom) The disparity is determined when minimum value of SSD is found.

There exists various kinds of algorithms to achieve image matching, two of the most popular and simplest algorithms are the sum of squared difference (SSD) and cross correlation (CC). Applying the matching window, the formula is summarized below:

$$SSD(d) = \min \left| \sum_{(x,y)\in W} [I_L(x,y) - I_R(x-d,y)]^2 \right| \tag{1.17}$$

$$CC(d) = \max \left| \sum_{(x,y)\in W} [I_L(x,y) \cdot I_R(x-d,y)] \right| \tag{1.18}$$

Similar to the AD method, the SSD method looks for the minimum of the squared difference, this is illustrated in figure 1.7. For the CC, it looks for the maximum value. Both methods attempt to find the highest similarity between the left and right images. In fact, the sum of squared difference is closely related to the cross correlation:

$$SSD(d) = \sum_{(x,y)\in W} [I_L(x,y) - I_R(x-d,y)]^2 \tag{1.19}$$

$$= \sum_{(x,y)\in W} I_L(x,y)^2 + \sum_{(x,y)\in W} I_R(x-d,y)^2 - 2\sum_{(x,y)\in W} I_L(x,y) \cdot I_R(x-d,y)$$

The circled term in equation 1.19 is the cross correlation.

It is also very common to normalize the image intensity of the left and right images, because the illumination, exposure of the left and right camera may vary in some cases. The SSD and CC method are still valid, only with the intensity of the images normalized:

$$\hat{I}(x,y) = \frac{I(x,y) - \bar{I}(x,y)}{\sqrt{\sum(I(x,y) - \bar{I}(x,y))^2}} \tag{1.20}$$

where $\bar{I}(x,y)$ bar is the mean value of the window intensities. By normalizing intensity to both left and right images, the result will be less sensitive to intensity difference between the left and right images.

However, in most of the situations, SSD and CC algorithms are still considered too simplified, the resulting disparity map tends to have errors. Therefore, stereo matching algorithm has been intensively studied, researchers aim to develop algorithm that produces disparity map not only with less error, but also with lesser computational time. Some of the well-known techniques include graph cuts [3], belief propagation [4], dynamic programming [5], etc, but the computation cost is considered fairly high.

1.6 Issues and challenges

The stereo matching is even more challenging in some situations, the more common ones are: photometric variation, occlusions, reflections, transparent objects, foreshortening, perspective distortion, featureless object, repetitive structures, etc. Some of the above

situations are illustrated by the stereo pair images below. (Image courtesy of research group "Computer Aided Medical Procedures" from Technical University of Munich [6])
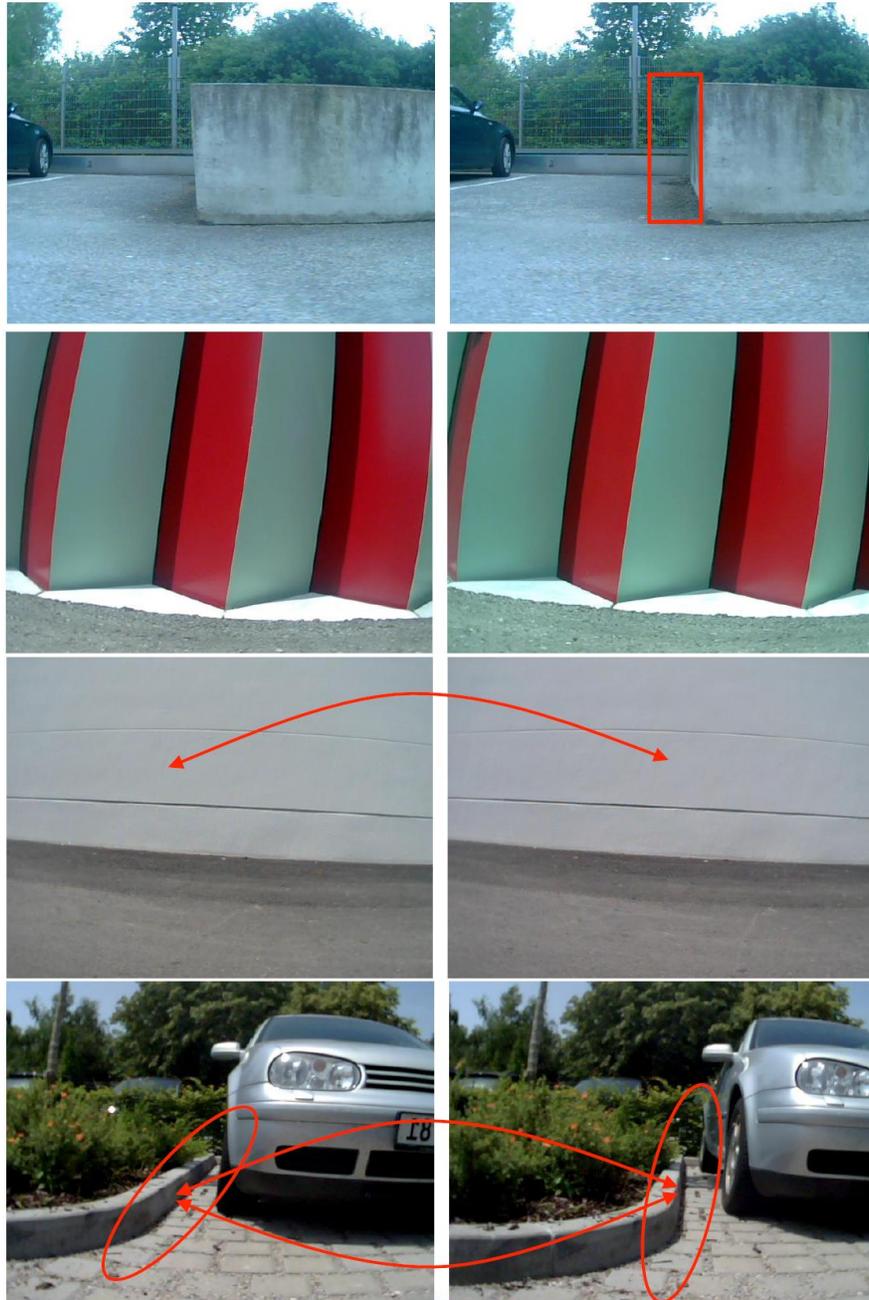
Figure 1.8: Some challenging situations in stereo matching, top to bottom: occlusion, repetitive structure, featureless object, and foreshortening effect [6]

In such situations, the constructed disparity map tends to have errors. Therefore, some of the more developed image matching method had included additional algorithm in order to deal with such challenging situations.

1.7 Case Study

Case study 1: Chocolate volume inspection

As discussed in previous sections, stereo vision can be computationally challenging when it comes to matching the left and right images. That is, when a stereo vision setup is used to measure depth of unknown scene, the result is not guaranteed to be good because of possibly unexpected situation.

However, if a stereo vision is designated for a specific task, the result can be quite reliable and repeatable. Here let us take a look at an example of 3D automatic inspection on chocolate height by LabVIEW [7]. LabVIEW is a system design platform with vision programming language by National Instrument.
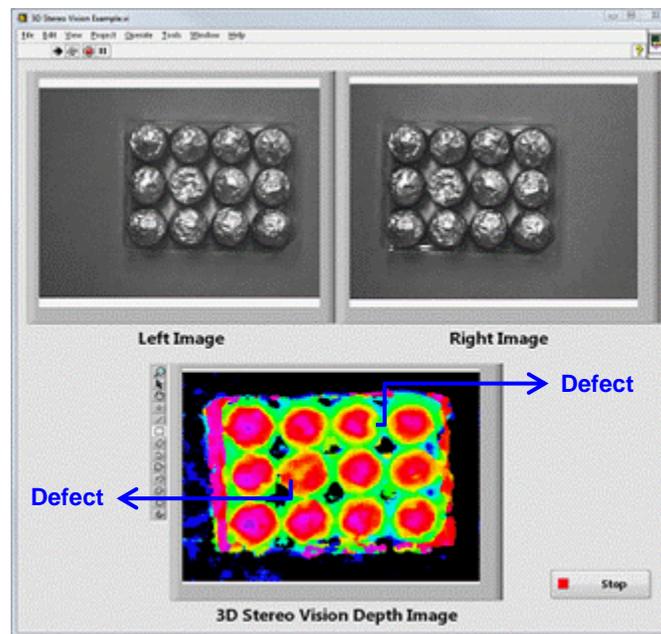
Figure 1.9: Quality inspection of chocolate in a box: 3D imaging allows much easier defect detection as compared to traditional 2D imaging. 3D depth image reveals serious defective chocolate at 2nd row, and slightly defective chocolate at 1st row.

Figure 1.9 shows a box of spherical chocolate is imaged by a stereo vision system which consists of a left and a right camera. By looking at the 2D image, it may look obvious to human brain that the 2nd chocolate of the 2nd row is wrecked/dented. However, with traditional 2D image, detection of such defective chocolate is difficult. With 3D stereo vision setup, after calibration, the depth information allows easier detection of such defective chocolate due to insufficient volume.

Case study 2: Wire bond inspection

In the manufacturing industry that relies on automatic 3D inspection for a certain kind of object, the inspection tends to be very specific. In this case, the algorithm of the 3D mapping can be optimized accordingly and the result is much more reliable. Another good example is the wirebond 3D inspection, as shown in figure 1.10. 3D stereo vision system has been routinely used for inspecting surface quality and height profile of the wire bonds on an integrated circuit board.
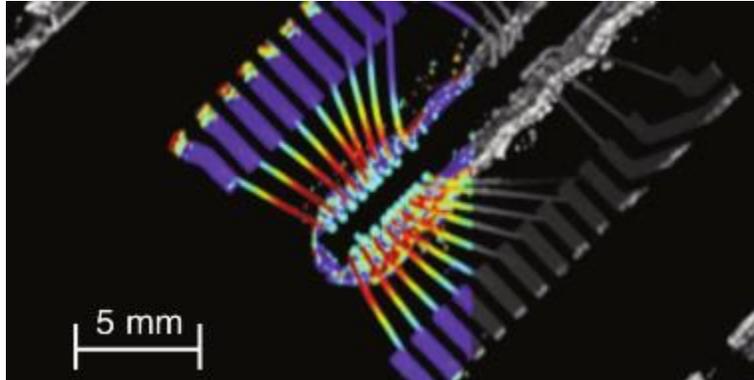
Figure 1.10: 3D Wire bonding inspection on an integrated circuit board.
(Image courtesy of Chromasens GmbH [8])

1.8 Conclusion

In conclusion, stereo vision provides rather simple hardware to extract depth information via principle of triangulation. However, stereo matching is in general computation costly, it can become rather challenging for various kinds of scenarios.

Chapter 2: Laser Scanning

Laser scanning is a widely used technique for 3D imaging application which relies on principle of triangulation. By observing the displacement of the laser light scattered on the object surface, we can calculate the distance of the object (range sensing). By doing scanning across the field of view, we can reconstruct the full 3D model of an object.

2.1 Single point range sensing

The basic principle of the laser scanning is based on the principle of triangulation. Here we start with determining the range of a single point. Figure 2.1 shows an object line profile lies in x-z plane to be imaged onto a sensor and a laser is shined towards the object and making intersection at point (x,z). Suppose the object is a diffuse surface, the illuminated spot is imaged onto the sensor.
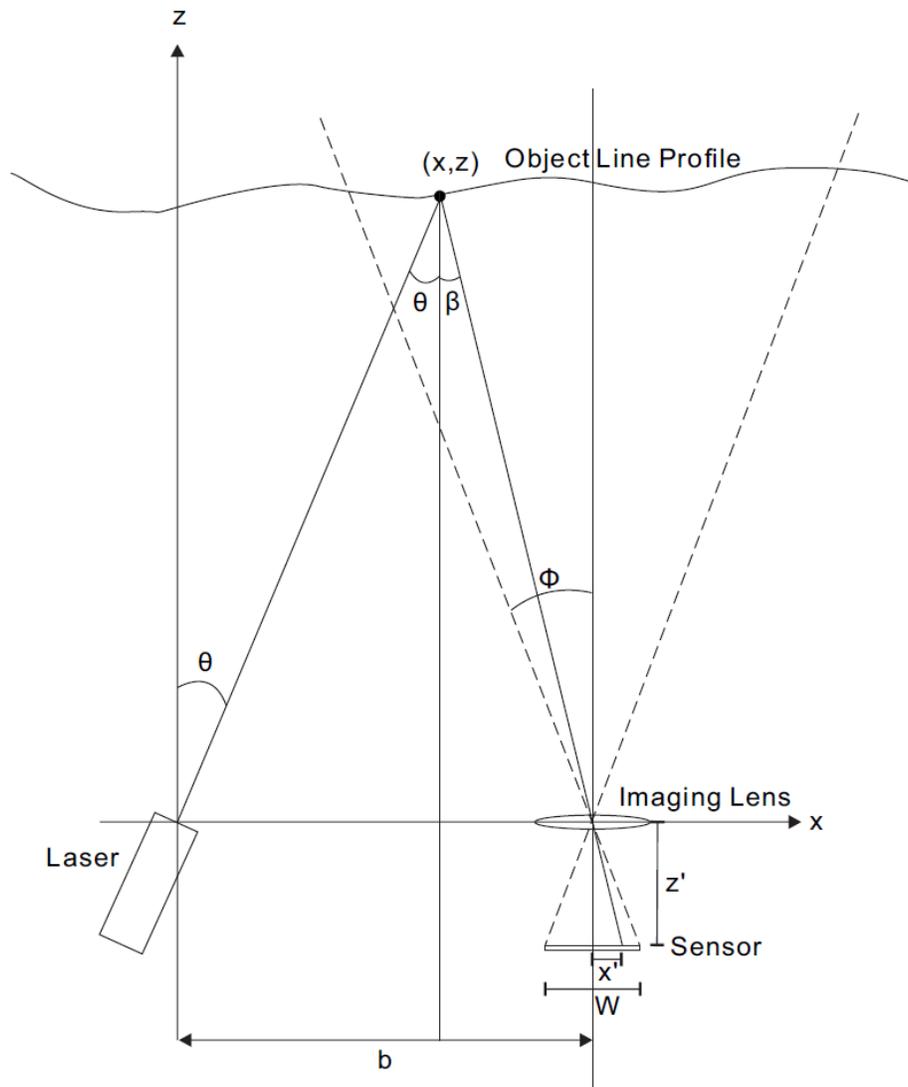
Figure 2.1: Schematic diagram of laser triangulation principle.

From trigonometry, we have:

$$b = z(\tan\theta + \tan\beta) \tag{2.1}$$

$$\tan\beta = \frac{x'}{z'} \tag{2.2}$$

$$z = \frac{bz'}{z'\tan\theta + x'} \tag{2.3}$$

From first order optics (Gaussian imagery), we know that for distant object, the image of the object point is close to the focal point, hence $z' \simeq f$. For finite conjugate, z' is proportional to the focal length of the imaging system, which is given by $z' = (1 + m)f$, where f is the focal length and m is the magnification (positive quantity).

$$z \simeq \frac{bf}{f\tan\theta + x'} \tag{2.4}$$

To describe the accuracy of the measurement, we can estimate the error in the range Δz. Its inverse quantity 1/Δz describes the accuracy [9]:

$$\Delta z = \left|\frac{z^2}{fd}\right|\Delta x' + \left|\frac{z^2}{d\cos^2\theta}\right|\Delta\theta \tag{2.5}$$

where Δx' is the uncertainty of the image location, which is essentially determined by the range sensor resolution (i.e. pixel size). Δθ is the uncertainty of the incident angle of the laser. In the most common setup, it is determined by the mechanical accuracy of the scanning mechanism.

The accuracy (1/Δz) of the range sensing is therefore proportional to the separation between the laser and the sensor, and the focal length of the imaging lens, but inversely proportional to the square of the distance.

Unfortunately, the focal length f and the separation d cannot be made as large as desired. If the focal length is increased, the field of view will be compromised. The field of view of the range sensor is simply:

$$FOV = 2\Phi = 2\tan^{-1}\left(\frac{W}{2z'}\right) \simeq 2\tan^{-1}\left(\frac{W}{2f}\right) \tag{2.6}$$

Equation 2.6 shows that the field of view is inversely proportional to the focal length of the imaging lens. This implies that the trade-off between the field of view and the range accuracy must be considered.

If the separation between the laser and the range sensor is increased, the overall mechanical structure will not be compact, this could be a potential issue for certain applications. Moreover, if the separation d is increased, the shadow effect will become

worse, as shown in figure 2.2. The shadow effect arises due to the slanted view angle towards the object. We will discuss more about the technique to eliminate the shadow effect.
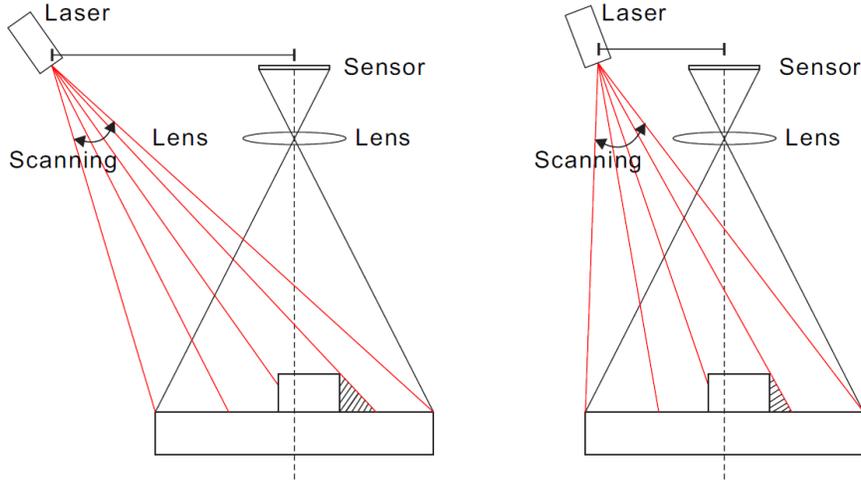
Figure 2.2: Shadow effects for different separation between the laser and the sensor.

From figure 2.1, once we determine the distance z of the particular point (x,z) of the object, the position x is then given by

$$x = z \tan \theta \tag{7}$$

Its error is therefore

$$\Delta x = |\tan \theta| \Delta z + \left| \frac{z}{\cos^2 \theta} \right| \Delta \theta \tag{8}$$

From the above calculation, we obtain the position of a single point (x,z) of the object. If we scan the laser across the field of view of the sensor, we obtain a line profile of the object with depth information.

Up to this point, we have discussed how the position of a single point can be determined via principle of triangulation. To obtain 3D image, there has to be certain mechanism to scan the laser spot across the field of view. The simplest way is to do a two directional (x-direction and y-direction) linear scanning of the object, but this approach is extremely inconvenient (moving object in x-y plane) and time consuming. Here we will discuss some of the more practical approaches.

2.2 Line scanning
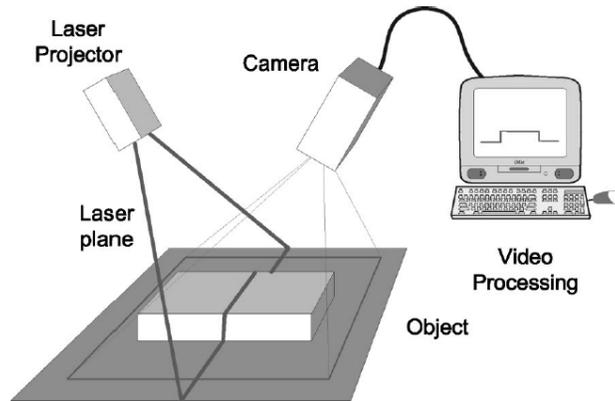
Figure 2.3: Schematic diagram of laser triangulation with a line laser [10].

The most common laser triangulation technique is to scan the object with a line laser, as shown in figure 2.3. For every single image acquisition of the line laser profile, we obtain the line profile of the object. By performing scanning in one direction, we obtain the 3D model of the scanned object. Because of its mechanical simplicity with just one required scanning mechanism, this method is often the most economical and hence most widely used.

It's well known that the output of a laser is usually a circular Gaussian beam. We can easily expand the Gaussian beam in single direction that it becomes an elongated elliptical Gaussian beam with a cylindrical lens. We have to take note that this method produces a line laser, but the profile is not uniform (Gaussian profile). In most applications, such beam profile is acceptable. Depending on the situation and the level of performance, non-uniform laser line profile may not be desired as the signal to noise ratio across the field of view is not constant. The center portion of the laser line tends to be overly saturated, making the measurement not accurate. For the applications where high uniformity line laser profile is required, one can consider using specialized lens element to achieve it. Such lens element has complex aspheric surfaces that generate a tremendous amount of spherical aberration that redistributes the light along the line, as shown in figure 2.4 [11].
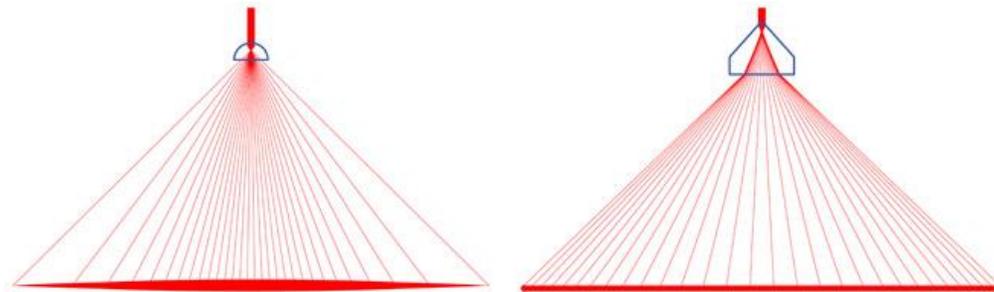
Figure 2.4: Line laser produced by conventional cylindrical lens (left)
vs. specialized optics (right), image courtesy of Laserline Optics [11].

For single point laser scanning, the sensor is simply a linear CCD/CMOS array. For laser line scanning, the sensor has to be a 2D CCD/CMOS array instead. With today's technology, 2D imaging sensors are readily available, making the line laser scanning technology an affordable solution to 3D imaging.

2.3 Line scanning mechanisms

Rotating the object

With the line laser and the image sensor staying stationary, placing the object of interest on a 360° turn table allows the entire object to be fully scanned. Such setup can be realized with a video camera, a line laser and a rotational motorized stage. When the scanning is completed, each frame of the video contains the "line image" captured at different angle of the object. By analyzing the "line images" and applying the principle of triangulation, 3D model can be reconstructed. This is a common laser scanning methods to obtain 3D models of objects such as antiques, model figures, fabricated parts, etc.

Rotating the line laser

Depending on the application, sometime we are only interested on the surface topography of an object. By rotating the angle of the laser (e.g. with stepper motor), the line laser will be scanned across the field of view, thus a 3D image of the front surface of the object will be obtained. This method does not require the movement of the object, but it only scans one side of the surface profile. For example, it can be used to examine the engravings of an artwork.

Moving the object

Similarly to the above method, instead of scanning the laser (by rotation) across the field of view, the object of interest is to be scanned linearly. This allows the 3D image of the front surface of the object to be reconstructed. Such method is most often used in the manufacturing industry where the rapidly manufactured end-products have to be inspected for their surface topography. After the end-products are manufactured, they may be transferred to the packaging station via conveyor belt. In this case, prior to the packaging station, the line laser and the image sensor can be installed on top of the moving end-products, and then the 3D images can be reconstructed.

Handheld laser scanner

The laser source and the sensor can be integrated as a single handheld device. By manually scanning through the object of interest, 3D image can be reconstructed. Usually with single scanning, the reconstructed 3D image may contain "holes" due to obscuration or shadow effect. By overlapping 3D images with multiple scanning attempts from different angles, a solid 3D model can be reconstructed. Handheld laser scanner is especially useful for obtaining 3D model of large objects such as high precision parts of automobiles and aircrafts.

2.4 Multiple lasers or multiple cameras setups

As mentioned earlier, shadow or obscuration effect causes missing points in the 3D "point clouds". A straightforward technique is to use two lasers, one of the lasers is shined towards the object at the opposite direction from the other laser. This setup addresses the shadow effect issue as shown previously in the figure 2.2.

There exists another form of setup –"Dual-view triangulation" which comprises of two cameras and a single laser beam to improve the range accuracy by a factor of $\sqrt{2}$ [10]. In addition, the image of the laser will be symmetric on the two sensors while an outlier (e.g. specular reflection) will be asymmetric, making the data acquisition less sensitive to environmental light.

2.5 Scheimpflug configuration
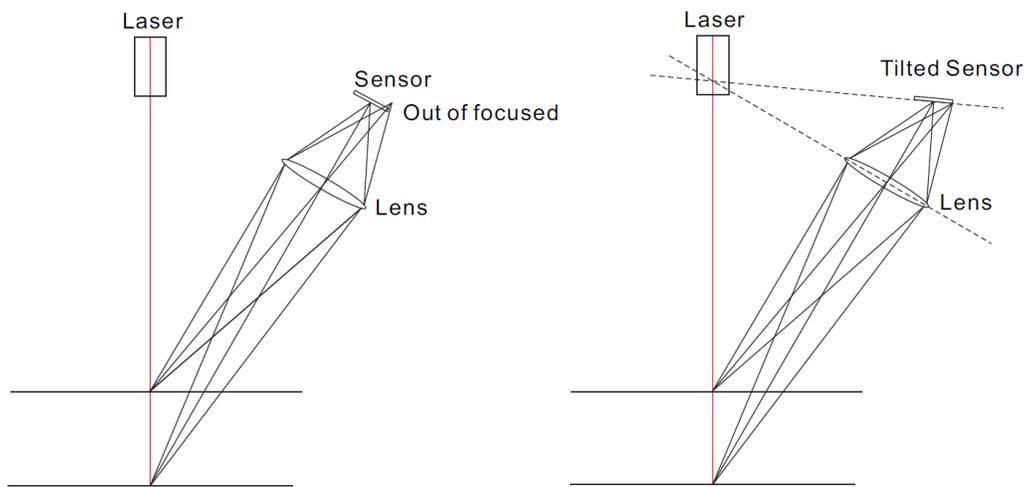
Figure 2.5: Laser triangulation with conventional setup (left)
vs. Scheimpflug condition (right).

Figure 2.5 shows the difference of the laser triangulation technique between the conventional setup and the Scheimpflug condition. The Scheimpflug condition is achieved when the angles of the sensor, imaging lens, and laser's line of sight are oriented in such a way that the extended lines coincide to a single point.

We can see that the Scheimpflug condition allows the range sensing with larger depth. For the conventional setup, the depth of the range sensing is dictated by the geometrical blur when the object distance is beyond the depth of focus of the imaging system.

The Scheimpflug condition is especially useful for the case when the depth of focus is limited, e.g. 3D imaging of small object with high magnification. However, the Scheimpflug condition is not often used as it introduces complexity that the associated cost may not be justifiable for general laser scanning applications.

2.6 Calibration

Calibration is required in order to obtain the measurements of the scanned objects. A priori known reference geometry, such as a checkerboard, can be placed into the field of view of the camera for calibration purpose. The algorithms for laser scanning calibration were well studied in the past, different methodologies were proposed, they are applicable to different applications [12,13,14,15].

Usually, calibration is to be done with the software packaging prior to the 3D scanning acquisition of the object. DAVID Laserscanner [16] is a popular software which allows user to obtain 3D model by scanning (manual or motorized) a line laser across the object at low cost. Figure 2.6 shows an example of the object to be scanned placing in front of the background cardboard with printed dots of known spacing for calibration purpose.
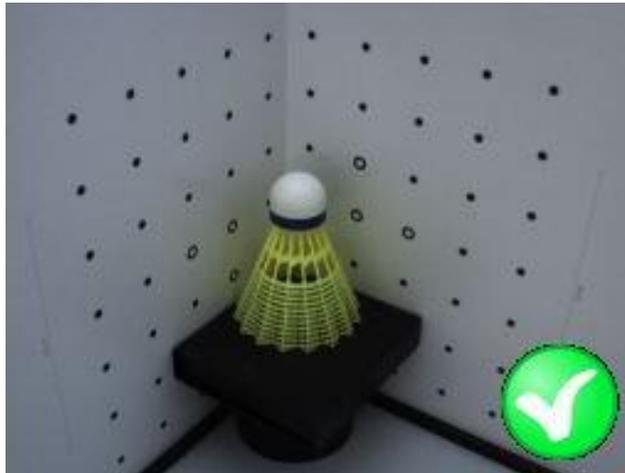
Figure 2.6: Calibration procedure by placing the object in front of known background, image courtesy of David laser scanner's user manual [17].

2.7 Single point flying spot laser triangulation scanner

Range sensing back on single laser spot triangulation technique was developed back in the 1970s. As mentioned earlier, single point triangulation requires the laser spot to be scanned over the object, in both horizontal and vertical direction, as shown in figure 2.7.
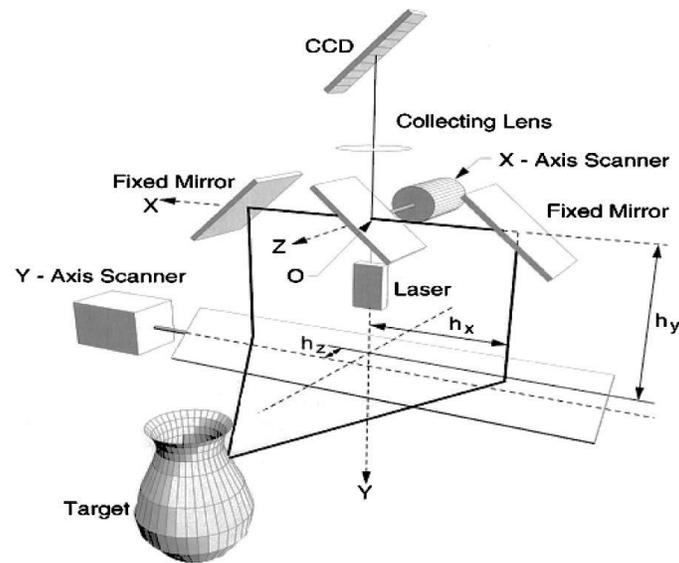
Figure 2.7: Single-point laser scanning using longitudinal synchronization [10].

The associated mechanical parts and the electronic devices lead to much higher cost comparing to the line scan based triangulation method. Moreover, the advancement of the 2D array sensor and the laser source (higher power) made the line scan based triangulation much more favorable than single point based triangulation.

As mentioned earlier, there is compromise between the field of view and the depth resolution as seen from equation 2.5 and 2.6. It was demonstrated that synchronized scanner techniques [9,18,19] can achieve large field of view scanning without sacrificing the depth resolution. Furthermore, for line scanning object surface with non-uniform reflective characteristic, it is difficult to find a setting of laser power and camera integration time across the laser line [19], which leads to false interpretation of data points. Single point based triangulation scanning essentially eliminates this issue because the data is acquired for every single point of the object surface. For line scan based triangulation, the power of the laser has to be spread into a line, hence the signal to noise ratio is reduced. While this is not a major concern for indoor environment, it becomes relevant for outdoor application where there is strong ambient light in the surrounding.

Today, most of the consumer level laser triangulation scanners are line scan based because of the overall robustness and cost effectiveness. Single point based triangulation scanning is applicable to more demanding situations, such as industrial or research level application.

2.8 Difficulties and limitations

As laser scanning requires mechanical scanning mechanism, the process of 3D image acquisition is considered time consuming. Another potential issue with usage of laser is the speckle noise. Speckle patterns typically appear when laser light is incident on diffuse surface, and they can be suppressed by the use of speckle reducer if necessary. However, diffuse surface is usually much more preferred than shiny surface. The specular reflection from a shiny surface may potentially lead to erroneous measurement in a simple setup. Lastly, laser scanning method does not retain the color information, unlike the stereo vision approach.

2.9 Case study

Case study 1: CMM Laser scanner

One application of laser scanning technology is obtaining 3D model of an object in place of conventional coordinate measuring machine (CMM). The conventional CMM is a device that allows the user to measure the physical geometry of an object with a mechanical probe. CMM is often used to verify if a fabricated part (by machining, injection molding, etc) is made correctly.

Conventional touch-probe based is time consuming, and it requires the user to be well trained in using the machine. On the other hand, laser scan based CMM is able to obtain the measurements with much shorter time scale. Furthermore, laser scan based CMM is a non-contact based measurement device, it is especially useful for measuring fragile or elastic objects, and parts with geometry that is difficult to access with a mechanical probe.

Figure 2.8 shows the difference between a laser scan based CMM and a conventional touch-probe based CMM. On the left, the laser scanner (Nikon LC15Dx) projects a line laser and scans through the fabricated in one go. On the right, the touch probe needs to take multiple measurements for each hole, hence it takes much longer time to complete the measurement.
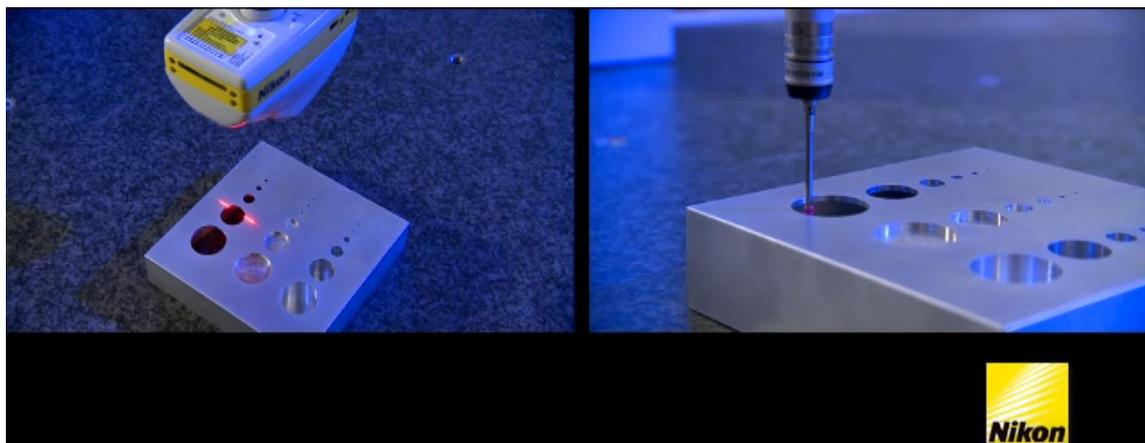
Figure 2.8: A snapshot from the promotional video of the Nikon's product:
Nikon LC15Dx CMM laser scanner (Image courtesy of Nikon Metrology [20])

Case study 2: Production in-line profilometer

As discussed previously in section 2.3, it is common to see the integration of the laser line scanner into a production line. One example is the product of Keyence, LJ-V7000 which is the first 3D inspection system which utilizes a blue laser. Due to the diffraction, a blue laser is able to produce a sharper line beam comparing to a conventional red laser. Figure 2.9 shows a series of keys passes through the laser scanner, and the measurement result not only shows the shape of the key, but also the fine engraving on the top surface.
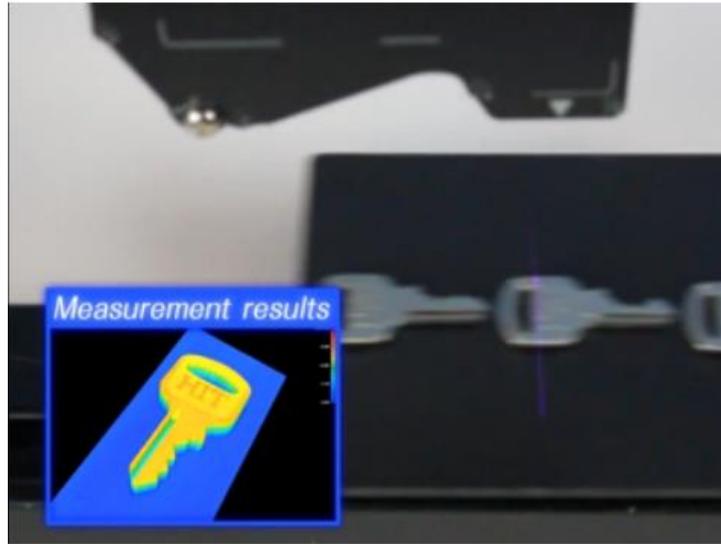
Figure 2.9: A snapshot from the promotional video of Keyence's product: Keyence LJ-VJ7000 Series. (Image courtesy of Keyence [21])

2.10 Conclusion

In conclusion, laser scanning is a reliable method for 3D imaging purpose, but in general the 3D image acquisition takes time, which may not suitable for some applications.

Chapter 3: Structured Light

Structured light technique is another type of 3D imaging method which relies on principle of triangulation. As discussed from the previous chapter, the major drawback of laser scanning method is that there has to be some scanning mechanisms in order acquire the 3D image, which may be difficult to implement in some applications. Furthermore, due to the scanning process, the 3D image acquisition time is considerably long. The idea of structured light method is to project a single or multiple encoded light patterns onto the object instead of a single laser line projection. By eliminating the need of scanning process, the aim of structured light technique is to acquire the 3D image with a much shorter time scale, and possibly real time performance.

Instead of projecting a single laser line, the structured light technique projects stripe patterns (usually black and white stripes) onto the object. The series of line segments allows us to compute the 3D points by using the principle of triangulation as discussed in the previous chapter. However, since the line segments are essentially identical, there is a problem to identify which distorted stripe corresponds to which projected stripe.

In order to identify the stripes (or other patterns), it is necessary to implement certain kind of coding strategy so that each pixel is distinguishable. This problem is similar to the correspondence problem as discussed earlier in the stereo vision. Hence, the concept of structured light technique can also be considered as "active stereo vision". While the passive stereo vision consists of two camera modules, the structured light replaces one of the cameras to a projector. There exists different coding strategies, some methods require multiple images, while some methods only need a single image to solve for the correspondence problem. Here we will discuss different kinds of techniques, and their strengths and limitations.

3.1 Time-multiplexing scheme

One of the most commonly used techniques is to project a series of light patterns successively, which is known as time-multiplexing coding. As such, the position of each pixel is encoded in the time domain. Here we will discuss different approaches under this category.

<u>Binary coding</u>

Binary coding is the simplest form of coding technique which projects black and white stripes to form a sequence of projection patterns. Each point on the surface of the object will be encoded with a unique binary code. This concept was first proposed by Posdamer and Altschuler in the early 1980s [22]. The idea is to project N patterns that can code 2^N stripes. Figure 3.1 shows a simplified 5-bit projection pattern [23]. This sequence of projected patterns allows us to code the scene with 32 unique line segments. For example, a particular line segment of the scene can be coded as "01101" for the line segment highlighted in the figure 3.1, where 0 corresponds to the black stripe and 1 corresponds to the white stripe.

With this encoding method, the 3D image can be constructed using the principle of triangulation for all 32 points along each horizontal line.
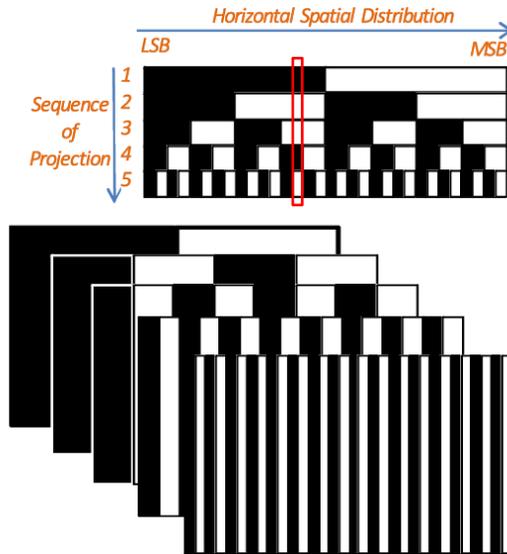
Figure 3.1: A simplified 5-bit projection pattern [23].

Gray coding

After a while, Inokuchi et al. introduced the Gray code [24], the pattern is very similar to the binary code, it still consists of black and white stripe, but the stripes are offset as shown in figure 3.2. As highlighted in the figure 3.2, the same particular line segment as in figure 3.1 is now represented as "01011". For the Gray code scheme, the number of requrired projections (images) and the coded line segments is the same as in the binary coding scheme, but it offers better performance. The reason is because the successive numbers of the Gray code vary exactly in one bit, when there is wrong decoding, it introduces only a misplacement of at most one resolution unit [25]. For this reason, most of the structured light systems use Gray code instead of the binary code.
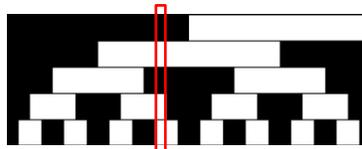
Figure 3.2: A 5-bit equivalent Gray code.

N-ary coding

While binary and gray coding scheme works well in general, it could potentially take large amount of images for the 3D image acquisition, which is time consuming. The question we might ask is: Can we reduce the number of images? The answer is the N-ary coding scheme, it similarly projects stripes pattern, but instead of 2 level (black or white), it contains N distinct levels of intensity. Figure 3.3 shows an example of stripe pattern with N-ary coding

[23]. For M patterns projection, it will be able to code N^M number of unique stripes. For instance, if M = 5, N = 4, the total number of unique stripes is 1024. For binary coding scheme, it requires 10 projections to produce 1024 unique stripes.
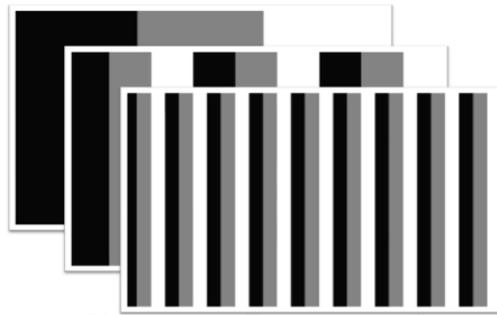
Figure 3.3: N-ary coding stripes with N = 3, M = 3 [23].

The reduction of the images leads to faster 3D image reconstruction. However, there is a drawback with such approach: the contrast of the pattern is reduced. This is not an issue for a uniformly white object (e.g. statue), because it is still relatively easy to extract the edges of the stripes so that the coded stripes can still be confidently detected. If we have an object that its surface does not have uniform reflectance properties, it will be difficult to detect the coded stripes correctly. For example, if a dark grey stripe is landed on a bright area or a bright grey stripe is landed on a dark area, there will be error in identifying the code of the stripe. For binary coding scheme, this problem is less of a concern because the high contrast patterns allow us to detect the edges of the stripes much more confidently.

From binary to N-ary coding, we can think of it as a trade-off, while we gain from the decreased number of the projected patterns (images), the signal to noise ratio of the system will be reduced.

Phase shifting

Phase shifting is another common type of pattern projection method for obtaining 3D image. One key advantage of phase shifting technique is it offers higher range resolution than binary coding [26]. The idea is by projecting sinusoidal pattern on the object, we observe the distorted (i.e. locally shifted) pattern due to the local height variation. The sinusoidal pattern is shifted linearly with the local height, in other words, we can obtain the depth information by recovering the phase information. One method to obtain sinusoidal pattern is by defocusing the projection of black and white stripes, the defocused image will be approximately sinusoidal.

To recover the phase, the common method is to project multiple phase-shifted sinusoidal patterns at different time frame as shown in figure 3.4. We can then compute the phase angle of a pixel based on the recorded intensities.
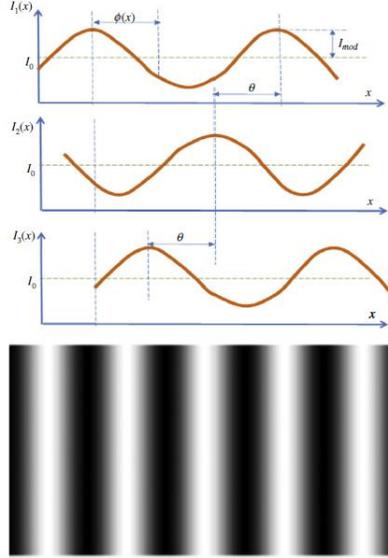
Figure 3.4: An example for phase shifted sinusoidal fringe patterns [23].

Here we will examine the 3-step phase shifted projection algorithm:

$$I_1 = I_0 + A \cos(\phi) \tag{3.1}$$

$$I_2 = I_0 + A \cos\left(\phi + \frac{2\pi}{3}\right) \tag{3.2}$$

$$I_3 = I_0 + A \cos\left(\phi - \frac{2\pi}{3}\right) \tag{3.3}$$

where I_0 is the background intensity and A is the amplitude of the sinusoidal pattern. Here each sinusoidal pattern is shifted equally by 120°. By trigonometry, we can find out the phase:

$$\phi = \tan^{-1}\left(\frac{\sqrt{3}(I_3 - I_2)}{2I_1 - I_2 - I_3}\right) \tag{3.4}$$

For 4-step phase shifting, with each phase difference equals to 90°, by similar construction, we can find out the phase as follow:

$$\phi = \tan^{-1}\left(\frac{I_4 - I_2}{I_1 - I_3}\right) \tag{3.5}$$

In either case, we can obtain the phase information, but there is issue with the phase ambiguity due to the periodic nature of the phase. This phenomenon is known as wrapped phase, that is the phase repeats itself every 2π (one period length). The discontinuity of the phase map can be removed by adding or subtracting multiples of 2π as follow:

$$\phi' = \phi \pm 2n\pi \tag{3.6}$$

This process is called the phase unwrapping, but there is a potential problem with this process and there are multiple solutions satisfies the equation 3.6. If the surface of the object is smooth, the phase unwrapping process is rather simple: we add or subtract 2π so that the unwrapped phase is a smooth function. However, this simple implementation does not work well for surface with abrupt discontinuities, which is the major drawback of the phase-shifting method.

Line shifting

As mentioned previously, the Gray code sequential coding is a popular coding strategy that allows us to distinguish the stripe pattern, but the discrete nature of such patterns limits the range resolution. On the other hand, phase shifting method exploits higher spatial resolution since it project periodic intensity patterns several times by shifting them in every projection, but the periodic nature of the patterns lead to phase ambiguity [26]. Jens Gühring et al. combined these two methods to bring the advantages of both strategies, i.e. the unambiguity of the stripe patterns of the Gray coding and high resolution of the phase shifting method [27]. Subsequently, Jens Gühring introduced a slightly different approach, which is the line shifting method, the concept is shown in figure 3.5 [25].
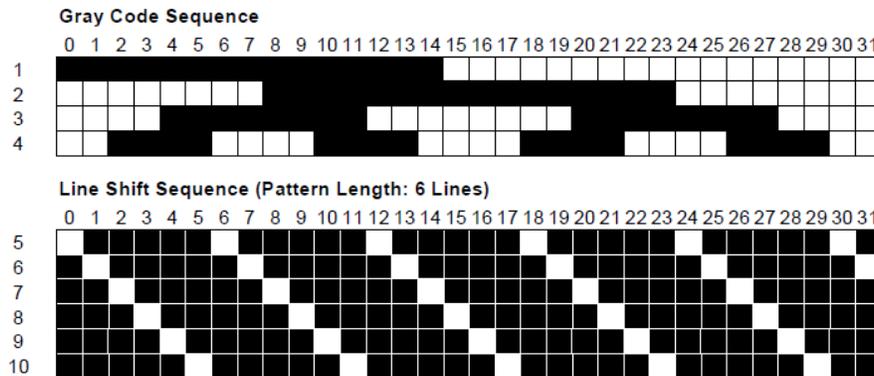
Figure 3.5: Line shifting method: Gray code sequence + line shift sequence [25].

Line shifting method is currently one of the best structured light techniques in terms of accuracy, but it requires significantly more images than other options. Joaquim Salvi et al. from University of Girona had provided a quantitative comparison between different pattern codification strategies in structured light system [26].

3.2 Direct codification

The previous time multiplexing methods require taking multiple images, hence it is difficult to achieve real time 3D imaging. To achieve real time 3D image acquisition, it is essential to reduce the number of images. Direct codification scheme attempts to encode every pixel by its own intensity or color.

Grey level coding scheme

The intensity of a pixel, which is represented by the grey level, contains the information that can be potentially used to identify the line stripes. However, as we might expect, this strategy is not robust as the scene is not necessarily uniformly white, any variation of the surface reflectance will lead to false information. Carrihill and Hummel demonstrated 3D imaging with this approach using 2 images [28], but the accuracy is significantly worse than alternative methods due to high sensitivity to noise and non-linearity of the projector device.

Color coding scheme

Another common type of coding scheme is by using color stripes, where each pixel in the image can be identified by its color. Tajima and Iwakawa firstly demonstrated 3D imaging with rainbow pattern generated with white light passing through a crystal prism [29]. Apparently, such approach works only for neutral color object.

Sato proposed using a single color stripe pattern instead of multiple binary or N-ary stripe patterns to identify the stripes [30]. As we know, if the object surface distorts the projected color as perceived by the camera, it will lead to erroneous results in the measurement. The idea was to project 3 shifted patterns so that the object color can be canceled, resulting in better reconstruction at the cost of 2 more images.

As discussed previously in the phase shifting scheme, it was required to project at least 3 images of phase shifted sinusoidal pattern at different time frame. Wust and Capson proposed the color sinusoidal fringe pattern in which red, green, blue fringes are simultaneously projected and phase shifted to one another [31]. Figure 3.6 shows the concept of color fringes pattern and its appearance when it is projected onto a ball.
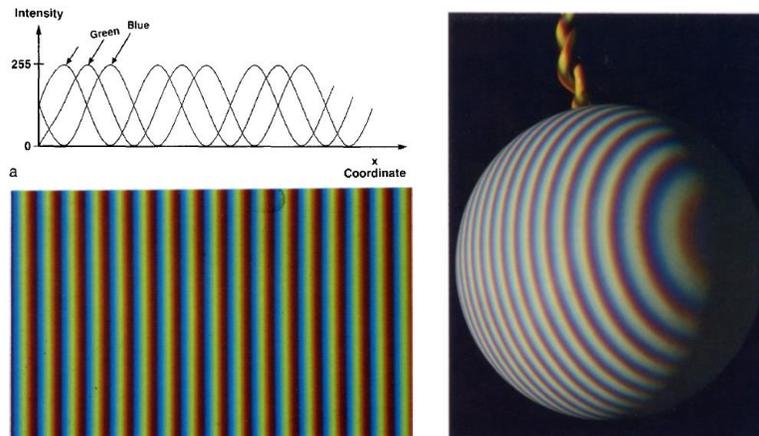
Figure 3.6: (Left) color fringe pattern
(Right) Photograph of color fringe pattern projected onto a ball [31].

3.3 Spatial neighborhood codification

The techniques in this category encode the information contained in a neighborhood (also known as window) with unique patterns. Such methods encode the information spatially instead of temporally. Therefore, it is possible to reconstruct the 3D image with a single shot instead of multiple images, making real time acquisition possible. However, it could be difficult to decode the images since the spatial neighborhood cannot always be recovered that errors of the measurement may arise. Here we will examine some of the methods.

<u>Stripe indexing</u>

The concept of stripe indexing is to arrange the stripe pattern is such a way that a unique stripe pattern can be distinguished by the algorithm. If more than 2 grey levels are used, the repetitive stripes patterns as discussed previously can be rearranged so that they form a unique sequence within the whole frame. For example, if 3 grey levels are used (B – Black, G – Gray, W – White), a pattern can be designed as follow (also illustrated in figure 3.7):

$$BWGWBGWGBGWBGBWBGW$$

This method was proposed by Durdle et al. [32]. The algorithm to decode the projected stripe pattern, which is distorted by the object, is to find the correlation peak between the image row and the template of the projected pattern for a certain size of matching window. For example of a 3 code sequence matching window, every pixel among a row will be matched with the set of sequences BWG, WGW, GWB, WBG, etc, and there will be one set of sequence which gives the highest correlation.
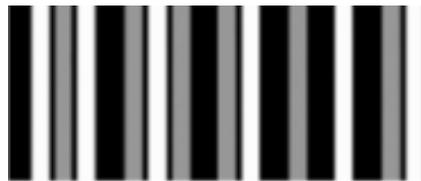
Figure 3.7: An example of stripe indexing pattern with 3 grey levels.

Another type of stripe indexing method is to use segmented line patterns instead of a conventional line patterns. For example, Maruyama and Abe [33] designed a pattern with vertical line slits that have small cuts at random locations. As shown in figure 3.8, the length of the line segment and the location of the cut form unique sets of information that can be served as matching purpose. However, this method can only be used for object with a smooth and continuous surface so that the distortion of the line segment pattern is not severe, it will be very difficult to recover the unique segment pattern when there is sharp discontinuity on the object surface.

Figure 3.8: An example of stripe indexing with segmented line pattern [23].

De Brüijn sequence

A De Brüijn sequence is a sequence with set of alphabets arranged in a form that each sub-sequence is unique. More specifically, a De Brüijn sequence of order m over an alphabet of n symbols is a cyclic string of length n^m that contains each substring of length m exactly once [26]. A simple example of a De Brüijn circle with m = 3, n =2 with alphabet {0,1} is shown in figure 3.9. In either clockwise or counterclockwise direction, the 3-digit patterns 000, 001, 010, 011, 100, 101, 110, 111 appear exactly once. This unique feature of the De Brüijn sequence has been often used in structured light patterns as the unique local patterns simplify the decoding task.
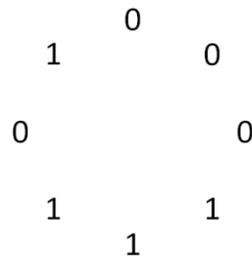
Figure 3.9: A simple example of De Brüijn sequence with m = 3, n = 2

Zhang et al. [34] developed a structured light pattern based on 125 vertical slits encoded with a De Brüijn sequence of 5 colors, and window size of 3 slits. The color indexed stripe were constructed in such a way that all neighboring stripes have different colors, this is shown in figure 3.10.
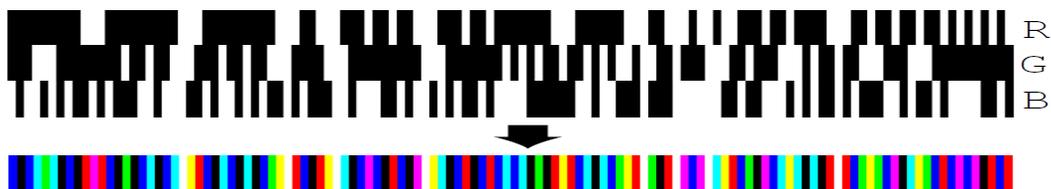
Figure 3.10: Coding scheme with De Brüijn sequence: Combination of binary R, G, B patterns produce a sequence with unique consecutive color transition (m = 5, n = 3) [34]

Using the De Brüijn sequence, the group was able to use only a single image to reconstruct the 3D image in real time. In addition, the group also extended their technique using the classical time-multiplexing paradigm. By projecting 7 time-shifted patterns, the slits can be located with subpixel accuracy, leading to high resolution 3D scanning for static scene. The difference between the single shot and 7 time-shifted patterns scheme are demonstrated in figure 3.11.
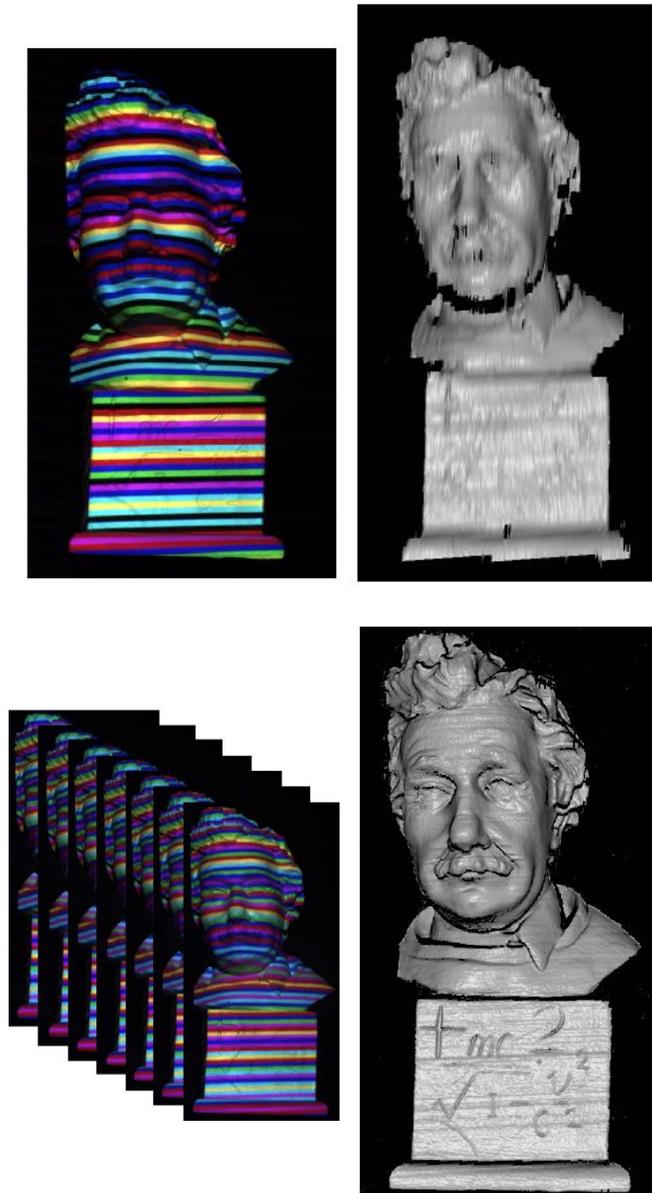
Figure 3.11: 3D imaging of Einstein bust. (Top) Single shot reconstruction (Bottom) 7 time-shifted patterns reconstruction [34].

2D Grid indexing patterns

The previous stripe indexing approach encodes the pattern in either column direction or row direction. 2D Grid indexing method can be considered as bi-dimensional extension of the stripe indexing using the De Brüijn sequence. The projected pattern is an array of symbols with each small window of size (w x h units) appears only once.

One example is projecting mini-patterns coded with shape primitives as shown in figure 3.12. This approach was proposed by Griffin et al. [35]. Each 3 x 3 window consists of shape primitives that are arranged uniquely. With these uniquely defined windows, one can match the array of projected patterns, hence reconstructing the 3D image. However, this method requires coding the patterns spatially that the effective sampling points are reduced, leading to inevitable loss of 2D spatial resolution.
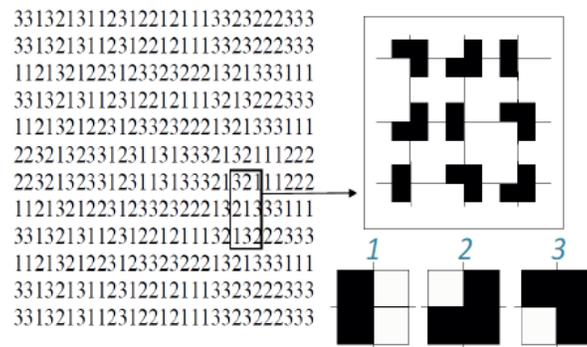
Figure 3.12: Projection shape primitives as codewords for grid indexing [23].

Likewise in the previous approach, using color information is an effective way of coding the projected patterns. The key idea is to generate pseudo-random arrays of color dots with unique sub windows. Morano et al. [36] presented a method of generating such pseudo-random arrays based on brute-force approach. Figure 3.13 shows an example of 2D array of color coded dots, each 3 x 3 window has unique color dots pattern for correspondence purpose.
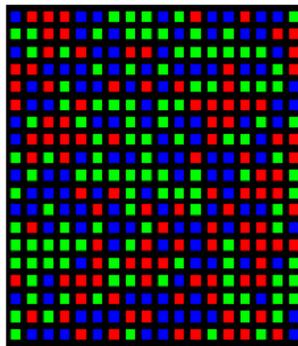
Figure 3.13: An example of 2D array of color coded dots [23].

3.4 Performance evaluation of 3D imaging with structured light

We have discussed a wide variety of structured light method for creating 3D images. The question we may ask is: "What is the best structured light method?" The answer to this question depends on the application. In other words, is the real time performance or the range accuracy more important? Figure 3.14 depicts the two extreme cases. The left extreme corresponds to the single laser line scanning method as discussed in chapter 2, the accuracy is high (no ambiguity problem) but the process is slow (scanning required). The right extreme corresponds to single shot structured light approach, the process is fast but the accuracy is compromised. In addition, single shot method tends to be much more fragile to the ambient light, non-uniform and multi-colored scenes may pose problems in reconstruction as well. With multiple images with multiple stripes pattern, a balance of high accuracy and high speed 3D imaging system can be realized.

Single Stripe ◄	Multi Stripes Multi Images	► Single Shot
- Scanning required - Static scene and time consuming. - High accuracy - Robust	- Scanning not required - Static scene or moving scene depending on number of images and camera speed	- Correspondence problem - Moving scene, real time performance. - Low accuracy - Fragile
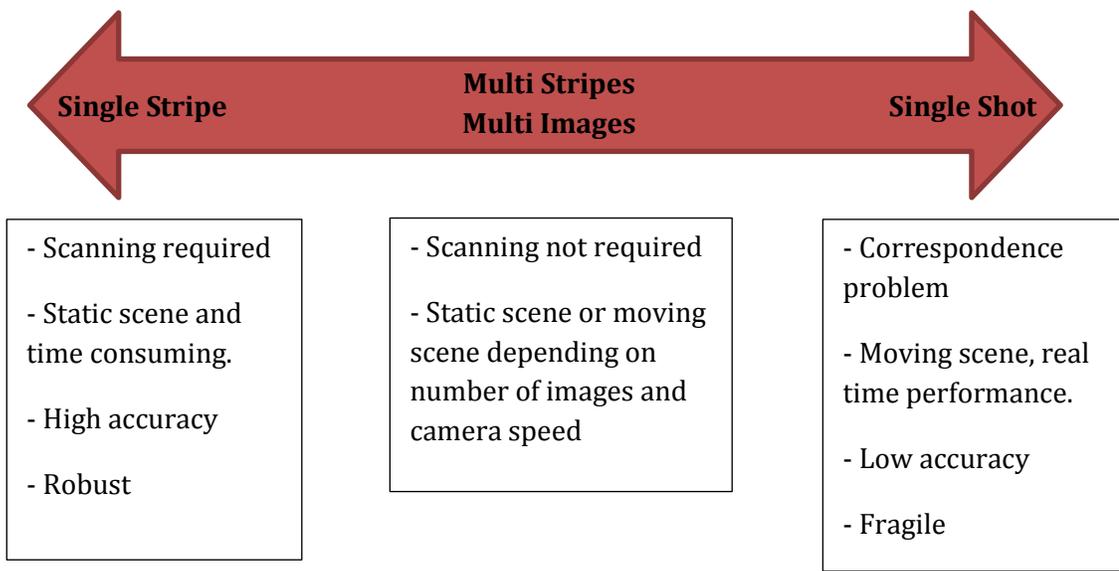
Figure 3.14: 3D imaging with triangulation method.

3.5 Case study

Case study 1: Motorcycle 3D imaging

One relevant application of structured light technology is to obtain 3D model of large object, such as automotive parts. The advantage of structure light technology is that the hardware setup is considered much more compact comparing to laser scanning technology.

For example, SMARTTECH [36] offers industrial 3D imaging solutions to specialized motorcycle accessories producer. Conventional time multiplexing structured light method is used in this case. Binary stripe patterns of different frequencies are projected onto the motorcycle consecutively. Figure 3.15 shows one of the captured frames with projected binary stripe pattern.

Figure 3.15: (Left) 3D imaging of motorcycle parts with structured light method.
(Right) Rendered 3D model after manual touch up [36].

Scanned 3D model is then imported to Solidworks. The scanned 3D model contains holes, it needs to be manually touched up with Solidworks, which final result is shown in figure 3.15.

Case study 2: Real time 3D imaging of music video

One direct application of real time 3D imaging is the entertainment purpose, such as gaming and video film industry. For example, the music video of Radiohead's music video "House of Cards" was filmed with real time 3D imaging technology developed by the XYZT Lab at Purdue University [37]. The technology used high speed camera which was capable of capturing images at a frame rate of 120 frames per second. The high speed camera was synchronized with the projector which projected 3 phase-shifted fringes consecutively, resulting in real time 3D imaging at 40 frames per second [38].
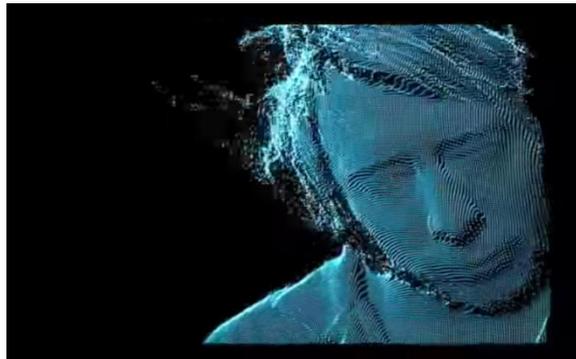
Figure 3.16: Screenshot of music video "Radiohead – House of Cards" [39].

3.6 Conclusion

In conclusion, structured light method offers faster 3D image acquisition comparing to laser scanning method. Structured light technology has been actively researched over the past decades. The main task is to design special patterns that can be decoded effectively, and there exists many creative solutions.

Chapter 4: Depth From Focus

      Depth from focus is perhaps the most intuitive method to obtain object's height information by focusing an object at different planes. Conceptually, a 3D image can be obtained by simply stacking multiple 2D images which focus an object at different levels. From geometrical optics, it's trivial that the depth of focus decreases as the magnification increases. Therefore, here we will discuss the depth from focus 3D imaging technique in the regime of microscope imaging system.

In addition, we will also discuss the working principle of confocal microscopy. Confocal microscope is a special type of microscope which is able to capture high quality 3D image with depth from focus technique as well.

4.1 Depth of focus

The definition of depth of focus can be controversial and often confused with depth of field. Here we define depth of focus in object space, and image space respectively, as illustrated in figure 4.1.
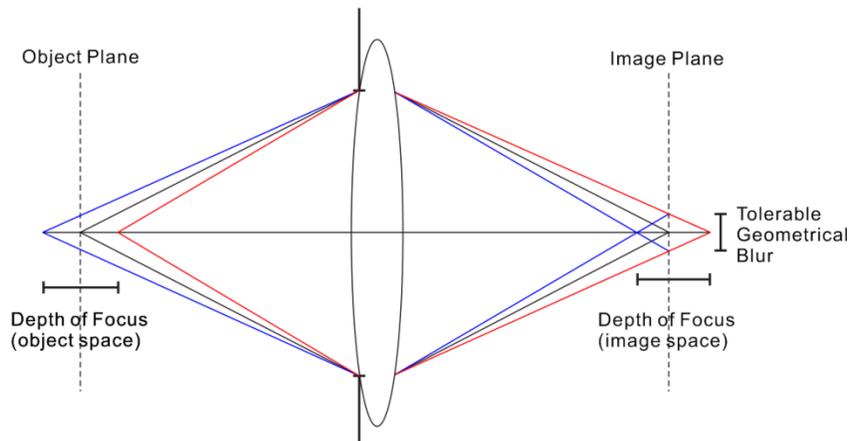
Figure 4.1: A simplified imaging system illustrating
the concept of depth of focus.

For depth from focus technique, we are interested in the image blur due to having object points at different object planes, hence the depth of focus is defined in the object space in this context. Depth of focus is the region in which the obtained image remains sharp, or within a tolerable amount of blur. By having thin depth of focus, it becomes possible to determine the object surface's depth profile by analyzing the associated image blur.

One of the most well-known formulas describing the depth of focus was provided by Born & Wolf [40] as follow:

$$\Delta z = \frac{n\lambda}{2NA^2} \tag{4.1}$$

where n is the refractive index of the medium, and NA is the numerical aperture of the objective lens, which is defined as $n \sin U$.

Unfortunately, the formula in equation 4.1 is not accurate for high numerical aperture, diffraction limited optical system, which was pointed out by I.T. Young et al. [41]. They provided a more accurate formula of depth of focus as follow:

$$\Delta z = \frac{\lambda}{4n \left(1 - \sqrt{1 - \left(\frac{NA}{n}\right)^2} \right)} \tag{4.2}$$

They tested the formula above experimentally and proved that it is more accurate than equation 4.1. Here we derive the depth of focus formula starting from the wave aberration theory, in which the defocus aberration can be evaluated by the optical path difference from a reference sphere as indicated in figure 2.
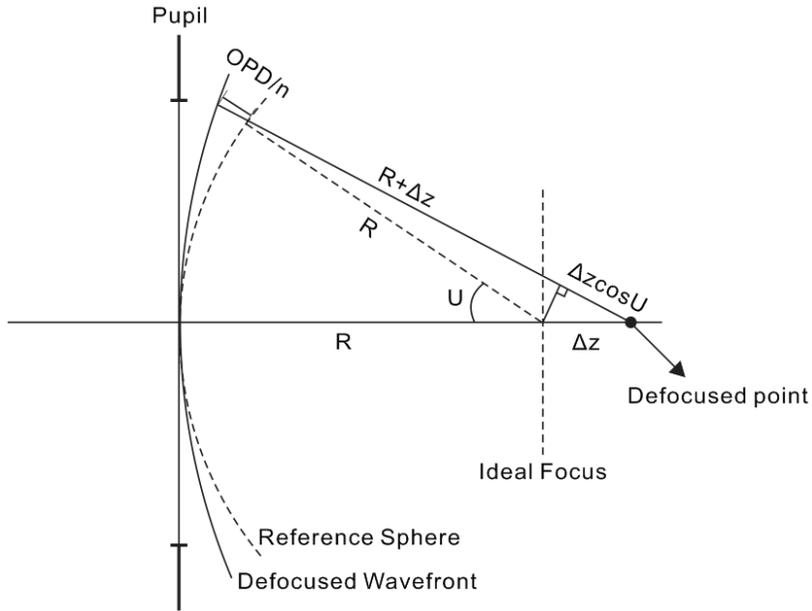
Figure 4.2: Defocused imaging system as described
with optical path difference.

Figure 4.2 describes a diffraction limited imaging (free from geometrical aberration) system which images on-axis light ray from infinity to a focused point. At the ideal focus point, the image is free from aberration, and then we can construct a reference sphere which converges a spherical wavefront to this point. If the observation point is moved away from this ideal focus point, defocus aberration is introduced to the imaging system, and the imaging system is no longer perfect, but the defocused wavefront is still spherical.

The degradation of the imaging system can be described by the optical path difference of the defocused wavefront with respect to the reference sphere. The optical path difference can be expressed as:

$$\frac{OPD}{n} = (R + \Delta z) - R - \Delta z \cos U \tag{4.3}$$

The $\cos U$ can be expressed in terms of $\sin U$ (as subsequently related by the NA),

$$\cos U = \sqrt{1 - \sin^2 U} \tag{4.4}$$

The optical path difference becomes:

$$OPD = n\Delta z \left(1 - \sqrt{1 - \left(\frac{NA}{n}\right)^2} \right) \tag{4.5}$$

Using Rayleigh's quarter-wave criterion, which says that if an aberrated wavefront has optical path difference no more than a quarter wavelength ($\lambda/4$), the imaging system is still considered diffraction limited. Hence we have:

$$\Delta z = \frac{\lambda}{4n \left(1 - \sqrt{1 - \left(\frac{NA}{n}\right)^2} \right)} \tag{4.6}$$

which is identical to the equation 4.2. For low numerical aperture system, one can do the following approximation:

$$\sqrt{1 - \left(\frac{NA}{n}\right)^2} \approx 1 - \frac{1}{2}\left(\frac{NA}{n}\right)^2 \tag{4.7}$$

which leads the equation 4.7 to the equation 4.1.

The derivation here assumes light from infinity focus to a point with depth of focus Δz in the image space. In a microscope system, the optical system is reversed, and the depth of focus Δz is defined in the object space. We expect the depth of focus decreases dramatically with increasing numerical aperture, here we tabulate the depth of focus for some common objective lens, assuming $n = 1, \lambda = 550$nm. Note that the numerical aperture for a given magnification objective lens is not necessary the same as stated in table 1, the numerical aperture can be larger or smaller depends on the specification of a particular objective lens.

Magnification	4X	10X	20X	40X	60X
NA	0.1	0.25	0.4	0.65	0.85
Δz (μm)	27.43	4.33	1.65	0.57	0.29

Table 4.1: Depth of focus for some common objective lens.

As we can see from table 4.1, for high magnification objective lens with large numerical aperture, the depth of focus is extremely short. Therefore, we can expect sub-micron axial resolution for depth from focus microscope system.

Lastly, equation 4.6 estimates the depth of focus solely based on the diffraction calculation in the object space. In an actual depth from focus microscope system, where the image is digitalized by the imaging sensor, the actual depth of focus may also depend on the pixel size if the imaging system is limited by the pixelization rather than the diffraction.

4.2 Working principle of depth from focus microscopy

The working principle of depth from focus microscopy relies on the limited depth of focus as discussed earlier. Since the depth of focus is limited, if we scan a sample vertically and determine best focus position, a depth map can then be generated. In a typical setup, the object is to be placed on a computer controlled vertical stage (e.g. piezo actuator), and then 2D image will be acquired at each vertical position.

To obtain the depth information for each point of the object surface, the simplest method is to determine the "focus measure" for each point in the image. The contrast of the image is highest when it is in the best focus, by calculating the standard deviation for a small region around a point of interest (illustrated in figure 4.2), the focus measure can be determined.
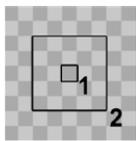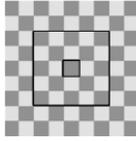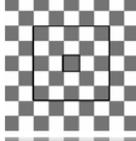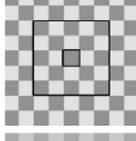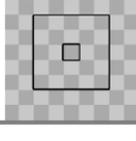
Figure 4.2: Region 1 as the point of interest, whereas region 2 of 5 x 5 neighborhoods is used to calculate the standard deviation [42].

After the focus measures are obtained for various planes, the next step is to find the peak of the focus measure, which yields the depth information. The peak position of the focus measure does not necessarily coincide one of the measured points, hence it may be necessary to do a polynomial curve fitting in order to find the maximum point of the focus measure. Figure 4.3 shows a typical focus measurement with fitted curve. From the interpolated polynomial curve, the maximum point of the focus measure can be determined more accurately.
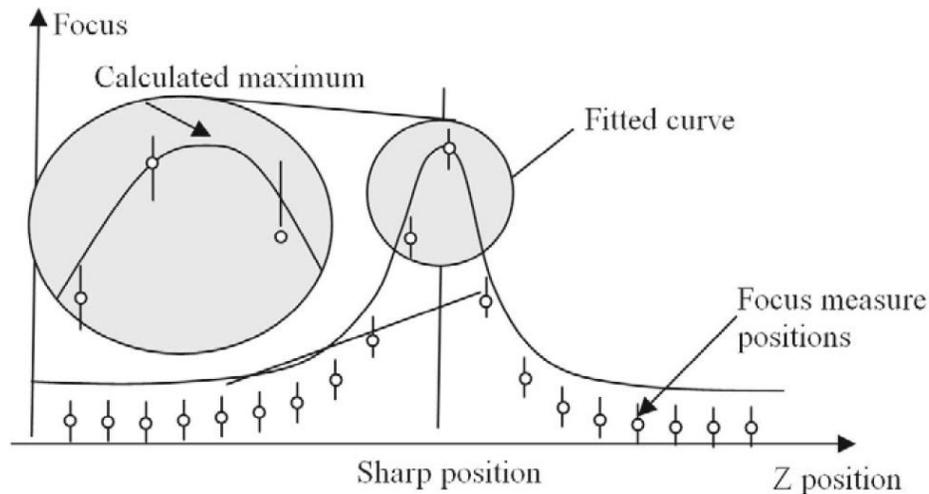
Figure 4.3: Finding maximum focus measure by curve fitting [42].

Once the maximum focus measure is determined, it is then translated into height profile for every pixel of the 2D image, and a 3D image can be reconstructed. However, one limitation with depth from focus technique is that it is not able to work with smooth surface (e.g. a polished sphere), because there is no feature for contrast detection. Depth from focus technique works by contrast detection, hence it is especially useful for obtaining 3D image of a rough surface.

4.3 Confocal microscopy

Confocal microscopy is another variant of microscopic technique which is widely used in biological imaging application. Although its working principle in generating 2D image is different from conventional microcopy, it is able to measure the depth profile with similar approach as discussed previously.

Confocal microscopy was firstly invented and patented by Marvin Minsky in 1957 [43]. It was aimed to overcome some limitations in a conventional fluorescence microscope. When a biological sample is illuminated by an excitation light source, the fluorophores absorb the light and emit light of longer wavelengths. By using appropriate excitation and emission filters, one can obtain fluorescence image for a biological sample. However, in conventional fluorescence microscopy, as the entire biological sample is illuminated evenly with the excitation light source, the resulting fluorescence signals detected by the camera include

large portions of unfocused background which reduce the contrast of the fluorescence image significantly.

As shown in figure 4.4, a confocal microscope uses a point illumination and a pinhole in an optical conjugate plane (hence the term "confocal") in front of the detector to eliminate out of focus signal.
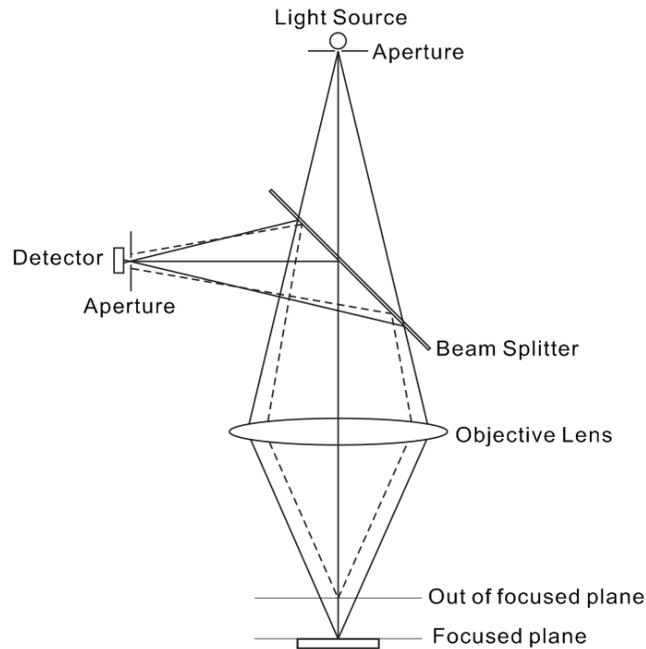
Figure 4.4: Schematic of a confocal microscope.
(Optical filters not shown)

By suppressing the out of focus signals, the confocal microscope is able to generate high quality image. However, it is a point imaging technique, hence it is required do a horizontal scanning in the x-y plane across the field of view of interest. A common configuration to implement the horizontal scanning is by spinning a plate with multiple pinholes. The spinning pinhole is also known as the Nipkow disk. This configuration allows taking multiple point images at one time, and the spinning pinholes mechanism scans through the entire 2D plane rapidly. However, the throughput of traditional Nipkow disk is only about 1% of the light shining onto the disk. Yokogawa Electric came out with a design using microlens array to focus light into individual pinhole, resulting in much better transmission efficiency [44]. Figure 4.5 shows the working principle with Yokogawa spinning Nipkow disk configuration for confocal microscopy.

Lastly, using vertical scanning method, multiple "slices" of 2D images can be obtained at different depth within a sample, this process is also known as optical sectioning. 3D image can be reconstructed by stacking these 2D images, this is illustrated in figure 4.6. The algorithm to stack or merge the 2D images can be simply registering the maximum pixel intensity among the whole set of images. Alternatively, it can be done by calculating the

standard deviation over a small region of image, similar to the technique as discussed in section 4.2.
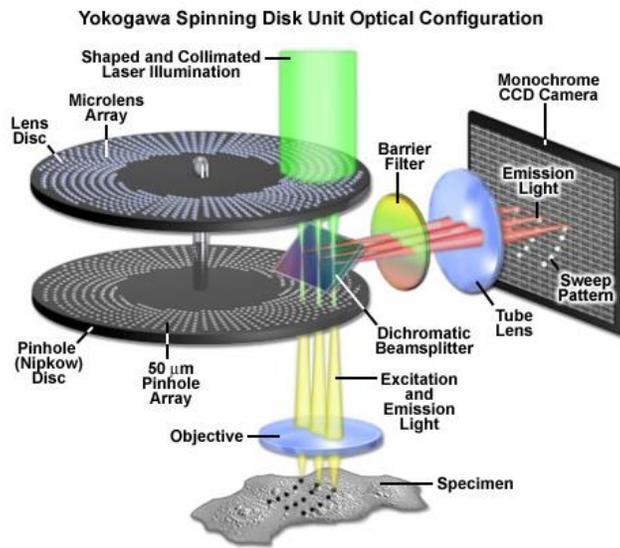
Figure 4.5: Schematic of a confocal microscope with Yokogawa spinning disk configuration. (Image courtesy of Zeiss [45])
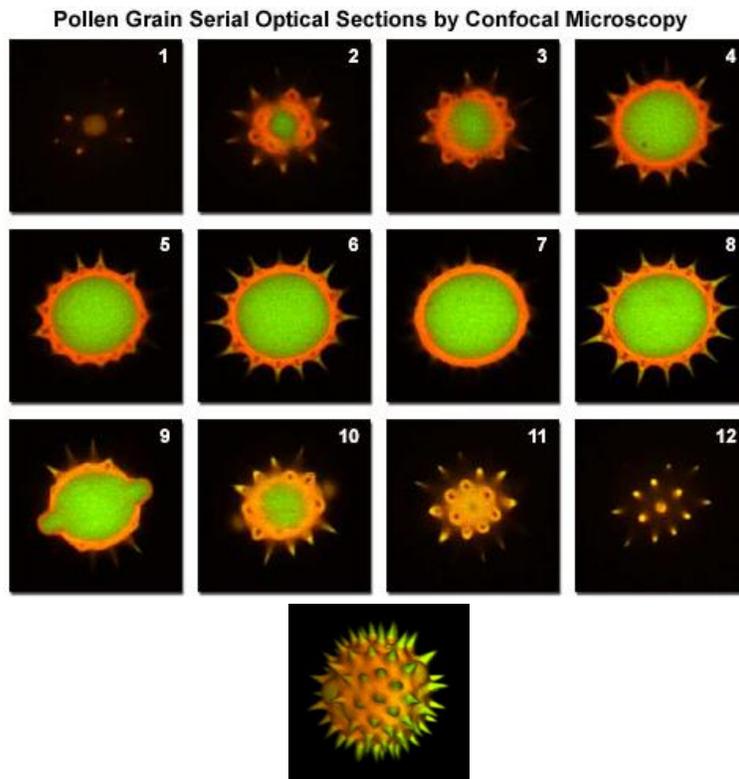
Figure4.6: (Top) Optical sectioning of a pollen grain with confocal microscopy. (Bottom) Reconstructed 3D structure. (Image courtesy of Olympus [46])

4.4 Chromatic Confocal Microscopy

Chromatic confocal microscopy is another interesting configuration of confocal microscopy which is worth studying. To obtain 3D image with confocal microscopy, it is necessary to scan the object vertically, such as putting the object on vertical translational stage, or moving the objective lens vertically. In any case, mechanical movement is required.

Chromatic confocal microscopy eliminates the need of mechanical scanning by using the effect of chromatic aberration. A conventional microscope objective lens is color corrected, that is the having red, green and blue light focused at the same plane. It is possible to design a microscope objective lens which is not color corrected, or purposely made to have strong chromatic aberration. A lens with chromatic aberration is illustrated in figure 4.7.
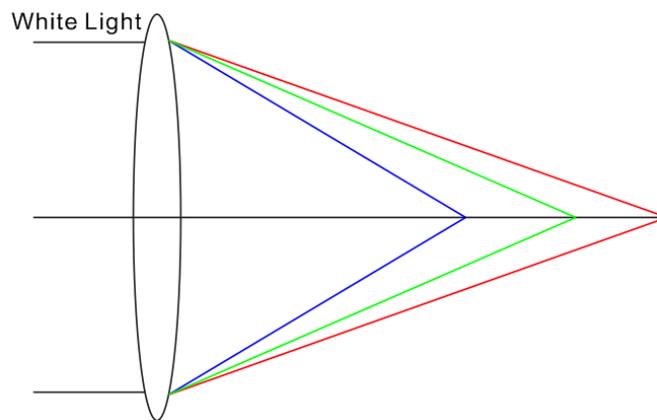
Figure 4.7: A lens with chromatic aberration, blue light is
focused closer than red light due to dispersion.

The working principle of chromatic confocal microscopy is shown in figure 4.8. In the figure, a white light source is used to illuminate the object. Due to chromatic aberration, the blue, green, and red lights are focused at different planes. In the example, point A is confocal with the aperture, so that the spectrometer detects mostly green light signal. When the confocal imaging system is shifted to point B and point C, the spectrometer detects mostly blue light signal and red light signal, respectively. This implies that, by analyzing the peak wavelength of the received signal using the spectrometer, depth or height information can be obtained. In other words, the z measurement is spectrally encoded. This is a good example of how depth information can be retrieved using alternative approach.

For obvious reason, chromatic confocal microscopy is not meant for fluorescence imaging as discussed earlier. However, it can be used for profile measurement, surface texture measurement, integrated circuit inspection, etc. In addition, color information of the object surface will not be recorded.
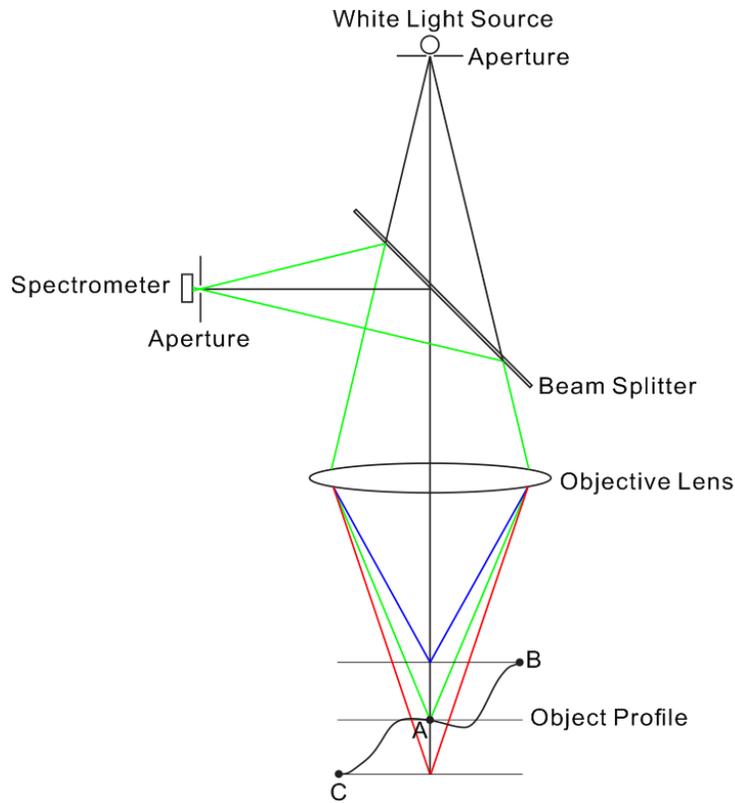
Figure 4.8: Schematic of a chromatic confocal microscope.

4.5 Case study

<u>Case study 1: Modern depth from focus digital microscope</u>

Constructing 3D image with depth from focus technique is commonly seen in modern digital microscope. In this case study, we look at 3D imaging capability from Keyence's digital microscope VHX-1000 series [47]. As shown in figure 4.9, a fully focused image and a complete 3D image of an uneven surface can be obtained by compiling images of different focus positions. The quality of the 3D image or measurement with depth from focus technique is not much dependent on the hardware, but the software algorithm for 3D reconstruction. Keyence claimed their new algorithm produces better 3D image from conventional approach.
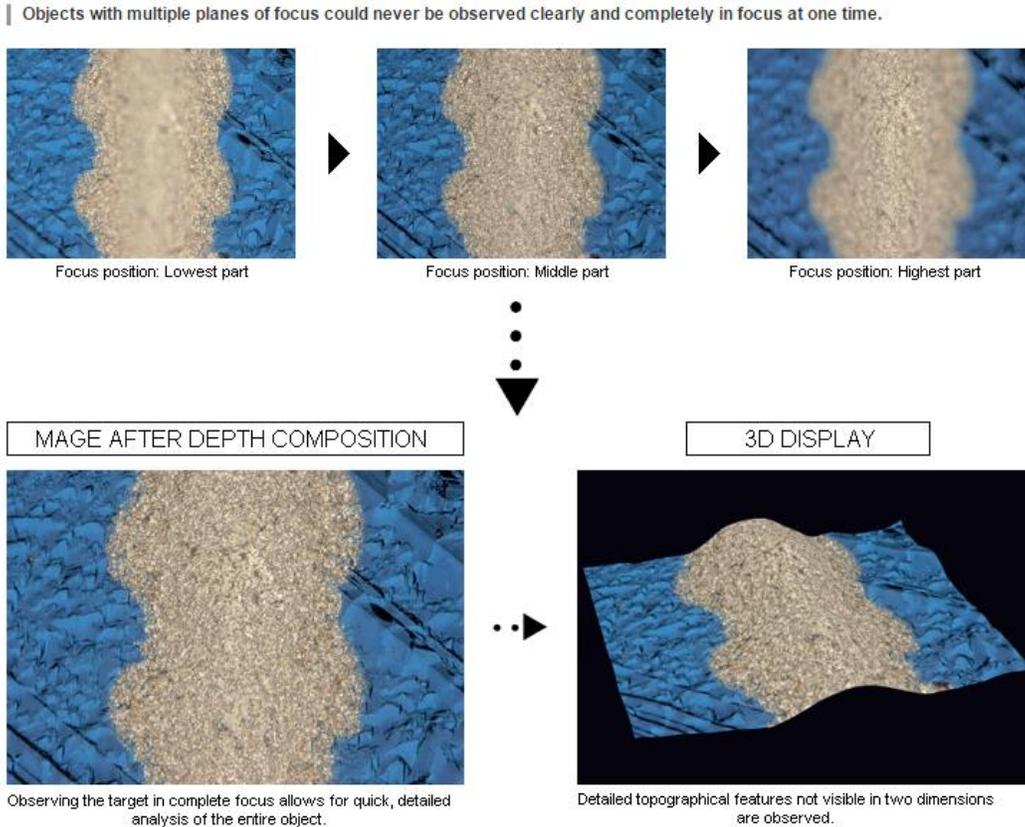
Figure 4.9: Reconstructing 3D image by stacking multiple 2D image.
(Image courtesy of Keyence [47])

Case study 2: Axial scanning with tunable lens in confocal microscopy

Confocal microscopy is very often used for imaging biological samples, owing to its excellent image quality by eliminating out of focused signal. However, one of the drawbacks is the image acquisition time is considerably long due to the nature of scanning operation. One method to improve the acquisition speed is to increase the vertical scanning rate is by using liquid tunable lens instead of mechanical stage to focus the object.

In this case study, we look at the liquid tunable lens technology from Optotune. Liquid tunable lens changes its focal length when voltage is applied to it. By integrating liquid tunable lens into a microscope system, the microscope system will be able to capture image at different object planes.

Figure 4.10 shows the working principle and its typical response time. A liquid tunable lens consists of an elastic polymer membrane, which shape is altered when a voice coil actuator exerts pressure on a liquid reservoir surrounding the aperture of the lens, thereby changing its focal length. The response time of such electrically tunable lens is less than 10ms, so that scanning rate in the order of 100Hz could be realized. On the other hand, for mechanical

translation with piezoelectric actuator, it scanning rate is limited by the mechanical inertia of the moving parts, scanning rate in the order of 10Hz is challenging [48].
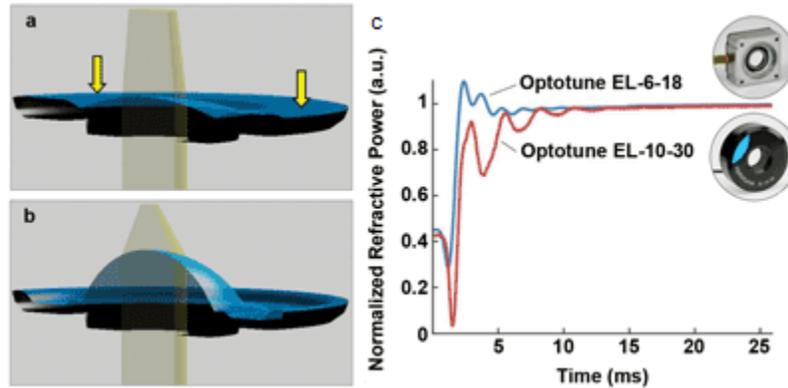
Fiugre 4.10: (Left) Working principle of an electrically tunable lens.
(Right) Typical response time of an electrically tunable lens.

For most 3D microscopy application, the working distance has to be increased or decreased to cover the whole thickness of the sample. For example, Optotune EL-10-30 has focal length ranging from 200mm to 100mm, an offset lens with negative focal length is required so that the total focal length can range from -667mm to 286mm, making it suitable for 3D microscopy application. Figure 4.11 shows a possible configuration for integrating electrically tunable lens into a conventional infinity corrected microscope system. This configuration uses 4f relay system with two achromats in such a way that the electrically tunable lens is placed at the conjugate pupil plane of the objective lens. Such configuration allows the microscope system working with variable working distance without changing the image magnification or field of view [48].
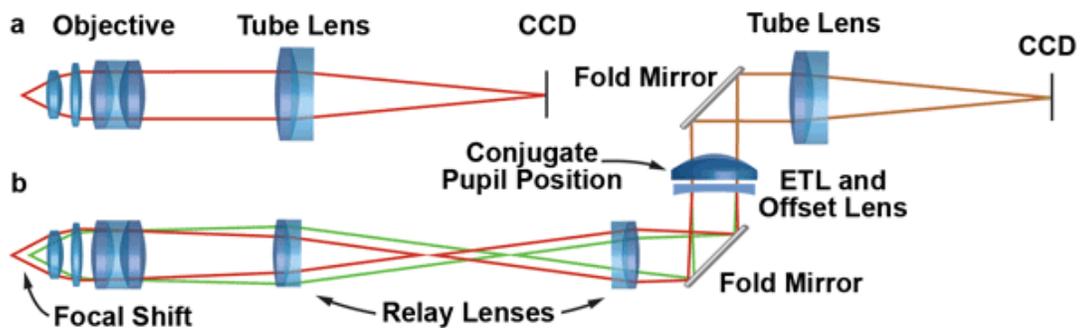
Figure 4.11: Integration of electrically tunable lens (ETL)
into an infinity corrected microscope system [48].

Optotune had demonstrated the integration of electrically tunable lens into a Yokogawa spinning disk confocal system [49]. Optical axial scanning with electrical tunable lens was also demonstrated by biomedical research group in Texas A&M University [50].

4.6 Conclusion

In conclusion, depth from focus is a simple technique for obtaining 3D image based on contrast detection, hence it is especially useful for measuring topology of a rough surface. The concept of stacking multiple 2D images is extensively used in confocal microscopy for generating high quality 3D image.

Chapter 5: White Light Interferometry

        White light interferometry is another type of 3D imaging method which is based on interference effect. The term white light refers to spectrally broadband light source, which may not necessarily be a true white light illumination. For example, a light emitting diode (LED) can be considered as white light source in this context. It is sometimes known as coherence scanning interferometry due the scanning operation which will be discussed later.

Consider a traditional interferometric setup where a laser light is split into two paths, upon recombination, constructive or destructive interference will be observed as long as the optical path difference between two paths are in integer of wavelength, or half integer of wavelength, respectively. We can observe repetition of bright and dark fringes repeating themselves as we vary the optical path difference. This is because a laser is a coherent light source with well-defined wavelength, in other words, the coherence length is much longer comparing to wavelength.

On the contrary, a white light source is not coherent, it has much shorter coherence length. In this case, we can still produce interference effect, but we can only observe the bright and dark fringes when the optical path difference is approximately equal. This property is the key principle of white light interferometry, i.e. measuring the height profile of a surface based on the occurrence of the interference effect. In this chapter, we will discuss the working principle starting from the interference effect of monochromatic light.

5.1 Interference of monochromatic light

We first consider interference of monochromatic light (e.g. laser) in the simplest form of interferometer, which is the Michelson interferometer as shown in figure 5.1.
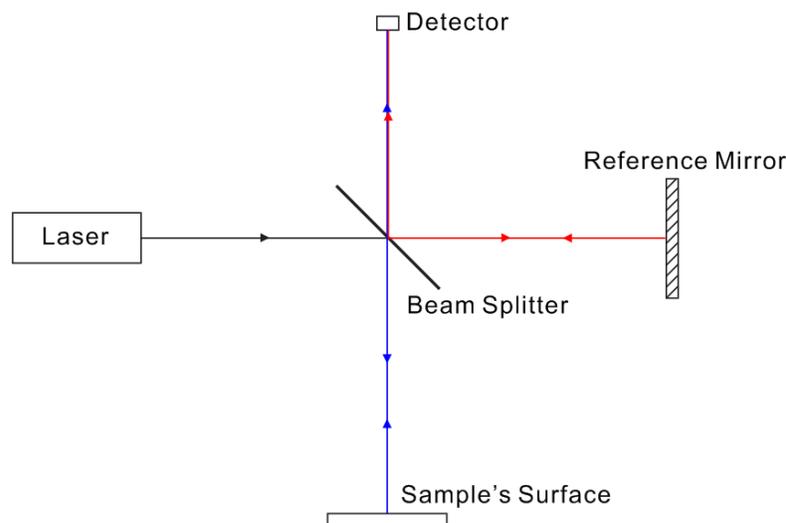
Figure 5.1: Michelson interferometer configuration.
(Color indication for split light path)

For collimated light, the electric field can be described by the plane wave:

$$u(t) = E_0 \cos(kx - \omega_0 t) \tag{5.1}$$

Upon reaching the beam splitter, it is equally split to two arms with same phase, one is the reference arm, and another one is the test arm which is used to probe the height information of the object of interest. Here the thickness of the 45° beam splitter is ignored. In reality the beam splitter has finite thickness, a 45° compensator can be inserted into one of the paths to compensate for the optical path difference, as well as the dispersion effect.

The length of test arm is to be set approximately equal to the reference arm, with path difference of Δl. Upon reflections from the reference mirror and test sample, the amplitudes of the electric field are reduced by factors of K_1 and K_2, respectively. These factors depends on the reflectively of the mirror and the surface of the sample.

The reflected light beams are then recombined at the beam splitter, assuming the beams are perfectly aligned, we are interfering one beam and another replica of itself, but with time delay due to the difference of the optical path. The detector measures the energy of the interference signal rather than the electric field. The energy is proportional to the irradiance, which is proportional to the modulus square of the electric field:

$$I_D = \left\langle \left| K_1 u(t) + K_2 u\left(t - \frac{2\Delta l}{c}\right) \right|^2 \right\rangle \tag{5.2}$$

Here the $\langle \dots \rangle$ denotes infinite time average, because the integration of the detector time is much longer than the oscillation at optical frequency. Since the detector measures real quantity, equation 5.2 can then be expanded as:

$$I_D(\Delta l) = K_1^2 \langle |u(t)|^2 \rangle + K_2^2 \left\langle \left| u\left(t - \frac{2\Delta l}{c}\right) \right|^2 \right\rangle + 2K_1 K_2 \text{Re} \left[\left\langle u(t) u^*\left(t - \frac{2\Delta l}{c}\right) \right\rangle \right] \tag{5.3}$$

The first two terms are both infinite time averages, which are just constant (DC) terms:

$$I_0 = \langle |u(t)|^2 \rangle = \left\langle \left| u\left(t - \frac{2\Delta l}{c}\right) \right|^2 \right\rangle \tag{5.4}$$

The last term is the autocorrelation function:

$$\gamma(\tau) = \langle u(t) u^*(t - \tau) \rangle \tag{5.5}$$

where τ is the time delay, which is $2\Delta l/c$ in this case. This term describes the interference of two beams.

For monochromatic plane wave, we have the autocorrelation function:

$$\gamma(\tau) = \left\langle \cos(2\pi v_0 t) \cos\left(2\pi v_0 \left(t - \frac{2\Delta l}{c}\right) \right) \right\rangle \tag{5.6}$$

$$\gamma(\tau) = \left\langle \cos\left(\frac{4\pi\Delta}{\lambda_0}\right) + \cos\left(4\pi v_0\left(t - \frac{\Delta l}{c}\right)\right)\right\rangle \qquad (5.7)$$

The first terms is the sinusoidal interference pattern as a function of optical path difference, the second term is averaged to zero. As we vary the optical path difference, we will observe a series of bright and dark fringe pattern, this pattern is also known as the interferogram.

Consider a perfect monochromatic light with wavelength λ_0 = 550nm, the interference pattern is plotted below (without the DC terms):
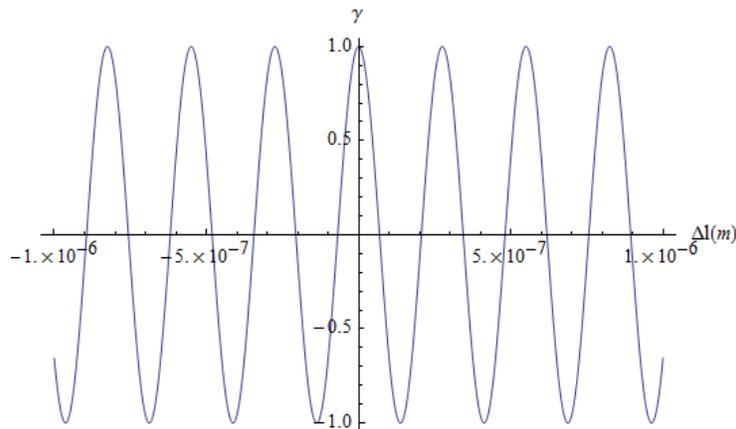
Figure 5.2: Interferogram with monochromatic light (550nm).

Constructive interference occurs when the optical path difference is in multiple of 275nm (due to double pass), likewise for the destructive interference. Figure 5.2 tells us some information of the height of the sample, if we observe maximum signal from the detector, we know the path difference between the reference arm and test arm is in multiple of half wavelength, but it is unclear what the exact difference is.

5.2 Interference with white light source

We have seen the interferogram with monochromatic light has bright and dark fringes repeat themselves as the optical path difference is varied. Before we model the equations with white light source, it is intuitive to see the effect of interferogram in the presence of multiple wavelengths. In this example, we add in light waves with λ_1 = 500nm, λ_2 = 600nm, on top of the original λ_0 = 550nm, the interferogram is shown in figure 5.3.

As we can see, the interference patterns of different frequency (wavelength) light coincide when the optical path difference is zero, and they are gradually out of phase when the optical path difference increases. Figure 5.4 shows an example of a photograph of white light fringe pattern, it is solely for illustration only.
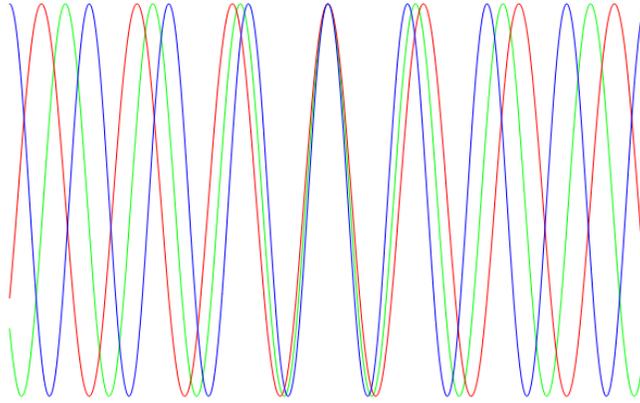
Figure 5.3: Interferogram with 3 monochromatic light, 500nm, 550nm, 600nm.
(Color indication for each component, not the actual wavelength)
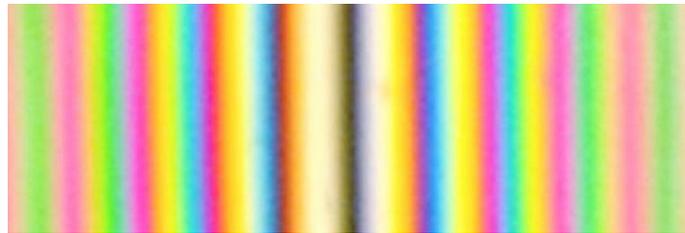
Figure 5.4: A photograph of white light fringe [51].

Now we add all the individual interferogram together, the resultant combination is as shown in figure 5.5.
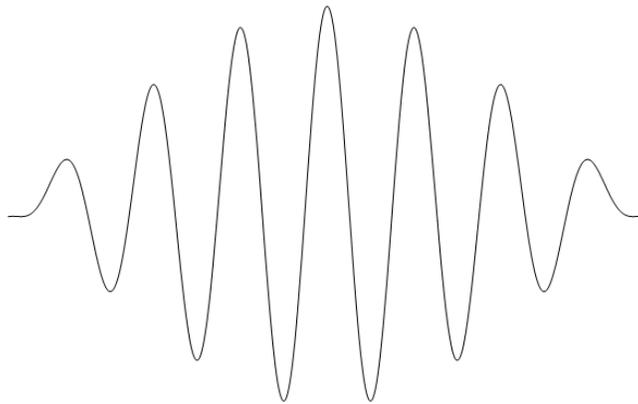
Figure 5.5: Resultant interferogram in the presence
of multiple wavelengths light.

From figure 5.5, we see that the interference pattern becomes "localized". The interference pattern follows an envelope which will be described later. This is the special property of white light interferometry which allows us to precisely determine the height information of a sample of interest.

Now we consider the interferometry with a spectrally broad light source (not necessarily a true white light source). The spectrum can be complicated in nature, here we simplify the model by assuming the spectrum having a Gaussian profile:

$$S(v) = |E(v)|^2 = Ce^{-\left(\frac{v-v_0}{\Delta v}\right)^2} \tag{5.8}$$

here v_0 is the central frequency, Δv is the bandwidth of the light source, and C is a constant term which is related to normalization factor, which is not of interest here. It will be absorbed into a new constant term subsequently.

We are interested to know the interferogram, which contains the autocorrelation function. To derive the autocorrelation function, we invoke the Winer-Khinchin Theorem [52], which states that autocorrelation function is given by the Fourier transform of the signal in the frequency domain:

$$\gamma(t) = \int_{-\infty}^{\infty} |E(v)|^2 e^{-2\pi i v t} dv \tag{5.9}$$

Substituting the spectrum equation 5.8, and replacing the time as time delay $\tau = 2\Delta l/c$,

$$\gamma(\tau) = C \int_{-\infty}^{\infty} e^{-\left(\frac{v-v_0}{\Delta v}\right)^2} e^{-2\pi i v \tau} dv \tag{5.10}$$

$$\gamma(\tau) = Ce^{-\pi^2 \Delta v^2 \tau^2} e^{-2\pi i v_0 \tau} \tag{5.11}$$

Equation 5.11 says that the autocorrelation function is an exponential oscillation modulated by a Gaussian profile. Next, we want to examine the signal received at the detector, which can be referred back to equation 5.3. Equation 5.3 contains the DC (constant) terms which depend on the reflectivities from the sample and the reference arm, we will just represent them as background DC signal.

$$I_D(\Delta l) = DC + Ce^{-\frac{4\pi^2 \Delta l^2 \Delta v^2}{c^2}} \cos\left(\frac{4\pi \Delta l}{\lambda_0}\right) \tag{5.12}$$

We can further relate the bandwidth of the spectrum to the coherence length l_c of the light source according to [53]:

$$l_c = \sqrt{\frac{2 \ln 2}{\pi} \frac{\lambda_0^2}{\Delta \lambda}} = \sqrt{\frac{2 \ln 2}{\pi} \frac{c}{\Delta v}} \tag{5.13}$$

For simpler representation of equation 5.13, we approximate it as follow:

$$l_c \approx \frac{2}{\pi} \frac{c}{\Delta v} \tag{5.14}$$

Lastly, the equation 5.12 is simplified to:

$$I_D(\Delta l) = DC + Ce^{-16\left(\frac{\Delta l}{l_c}\right)^2} \cos\left(2\pi\left(\frac{2\Delta l}{\lambda_0}\right)\right) \tag{5.15}$$

Equation 5.15 contains two pieces of information of our interest, the exponential term (Gaussian profile) is the envelope of the interferogram, which is modulated by the cosine oscillating term. Here we plot of the interferogram (without DC terms) as shown in figure 5.6, assuming λ_0 = 550nm, $\Delta\lambda$ = 100nm.
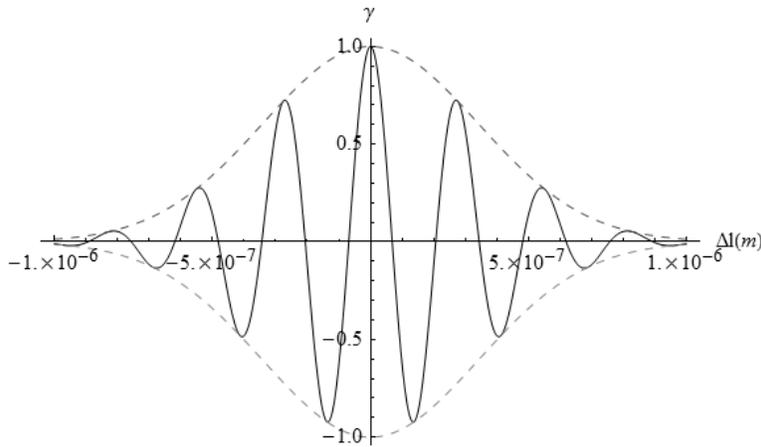
Figure 5.6: Interferogram from a spectrally broadband
light source with Gaussian spectral distribution.

The fringe spacing is $\lambda_0/2$, and the width of the envelope is inversely proportional to the coherence length (or proportional to the bandwidth) of the light source. In other words, the more incoherent the light source is, the more localized interferogram can be obtained, which leads to more accurate height determination.

5.3 Scanning white light interferometric microscopy

White light interferometry is often integrated to optical microscope system to extend the microscope's 2D imaging capability to 3D imaging with height measurement. White light interferometry works best for infinity corrected microscope, because a beam splitter can be inserted the space between the objective lens and the tube lens. As the light rays in between the objective lens and the tube lens are parallel, the beam splitter will not introduce optical aberrations to the imaging system. There exists multiple types of interferometric setup, the three major types are shown in figure 5.7.
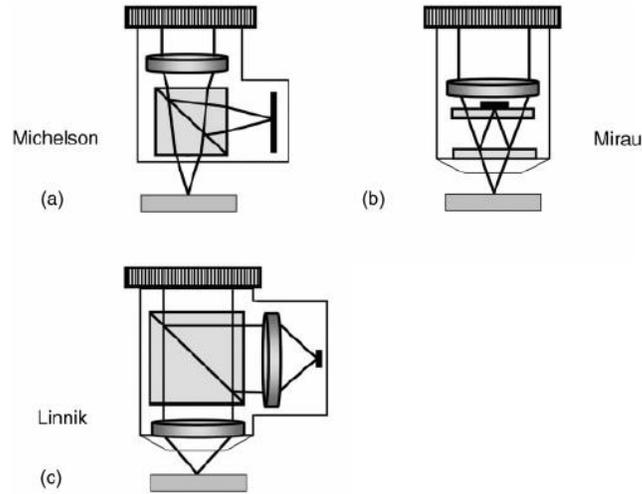
Figure 5.7: 3 major forms of interferometer configurations [54].

The light path coupled from the white light source is not shown in the figure 5.7. For the Michelson type, the beam splitter is outside of the objective lens which limits the working distance of the objective lens, so it is only suitable for low magnification (1X ~ 5X). The Mirau type contains the miniaturized reference mirror right in front of the objective lens. This design is compact, and hence it allows higher magnification (10X ~ 50X). The miniaturized mirror is to have the same size as the illuminated surface of the object. For high magnification operation, the shadowing effect is minimal. For the Linnik type, the reference mirror does not limit the working distance, it can work with any magnification by using two identical objective lenses. However, such kind of interferometer is very difficult to adjust, hence its commercial use is rather limited [54].

The most common interferometer type is the Mirau type, hence we will discuss the working principle of white light interferometric microscope based on Mirau configuration. Figure 5.8 shows the schematic diagram. In the schematic diagram, most of the optical components are simplified and the ray displacements due to window thickness are ignored. A point like white light source is collimated by the condenser lens, illuminating the sample surface co-axially after being reflected by the beam splitter. As in conventional microscope, the sample is imaged by the objective lens into infinity space, the collimated beam is then imaged by the tube lens onto a 2D imaging sensor. The 2D magnification is given by the ratio of the focal length of the tube lens to the focal length of the objective lens.

The interference paths at the Mirau objective lens is shown in figure 5.9. In the schematic diagram, only 1 ray coming from the white light illumination is shown. Upon hitting the beam splitter, the partial mirror splits the light path into reference arm and test arm, allowing interference effect to be recorded at the imaging detector. The partial mirror is placed halfway between the objective lens and the sample focused position, this implies that the peak of the interferogram is obtained when the sample is at the best focusing position.

The Mirau objective lens is mounted on a high precision positioning stage, such as a piezo-electric actuator, which allows z-scanning throughout the surface profile of the sample. Since the microscope system is infinity corrected, the displacement of the objective lens does not change the image magnification. From the z-scanning, the intensity of each pixel is recorded on the imaging sensor, this process produces the interferogram.
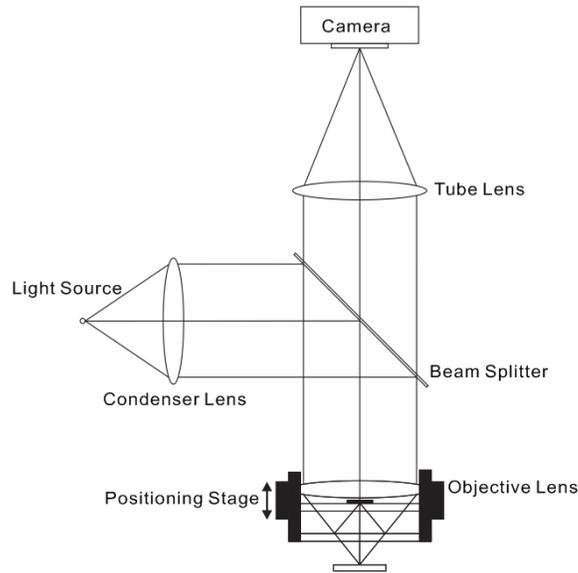
Figure 5.8: A Mirau type interferometric microscope system.
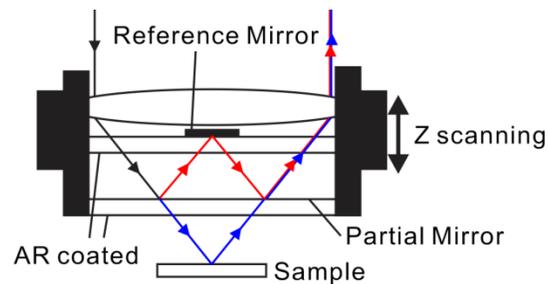
Figure 5.9: Close up look at the Mirau objective lens
(Color indication for split light paths)

As discussed earlier, the signal of the interferogram contains DC components, they are to be removed by digital high pass filter. The next task is to find the peak of the interferogram, in the real situation, the peak of the interferogram does not necessarily coincide with the peak of the envelope function. Moreover, this also depends on the sampling step size, but it will be impractical to use extremely fine sampling step to locate the peak of the interferogram.

There exists different kind of algorithms to determine the peak of the interferogram, and hence the surface height information. Hilbert transform algorithm [55] developed by Kino and Chim, direct quadratic polynomial fitting algorithm developed by Park and Kim [56] are

some of the notable examples. The algorithm of finding the peak of the interferogram is crucial in white light interferometry, because it is one of the parameters that determines the height measurement accuracy.

5.4 General advantages and disadvantages

The working principle of white light interferometric microscopy is similar to the depth from focus microscopy due to the z-scanning. However, one clear advantage of white light interferometric microscopy is that it works with smooth and shiny surface. Depth from focus method is not able to sense the height information for such surface with no feature for focus detection. Another advantage is the depth resolution of white light interferometric microscopy does not depend on the 2D image magnification, this allows high resolution depth measurement with wide field of view.

On the other hand, white light interferometry is also similar to the phase shifting interferometry, except that phase shifting interferometry uses monochromatic light in most cases. One common use of phase shifting interferometry is the optical testing, such as measuring the surface profile of polished mirrors and lenses. The accuracy of phase shifting is generally much better than white light interferometry, but there are two major limitations of the phase shifting interferometry. The first limitation is that the surface has to be shiny in order to obtain good contrast fringes. The second limitation is that the surface's profile needs to have rather smooth profile. Any discontinuity of the surface profile leads to discontinuous fringes, leading to phase ambiguity. This is also known as 2π ambiguity, this problem is similar to what discussed in the chapter 3, except that the fringes were obtained in different ways.

In general, white light interferometry is considered a robust and versatile solution to obtain 3D measurements of various objects as it works with both rough and reflective surfaces, as well as structures with steep discontinuities. However, the major drawback is that the time acquisition of 3D measurement is rather long, which is due to the z-scanning process. This disadvantage is rather the fundamental limit of optical system, the depth of field is extremely limited especially for high magnification operation. The z-scanning is necessary in order to get perfectly focused images over large height variation.

5.5 Case study

The non-contact 3D optical profiling instrument based on white light interferometry offers wide range of metrology applications, such as measuring object's 3D profile, characterizing surface roughness, surface parallelism, etc. It is often used for inspecting products such as wafers, MEMS device, solar cell, microelectronics, etc.

In this case study, we take a look at the operation of measuring the topography of a micro-machined sample with Polytec's instrument [57]. Some of the steps of obtaining the 3D topography of the sample are shown in figure 5.10. When an object is initially placed under the interferometric microscope objective, the operator needs to firstly focus the object. Next, the operator needs to adjust the tip-tilt of the reference mirror. When the reference mirror

is not parallel to the sample surface, the fringes will be closely packed due to the tilted wavefronts between the two-wave interference, this effect is shown in the top left of the figure 5.10. When the tilt of the reference mirror is adjusted, the fringe pattern will look like the top right of the figure 5.10. The operator can select the region of interest to perform the z-scanning within a range (middle of figure 5.10). After the z-scanning process is completed, the software calculates the height information, and the operator can choose to view the 3D model from arbitrary viewing angle (bottom of figure 5.10).
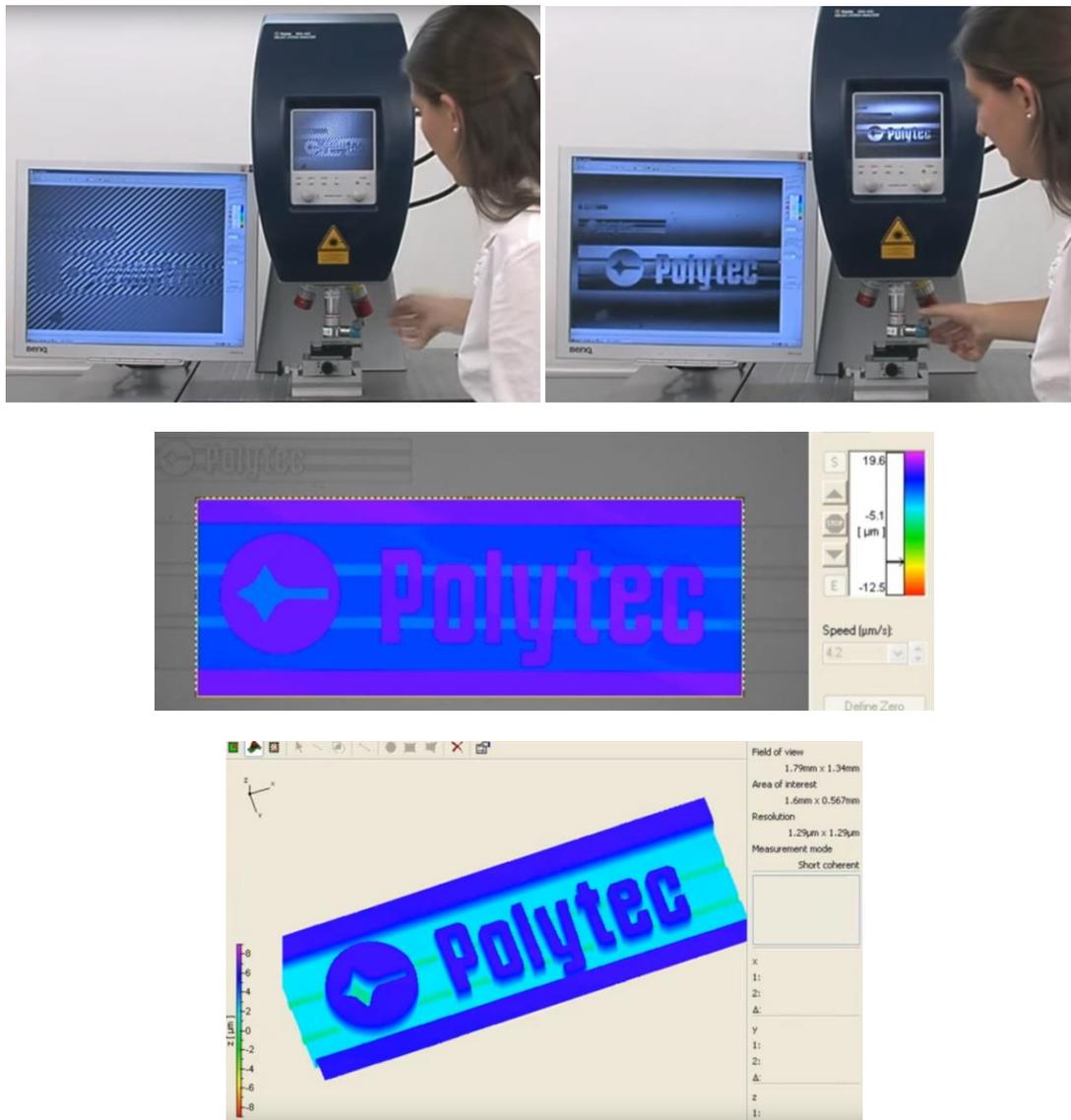

Figure 5.10: Steps toward obtaining 3D topography of a micro-machined sample, image courtesy of Polytec GmbH [58].

5.6 Conclusion

In conclusion, white light interferometry is a robust 3D imaging technique as it works for various kinds of surfaces. For this reason, it is often used for metrology application.

Chapter 6: Optical Coherence Tomography

Optical coherence tomography is another type of 3D imaging technique which is widely used in medical imaging. The working principle of optical coherence tomography is very similar to the white light interferometry, which is the use of low coherence light source. It is sometimes referred as low coherence interferometry.

By creating interference effect with low coherence light source, one can obtain the depth information in the same way as white light interferometry. The main difference is that optical coherence tomography uses long wavelength light source to penetrate into scattering medium, while white light interferometry works with surface reflection (specular surface) or surface reflectance (diffuse surface). When the light is scattered inside a scattering medium, the backward scattered light interferes with the reference light. As we scan through the sample, an interferogram is produced and hence the depth information can be determined. The classical application of optical coherence tomography is the in vivo retinal imaging, as illustrated in figure 6.1.
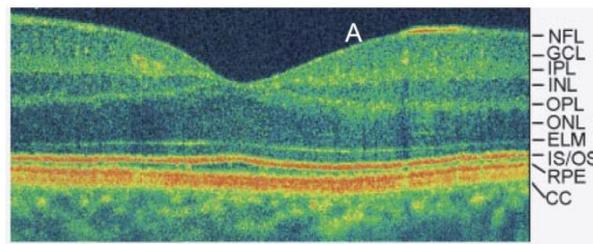
Figure 6.1: Cross-sectional imaging of retina (fovea) reveals
multiple layers of cell structures [59].

Besides medical imaging application, optical coherence tomography is also applicable for industrial application, such as nondestructive testing, volumetric measurement, thickness measurement of silicon wafer, etc.

Optical coherence tomography (OCT) is classified into two categories, namely the time domain OCT, the Fourier domain OCT. The Fourier domain OCT is further classified into the spectral domain OCT and the swept source OCT which will be discussed later. Time domain OCT is considered the traditional method for OCT imaging, it requires mechanical scanning of the reference mirror which is much more time consuming comparing to the Fourier domain OCT. Nowadays, most of the commercial OCT system is based on Fourier domain OCT owing to the faster image acquisition.

The working principle of time domain OCT is very similar to the white light interferometry which was discussed in the previous chapter. In this chapter, we will focus on working principle of Fourier domain OCT. The book "Optical Coherence Tomography – Technology and Applications" written by Drexler and Fujimoto [60] provides theoretical model and comprehensive explanations of OCT technology, most of the materials in this chapter will be referenced from this book.

6.1 General setup

Figure 6.2 shows the general schematic setup of an OCT system with fiber optics. In an OCT setup, low coherence source with long wavelength is often used. Superluminescent diodes are the most commonly used light sources in OCT system, they are available in 800nm, 1000nm, 1300nm, etc, with bandwidth about 100 – 150nm.
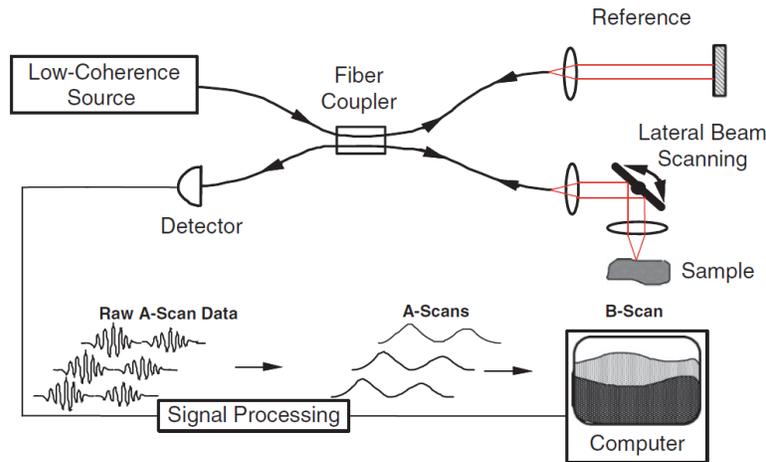
Figure 6.2: Schematic diagram for a general OCT system [60].

Low coherence light will be split at the fiber coupler, with one channel going to the reference mirror, and another channel going to the sample arm, forming a Michelson type interferometer. As the backscattered signal from the sample is much weaker from the mirror, the split ratio of the fiber coupler is often not 50:50, it can be a 20:80 for example.

The light coupled out from the fiber end will be re-collimated and focused on the sample of interest. The backscatterd light from the sample is then redirected back to the fiber and mixed with the reflected light from the reference fiber at the fiber coupler. Interference signal will then be received at the detector.

To obtain cross-sectional image, an axial scan which is also known as A-scan, is required to scan through a particular point of the sample. With traditional time domain OCT, the A-scan is performed by moving the reference mirror to obtain the interferogram. For Fourier domain OCT, the A-scan is achieved by taking the inverse Fourier transform of the signal recorded by the detector, this will be discussed later.

As the light is focused into a spot on the sample, when A-scan is performed, we only obtain the depth information for a single point. To obtain 2D cross-sectional image, the focusing spot needs to be scanned along a line by tilting a galvanometric actuated mirror as shown in figure 6.2, this scanning process is also known as the B-scan. The galvanometric mirror can be tipped and tilted in both x and y direction, and hence achieving a 3D volumetric imaging of the sample of interest. The scanning scheme is illustrated in figure 6.3.
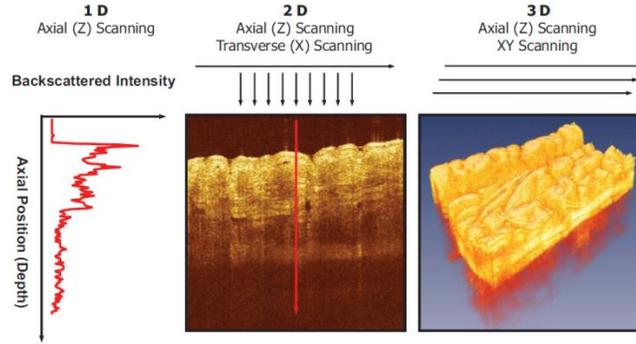
Figure 6.3: 1D, 2D, 3D scanning for OCT operation [60].

Similarly to white light interferometry, the bandwidth of the light source dictates the axial resolution, the wider the bandwidth, the better the axial resolution can be achieved. The center wavelength of the light source also plays an important role in OCT system. As a rule of thumb, longer wavelength light penetrates scattering medium (such as biological tissues) deeper, which can reveal the internal structure better. However, due to the diffraction effect, longer wavelength light will not be as tightly focused as shorter wavelength light. That is, using the same numerical aperture (NA) objective lens, the lateral resolution with longer wavelength light is lower.

Another important parameter of the OCT imaging system is the axial field of view, which is dictated by the depth of focus of the focused spot, in the diffraction limited regime. As we know the lateral resolution is inversely proportional to the NA, while the depth of focus is inversely proportional to the square of the NA, it implies that high lateral resolution comes with the penalty of small axial field of view (small working range). The definitions of lateral resolution, axial field of view, axial resolution, etc really depend on which criteria are used, figure 6.4 only serves to illustrate the main concept which describes the performance of an OCT system.
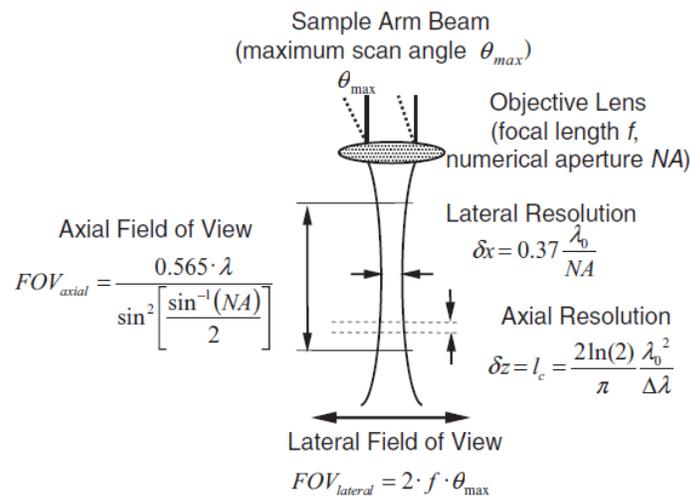
Figure 6.4: Relation of lateral, axial resolution and field of view [60].

6.2 Working principle of Fourier domain OCT
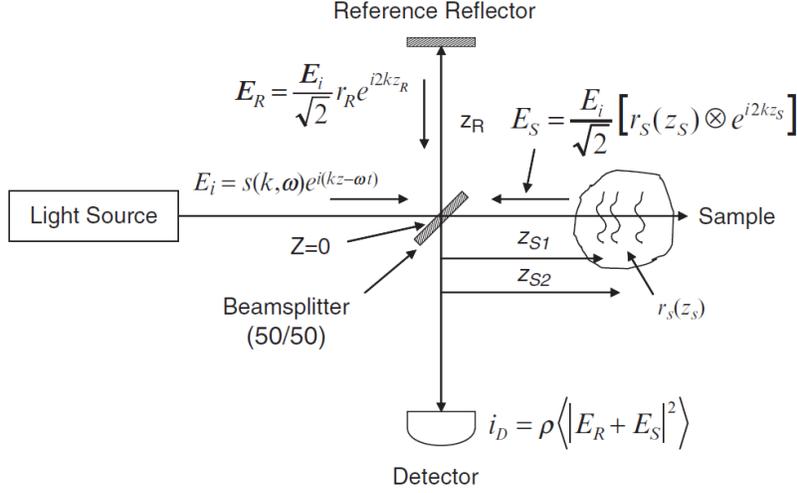
Figure 6.5: Schematic of a Michelson interferometer used in an OCT system [60].

Figure 6.5 shows the schematic diagram of a simplified Michelson interferometer for explaining the working principle of the OCT system. Consider the light coming out from the low coherence light source is a plane wave which electric field can be described as:

$$E_i = s(k, \omega)e^{i(kz - \omega t)} \tag{6.1}$$

where $s(k, \omega)$ is the electric field amplitude of the light source as a function of wave number ($k = 2\pi/\lambda$) and angular frequency ($\omega = 2\pi v$). The beam splitter splits the light equally into the reference arm and the sample arm. The return signals are attenuated by factors of r_R and r_S respectively. Since the distance from the beam splitter to the reference reflector is z_R, the accumulated phase is then $i2kz_R$ (double pass).

For the sampling arm, the sample is described by a stack of multi layers structures, it can be modeled as a series of delta functions:

$$r_s(z_s) = \sum_{n=1}^{N} r_{Sn} \delta(z_S - z_{Sn}) \tag{6.2}$$

and the return signals can then be described as convolution of the delta functions with the accumulated phase $i2kz_S$ as follow:

$$E_S = \frac{Ei}{\sqrt{2}} [r_s(z_s) * e^{i2kz_S}] \tag{6.3}$$

The resultant return signals are multiple plane waves with different phase delay depending on the depth of the scattering sites. The return signals from the reference reflector and the sample are combined at the fiber coupler and hence interfered at the detector. The detector generates photocurrent which is proportional to the square of the sum of the electric fields:

$$I_D(k,\omega) = \left\langle \left| \frac{s(k,\omega)}{\sqrt{2}} r_R e^{i(2kz_R - \omega t)} + \frac{s(k,\omega)}{\sqrt{2}} \sum_{n=1}^{N} r_{Sn} e^{i(2kz_{Sn} - \omega t)} \right|^2 \right\rangle \qquad (6.4)$$

where $\langle \dots \rangle$ denotes the integration time of the detector. Since the response time of detector is much slower than oscillation at optical frequencies, the temporal term will be left out as a product of time average. In addition, the detector measures real quantity, hence expanding equation 6.4 leading to:

$$I_D(k,\omega) = S(k)[R_R + R_{S1} + R_{S2} + \cdots]$$
$$+ 2\left[S(k) \sum_{n=1}^{N} \sqrt{R_R R_{Sn}} \cos[2k(z_R - z_{Sn})] \right] \qquad (6.5)$$
$$+ \left[S(k) \sum_{n \ne m=1}^{N} \sqrt{R_{Sn} R_{Sm}} \cos[2k(z_{Sn} - z_{Sm})] \right]$$

Here $S(k) = \langle |s(k,\omega)|^2 \rangle$ is the spectral distribution of the light source. Equation 6.5 is known as the spectral interferogram, it contains three distinct components:

1) The first line corresponds to the "DC terms", it is the background signals received by the detector which do not yield depth information of a sample.
2) The second line corresponds to the "Cross-correlation terms", it is the desired output for the OCT system, it reveals the depth information of a sample. Due to the square root dependency of the sample and reference mirror's reflectivities, these terms are typically smaller than the DC terms.
3) The third line corresponds to the "Autocorrelation terms", it is the interference that occurs between different layers of the sample. These signals are the artifacts in a typical OCT system. The magnitudes are often much smaller than the DC and the cross-correlation terms, because they depend on the reflectivities of the samples which are much weaker than the reference mirror.

Since the interferogram from equation 6.5 is in described the k-space (wave number space), by taking inverse Fourier transform, we obtain the interferogram in the z-space (position space). Using the Fourier relation below:

$$\gamma(z) \overset{F}{\leftrightarrow} S(k) \qquad (6.6)$$

$$\frac{1}{2}[\delta(z + z_0) + \delta(z - z_0)] \overset{F}{\leftrightarrow} \cos kz_0$$

as well as the convolution property of Fourier transform:

$$x(z) * y(z) \overset{F}{\leftrightarrow} X(k)Y(k) \qquad (6.7)$$

We obtain

$$i_D(z) = \gamma(z)[R_R + R_{S1} + R_{S2} + \cdots]$$
$$+ 2\left[\gamma(z) * \sum_{n=1}^{N} \sqrt{R_R R_{Sn}}\left(\delta\big(z \pm 2(z_R - z_{Sn})\big)\right)\right] \tag{6.8}$$
$$+ \left[\gamma(z) * \sum_{n \neq m=1}^{N} \sqrt{R_{Sn} R_{Sm}}\left(\delta\big(z \pm 2(z_{Sn} - z_{Sm})\big)\right)\right]$$

The convolution operation between $\gamma(z)$ and multiple delta functions become multiple copies of $\gamma(z)$ profiles at locations of the delta functions:

$$i_D(z) = \gamma(z)[R_R + R_{S1} + R_{S2} + \cdots]$$
$$+ 2\sum_{n=1}^{N} \sqrt{R_R R_{Sn}}\left[\gamma\big(2(z_R - z_{Sn})\big) + \gamma\big(-2(z_R - z_{Sn})\big)\right] \tag{6.9}$$
$$+ \sum_{n \neq m=1}^{N} \sqrt{R_{Sn} R_{Sm}}\left[\gamma\big(2(z_{Sn} - z_{Sm})\big) + \gamma\big(-2(z_{Sn} - z_{Sm})\big)\right]$$

Equation 6.9 is the result of A-scan for the Fourier domain OCT, it reproduces the sample profile (equation 6.2) with slight modification. The original input to the OCT system, which is the sample profile, is blurred out by the function $\gamma(z)$. This blurred output is analogous to the Point Spread Function (PSF) of an imaging system, except this applies to the axial resolution rather than the lateral resolution. A plot of the equation 6.9 is illustrated in figure 6.6. Since the detected interferometric spectrum is real, its inverse Fourier transform must be Hermitian symmetric. In addition, the DC terms and the autocorrelation terms give rise to artifact signals which are centered at zero path length should be filtered carefully.
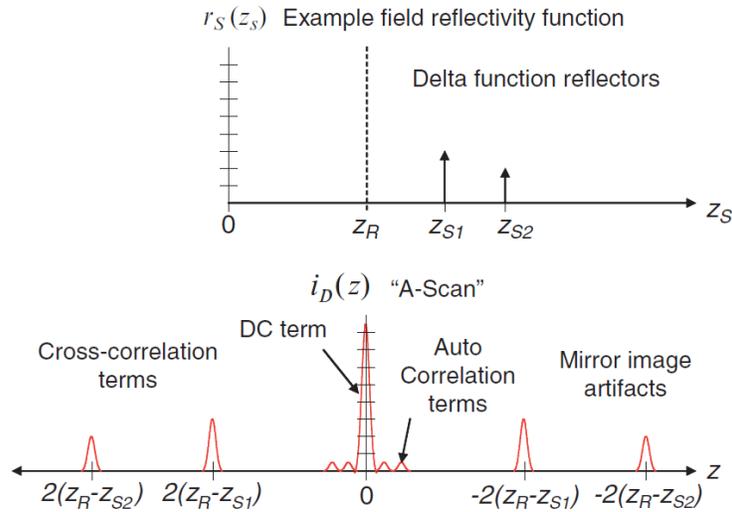

Figure 6.6: Illustration of the result of A-scan in Fourier domain OCT [60].

The function $\gamma(z)$ is the inverse Fourier transform of the spectral function $S(k)$. Since the bandwidths of the function $\gamma(z)$ and $S(k)$ are related by the "uncertainty principle", which means that performing OCT imaging with large bandwidth incoherent source leads to more precise axial measurement in the cross-sectional image. This idea agrees with our basic understanding of OCT imaging with the use of incoherent source with short coherence length.

6.3 Implementations of Fourier domain OCT

Spectral domain OCT

Equation 6.5, which is the spectral interferogram, assumes the detector is able to acquire the photon signals as a function of wavelength (or wavenumber, or optical frequency). In spectral domain OCT, when a broadband light source is used, all spectral components of the signals (equation 6.5) are captured by a detector array placed at the output of a spectrometer. Hence, spectral domain OCT is also known as spectrometer-based OCT. Figure 6.7 shows a schematic setup of spectral domain OCT.
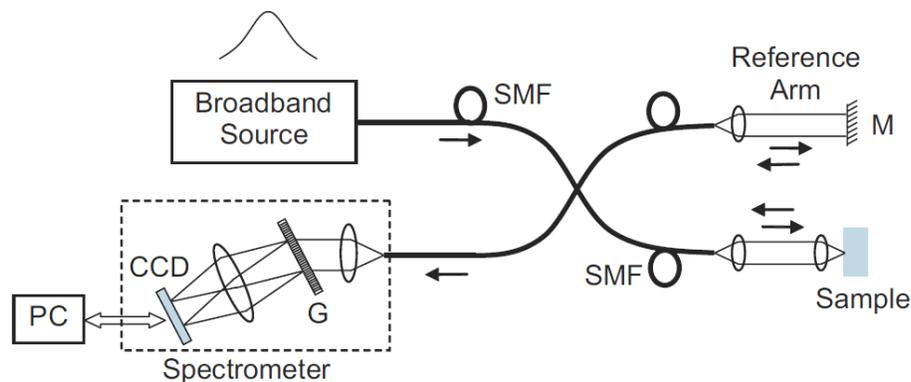
Figure 6.7: Schematic diagram for a typical spectral domain OCT [59].
SMF: Single Mode Fiber, M: Mirror, G:Grating

A typical spectrometer consists of a diffraction grating (transmission, or reflection based) which disperse the interference beam into a multiple orders of diffracted beams. A linear detector (e.g. CCD camera) then captures the interference signal of an OCT system as a function of wavelength. Usually, the spectral data is rescaled and resampled evenly in the wavenumber space (k-space), and then it will be inverse Fourier transformed to obtain the raw data of the depth profile of a sample.

Swept Source OCT

For swept source OCT, narrowband laser is used instead of broadband light source. A tunable narrowband laser is rapidly swept across a certain optical bandwidth. For example, tunable Fabry-Perot cavity tunable laser with above 20kHz swept rate allows detection of interference signal at high wavelength modulation rate.

In a typical swept source OCT setup as shown in figure 6.8, a circulator is used to collect out of phase interference signal. The out of phase signal is to be subtracted from the return interference signal, such that the desired interference signal is enhanced and excess noise from the light source can be cancelled [61].

Figure 6.8: Schematic diagram for a typical swept source OCT [59].

One advantage of swept source OCT over the spectrometer-based OCT is the reduced noise signal. This is because at each wavenumber, the output of a swept source is a narrowband laser with much smaller noise than that of a broadband light source [59].
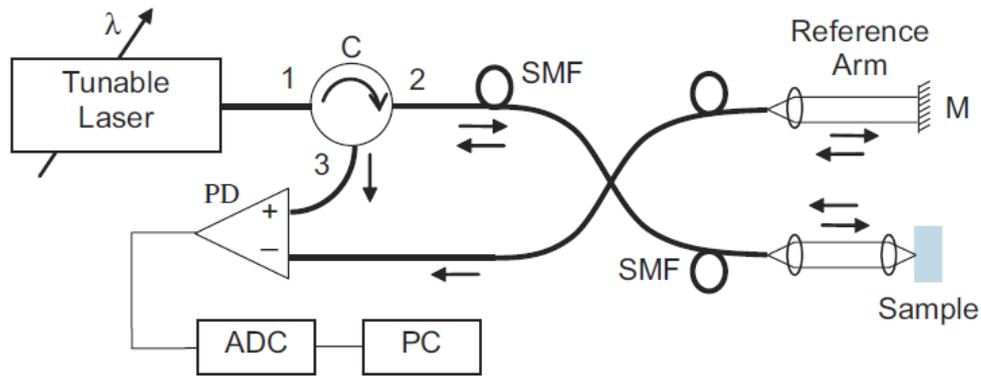
Dispersion compensation

For Fourier domain OCT, dispersion phenomenon needs to be paid attention as it degrades the quality of OCT images. As we know the wider the bandwidth of the light source (either broadband source or swept source laser), the better the axial resolution is, meanwhile the dispersion effect becomes more severe. The wavenumber dependent phase shift can be compensated by placing optical materials (such as BK7 glass block) in the reference arm, which had been demonstrated by M. Wojtkowski et al. [61]. The dispersion mismatch can also be compensated from software approach. B. Cense et al. demonstrated the phase shift due to dispersion mismatch can be measured and consequently corrected for retinal imaging application [62].

6.4 Comparison of OCT and other medical imaging techniques

As mentioned earlier, OCT is an imaging technology that is widely used in medical imaging. Here we take a look at two other medical imaging techniques, namely the confocal microscopy and the ultrasound imaging and compare their capabilities [60].

While confocal microscopy offers extremely high resolution imaging (~1µm) capability, its imaging depth in biological tissue is very limited (~few hundred micrometers) as the image's contrast is significantly reduced by optical scattering. On the other hand, OCT

detects the signals of scattered light within biological tissues, which allows much deeper image penetration depth (~2mm) comparing to confocal microscopy.

Ultrasound technology is a very well established medical imaging technique which is quite similar to OCT, except that it uses sound instead of light. It sends pulses of ultrasound waves, and determines the depth by measuring the "echo" time upon backscattered from human body. Typical clinical ultrasound imaging allows imaging deep inside human body with low resolution (0.1-1mm). While high frequency ultrasound leads to higher resolution imaging (~20μm), its penetration depth is more limited due to strong attenuation.

Overall in the medical imaging applications, OCT imaging technology fills in the gap in between the ultrasound technology and the confocal microscopy, which is illustrated in figure 6.9. In particular, it has become a clinical standard in ophthalmology due to the in-vivo imaging capability with considerably high resolution.
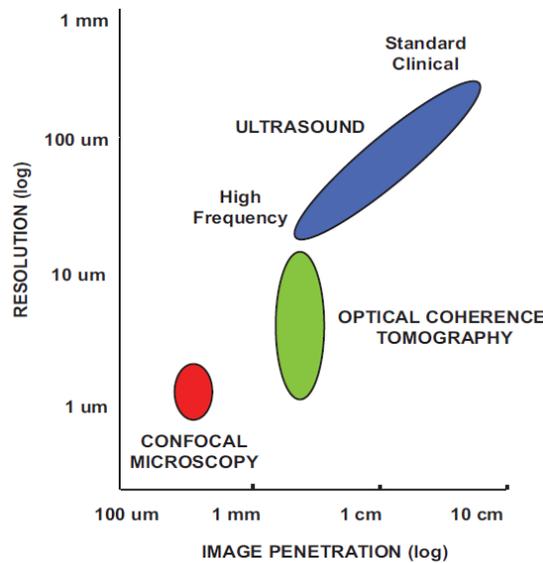
Figure 6.9: Comparison of resolution and penetration depth for ultrasound, OCT, and confocal microscopy [60].

6.5 Case study

In this case study, we again look at the integration of Optotune's liquid tunable lens which was introduced in the previous chapter. The ability of changing focal length with electrically tunable lens can lead to improvement of OCT imaging.

As discussed earlier in section 6.1, for a general OCT imaging system, a compromise has to be decided between the lateral resolution and the axial field of view because tighter focused spot leads to smaller depth of focus. The integration of electrically tunable lens allows high lateral resolution with extended depth of scanning, the concept is depicted in figure 6.10.
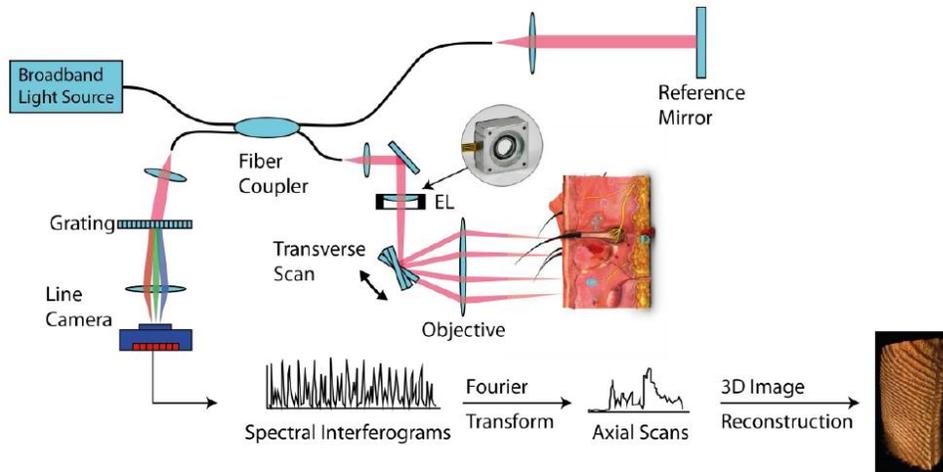
Figure 6.10: Schematic setup of a spectral domain OCT with integration of electrically tunable lens (EL) [63].
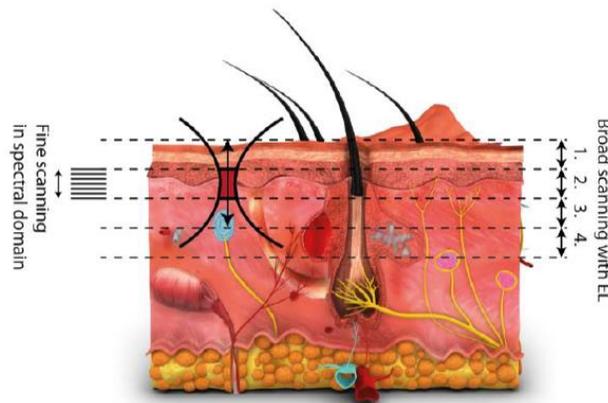
Figure 6.11: Illustration of stitching several A-scans in order to obtain large scan depth while maintaining high lateral resolution [63].

For this example, as illustrated in figure 6.11, a thick layer of human skin is to be scanned by a spectral domain OCT system described by figure 6.10. The electrically tunable lens changes the "penetration depth" of a tight focused spot, hence allowing broad scanning range across the skin tissue. High lateral resolution can be achieved with tight focused spot by using high NA objective lens in this case. The fine axial scanning is still achieved by performing A-scan, and the axial resolution is determined by the coherence length (or bandwidth) of the broadband light source. The result is a high lateral resolution OCT imaging system with large scan depth by using electrically tunable lens, which is faster than mechanical scanning.

6.6 Conclusion

In conclusion, optical coherence tomography (OCT) is a well-established imaging technique which is especially useful to generate cross-sectional images of a biological tissues. By penetrating through a biological sample, OCT offers high resolution cross-sectional imaging which is better than other alternatives, in-vivo retinal imaging is a good example.

References

[1] Theo Moons, Luc Van Gool, and Maarten Vergauwen, "3D Reconstruction from multiple images part 1: principles," Foundations and Trends in Computer Graphics and Vision: Vol.4, pp. 287-404 (2010)

[2] Charles T. Loop, and Zhengyou Zhang, "Computing rectifying homographies for stereo vision," Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Vol. 1, pp. 125-131 (1999)

[3] V. Kolmogorov, and R. Zabih, "Computing visual correspondence with occlusions using graph cuts," Proceedings Eighth IEEE International Conference on Computer Vision, ICCV pp. 508-515(2001)

[4] Jian Sun, Nan-Ning Zheng, and Heung-Yeung Shum, "Stereo matching using belief propagation," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 25, no. 7, pp. 787-800 (2003)

[5] Y. Ohta, and T. Kanade, "Stereo by intra- and inter-scanline search using dynamic programming," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. PAMI-7, no. 2, pp. 139-154 (1985)

[6] http://campar.in.tum.de/twiki/pub/Chair/TeachingWs09Cv2/3D_CV2_WS_2009_Stereo.pdf

[7] http://www.ni.com/white-paper/14103/en/

[8] http://www.nevis.com.tw/file/image/files/chromasens-app-note-wire-bond20140919.pdf

[9] Francois Blais, Marc Rioux, and J.-Angelo Beraldin, "Practical considerations for a design of a high precision 3-d laser scanner system," Proceeding SPIE 0959, Optomechanical and Electro-Optical Design of Industrial Systems, 225 (1988)

[10] Francois Blais, "Review of 20 years of range sensor development," Proceeding SPIE 5013, Videometrics VII (2003)

[11] http://www.laserlineoptics.com/powell_primer.html

[12] Lyubomir Zagorchev, and A. Ardeshir Goshtasby, "A paint-brush laser range scanner," Computer Vision and Image Understanding, Vol. 101, pp. 65-86 (2006)

[13] Winkelbach S., Molkenstruck S., and M. Wahl F., "Low-Cost Laser Range Scanner and Fast Surface Registration Approach," DAGM 2006, LNCS 4174, pp. 718-728, 2006, Springer Berlin Heidelberg (2006)

[14] Guan Xu, et. al., "An optimization solution of a laser plane in vision measurement with the distance object between global origin and calibration points," Sci. Rep. 5, 11928 (2015)

[15] Park, Y., Yun, S., Won, C.S., Cho, K., Um, and K., Sim, S., "Calibration between color camera and 3D LIDAR instruments with a polygonal planar board," Sensors 2014, 14, pp. 5333-5353 (2014)

[16] http://www.david-3d.com/en/support/downloads

[17] http://wiki.david-3d.com/david3_user_manual/camera_calibration

[18] G. L. Oomen, and W.J.P.A Verbeek, "A real-time optical profile sensor for robot arc welding," Proc. SPIE 0449, pp. 62-71 (1984)

[19] Hanlin Zhang, Yongjie Ren, Changjie Liu, and Jigui Zhu, "Flying spot laser triangulation scanner using lateral synchronization for surface profile precision measurement," Appl. Opt. 53, pp. 4405-4412 (2014)

[20] http://www.nikonmetrology.com/en_EU/Products/Laser-Scanning/CMM-scanning/LC15Dx-laser-scanner

[21] http://www.keyence.com/products/measure/laser-2d/lj-v/features/index.jsp

[22] J. L. Posdamer, and M. D. Altschuler, "Surface measurement by space-encoded projected beam systems," Computer Graphics and Image Processing Vol. 18, Issue 1, pp. 1-17 (1982)

[23] Jason Geng, "Structured-light 3D Surface imaging: a tutorial," Advances in Optics and Phonics, Vol. 3, pp. 128-160 (2011)

[24] S. Inokuchi, K. Sato, and F. Matsuda, "Range-imaging for 3-D object recognition," International Conference on Pattern Recognition (International Association for Pattern Recognition), pp. 806-808 (1984)

[25] Jens Guehring, "Dense 3D surface acquisition by structured light using off-the-shelf components," Proceeding SPIE 4309, Videometrics and Optical Methods for 3D Shape Measurement, 220 (2000)

[26] Joaquim Salvi, Jordi Pagès, and Joan Batlle, "Pattern codification strategies in structured light systems," Pattern Recognition, Vol. 37, Issue 4, pp. 827-849 (2004)

[27] Claus Brenner, Jan Boehm, and Jens Guehring, "Photogrammetric calibration and accuracy evaluation of a cross-pattern stripe projector," Proceeding SPIE 3641, Videometrics VI, 164 (1998)

[28] Brian Carrihil, and Robert Hummel, "Experiments with the intensity ratio depth sensor," Computer Vision Graphics and Image Processing, Vol. 32, Issue 3, pp. 337-358 (1985)

[29] J. Tajima, and M. Iwakawa, "3-D data acquisition by Rainbow Range Finder," Proceedings. 10th International Conference on Pattern Recognition, Vol. 1, pp. 309-313 (1990)

[30] Tatsuo Sato, "Multispectral pattern projection range finder," Proceeding SPIE 3640, Three-Dimensional Image Capture and Applications II, 28 (1999)

[31] Clarence Wust, and David W. Capson, "Surface profile measurement using color fringe projection," Machine Vision and Applications, Vol. 4, Issue 3, pp. 193-203 (1991)

[32] N. G. Durdle, J. Thayyoor, and V. J. Raso, "An improved structured light technique for surface reconstruction of the human trunk," Conference Proceedings. IEEE Canadian Conference on Electrical and Computer Engineering, Vol. 2, pp. 874-877 (1998)

[33] M. Maruyama, and S. Abe, "Range sensing by projecting multiple slits with random cuts," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 15, Issue 6, pp. 647-651 (1993)

[34] Li Zhang, B. Curless, and S. M. Seitz, "Rapid shape acquisition using color structured light and multi-pass dynamic programming," Proceedings. First International Symposium on 3D Data Processing Visualization and Transmission, pp. 24-36 (2002)

[35] Paul M. Griffin, Lakshmi S. Narasimhan, and Soung R. Yee, "Generation of uniquely encoded light patterns for range data acquisition," Pattern Recognition, Vol. 25, Issue 6, pp. 609-616 (1992)

[36] R. A. Morano, C. Ozturk, R. Conn, S. Dubin, S. Zietz, and J. Nissano, "Structured light using pseudorandom codes," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 20, Issue 3, pp. 322-327 (1998)

[36] http://smarttech3dscanner.com/smarttech-3d-scanning-technologies-are-used-by-the-enduro-tech-specialized-motorcycle-accessories-producer/

[37] http://www.xyztlab.com/

[38] Zhang S, and Huang PS, "High-resolution, real-time three-dimensional shape measurement," Optical Engineering, Vol. 45, Issue 12 (2006)

[39] https://www.youtube.com/watch?v=8nTFjVm9sTQ

[40] Born M, and Wolf E., "Principles of Optics (6th ed.)," Oxford Pergamon Press (1980)

[41] I. T. Young, R. Zagers, L. J. van Vliet, J. Mullikin, F. Boddeke, H. Netten, "Depth-of-focus in microscopy," Proceedings of the Scandinavian Conference on Image Analysis. Vol. 1 (1993)

[42] Richard Leach, "Optical Measurement of Surface Topography," Springer (2011)

[43] Marvin Minsky, "Microscopy apparatus," US patent 3013467 A (1961)

[44] Lawrence D. Favro, Robert L. Thomas, Pao-Kuang Kuo, Li Chen, "Confocal microscope," US patent 5162941 A (1992)

[45] http://zeiss-campus.fsu.edu/tutorials/spinningdisk/yokogawa/indexflash.html

[46] http://www.olympusmicro.com/primer/techniques/confocal/confocalintro.html

[47] http://www.keyence.com/products/microscope/digital-microscope/vhx-1000/features/feature-05.jsp

[48] https://www.photonics.com/Article.aspx?AID=57323

[49] http://www.optotune.com/applications/microscopy

[50] Joey M. Jabbour, Bilal H. Malik, Cory Olsovsky, Rodrigo Cuenca, Shuna Cheng, Javier A. Jo, Yi-Shing Lisa Cheng, John M. Wright, and Kristen C. Maitland, "Optical axial scanning in confocal microscopy using an electrically tunable lens," Biomed. Opt. Express 5, pp. 645-652 (2014)

[51] https://en.wikipedia.org/wiki/White_light_interferometry

[52] Weisstein, and Eric W., "Wiener-Khinchin Theorem." From *MathWorld*--A Wolfram Web Resource. http://mathworld.wolfram.com/Wiener-KhinchinTheorem.html

[53] Ceyhun Akcay, Pascale Parrein, and Jannick P. Rolland, "Estimation of longitudinal resolution in optical coherence imaging," Appl. Opt. 41, 5256-5262 (2002)

[54] Schmit, J., Creath, K., and Wyant, J. C., "Surface Profilers, Multiple Wavelength, and White Light Intereferometry, in Optical Shop Testing, Third Edition (ed D. Malacara)," John Wiley & Sons, Inc., Hoboken, NJ, USA. (2007)

[55] Stanley S. C. Chim, and Gordon S. Kino, "Three-dimensional image realization in interference microscopy," Appl. Opt. 31, 2550-2553 (1992)

[56] Park M, and Kim S, "Direct quadratic polynomial fitting for fringe peak detection of white light scanning interferograms," Opt. Eng. Vol. 39, Issue 4, 952-959 (2000)

[57] http://www.polytec.com/us/products/surface-metrology/

[58] https://www.youtube.com/watch?v=3ANWwqM4DSo

[59] Zahid Yaqoob, Jigang Wu, and Changhuei Yang, "Spectral domain optical coherence tomography: a better OCT imaging strategy," BioTechniques, Vol. 39, No. 6, pp. S6-S13 (2005)

[60] Wolfgang Drexler, and James G. Fujimoto, "Optical Coherence Tomography – Technology and Applications," 1st edition, Springer (2008)

[61] Maciej Wojtkowski, Vivek J. Srinivasan, Tony H. Ko, James G. Fujimoto, Andrzej Kowalczyk, and Jay S. Duker, "Ultrahigh-resolution, high-speed, Fourier domain optical coherence tomography and methods for dispersion compensation," Opt. Express 12, 2404-2422 (2004)

[62] Barry Cense, Nader A. Nassif, Teresa C. Chen, Mark C. Pierce, Seok-Hyun Yun, B. Hyle Park, Brett E. Bouma, Guillermo J. Tearney, and Johannes F. de Boer, "Ultrahigh-resolution high-speed retinal imaging using spectral-domain optical coherence tomography," Opt. Express 12, 2435-2447 (2004)

[63] http://www.optotune.com/applications/ophthalmology