ANALYSIS OF SENSITIVITY AND RESOLUTION IN PLASMONIC MICROSCOPES

Suejit Pechprasarn, BEng.

Thesis submitted to

The University of Nottingham for the degree of

Doctor of Philosophy

July 2012

Table of contents

CHAPTER 1: Introduction

CHAPTER 2: Background and literature review

CHAPTER 3: Methodology and validation

CHAPTER 4: SPR microscopy

CHAPTER 5: Quantitative analysis of SP microscope performance using RCWA

CHAPTER 6: SPR confocal microscopy

PAGE

CHAPTER 7: SP imaging using a spatial light modulator

CHAPTER 8: Discussion, conclusion and future work

PAGE

Acknowledgement

It would not have been possible to write this doctoral thesis without the help and support of the kind people around me, to only some of whom it is possible to give particular mention here.

First and foremost I would like to thank my supervisor Professor MG Somekh. It has been an honour to be his PhD student. I appreciate all his contributions of time, ideas, and funding to make my PhD experience productive and stimulating. The joy and enthusiasm he has for the research was contagious and motivational for me, even during tough times in my PhD.

I would like to thank the Department of Electrical and Electronic Engineering and the International Office of the University of Nottingham for providing the funding for this work.

I would also like to thank Professor Ian Harrison, Professor Paul O’Shea and Dr. Mark Pitter for their guidance and supervision during the first two years of my PhD. I appreciated my laser and optical system alignment training from Dr Richard Smith and Dr Ryan Qian. It has been an honour to work at the Institute of Biophysics, Imaging and Optical Science (IBIOS). I would like to thank everyone in IBIOS, especially Miss Bei Zhang for her valuable discussions and experimental results to back this thesis up and Mrs Kelly Vere for proof reading this thesis.

I would like to thank Professor Aaron Ho of the Chinese University of Hong Kong for providing facilities which enabled me to carry out a great deal of the work presented in this thesis.

I also acknowledge valuable discussions and support from Dr. See Chung Wah.

Lastly, I would like to thank my family for all their love and encouragement. For my parents, Ms Jantana Taweekijworachai and Mr Prasit Pechprasarn who raised me with a love and supported me in all my pursuits.

Suejit Pechprasarn

Conferences & Papers

Papers

- Zhang B, **Pechprasarn S**, Zhang J, Somekh MG. 2012. “Confocal surface plasmon microscopy with pupil function engineering”. *Opt. Express* 20:7388-97
- **Pechprasarn S** and Somekh MG. 2012. "Surface plasmon microscopy: resolution, sensitivity and crosstalk", *Journal of Microscopy*. doi: 10.1111/j.1365-2818.2012.03617.x
- Zhang B, **Pechprasarn S**, Zhang J, Somekh MG. “SP measurement with confocal microscope: experiment and simulation” (in preparation to publish)
- Somekh MG and **Pechprasarn S,** “Surface plasmon microscopy for biological applications” (in preparation to publish)

Conference Proceedings

- Zhang J, Huang Y, **Pechprasarn S**, Pitter MC, Somekh MG, “Thin gold films as contrast agents and their potential applications”; 22-26 May 2011 ICM—International Conference Centre Munich, Germany European Conferences on Biomedical Optics.
- **Pechprasarn S**, Smith RJ, Pitter MC, Somekh MG, “Understanding surface wave and surface plasmon imaging: through rigorous diffraction theory” 5th EOS Topical Meeting on Advanced Imaging Techniques, 29 July to 02 July 2010, Ramada-Treff Hotel Regina-Titlis, Engelberg, Switzerland.
- Somekh MG and **Pechprasarn S**; “Surface Plasmon microscopy: resolution vs. sensitivity”; Photonics Global Conference 2010, 14-16 December 2010, Singapore

- **Pechprasarn S** and Somekh MG; “Analysing surface plasmon microscopy with rigorous diffraction theory”; Functional Optical Imaging (FOI), 2011, Ningbo, China. 978-1-4673-0451-1/11 IEEE
- Zhang B, **Pechprasarn S** and Somekh MG; “Confocal surface plasmon resonance microscopy with pupil function engineering” ; Functional Optical Imaging (FOI), 2011, Ningbo, China. 978-1-4673-0451-1/11 2011 IEEE

Abstract

Surface plasmons (SP) are guided electromagnetic wave propagating along the surface of metal. The properties of SP are affected by the material attached to the metallic surface so they can be used as a very sensitive sensor capable of detecting the deposition of subnanometric layers of dielectric. SP has been widely investigated for biosensor applications and the theory is well established.

Although SP sensors have been well studied, integrating the SP to a microscope is a relatively young field. Since the SPs are surface waves; microscopy techniques to optimise the SP microscope performance will require totally different techniques to non-surface wave microscopy.

This thesis develops a theoretical framework to understand different types of SP microscope setups through the rigorous diffraction theory. The framework analyses the diffraction process through rigorous wave coupled analysis (RCWA) and a software package processes the diffracted orders to recover the microscope response for a range of different systems. In this thesis I will investigate the non-interferometric SP microscope, interferometric SP microscope and confocal SP microscope. I will show that the non-interferometric system exhibits a trade-off between lateral resolution and sensitivity, where an image obtained with a good contrast will have low lateral resolution. In order to get around the trade-off, the interferometric system can be employed; however, the main challenge for the interferometric setup is its optical alignment. I will show that a confocal SP microscope, which has been developed as a part of this thesis, can simplify the complexity of the interferometric system and give similar measurement performance.

For the interferometric and confocal systems, the SP measurements are normally carried out through the interference signal, which is interference between a reference beam and the SP. I will suggest a method to extract SP propagation parameters from

the interference signal by employing a spatial light modulator and also show that the SP propagation parameters do not only give us some insight to the SP effect for the interferometric system, but also gives us a new imaging mode to improve the resolution.

List of figures

CHAPTER 2 **PAGE**

CHAPTER 3 **PAGE**

CHAPTER 4 **PAGE**

PAGE

PAGE

CHAPTER 5

PAGE

PAGE

CHAPTER 6

CHAPTER 7

PAGE

PAGE

CHAPTER 8

Appendix A

Appendix B

PAGE

Appendix D

List of Table

CHAPTER 2 **PAGE**

CHAPTER 3

CHAPTER 4

CHAPTER 5

Appendix B

Appendix C

PAGE

Appendix D

Chapter 1

Introduction

In this chapter, I will provide an overview of this thesis. I will begin by asking what are surface plasmons (SPs) and why do we need them? I will then explain the current stage of the SP microscopy research, explain the current problems and summarise the key objectives of the study.

1.1 What are surface plasmons and why do we need them?

In this section, I will summarise the key features of SP. More details can be found in chapter 2, which explains the background of SP and provides a literature review.

Surface plasmons (SP) are guided electromagnetic waves propagating along the surfaces of metal. The properties of SP correspond to the material attached to the metallic interface, so it can be used as a very sensitive sensor to detect changes at the metal interface for nano scale thickness (Raether 1988). SP can be excited with visible light using high refractive index coupling such as a prism (Kretschm.E & Raether 1968); therefore optical science has made use of SP by using them as a sensor. SP sensor research has been growing massively and the theory of the SP phenomenon is well established (Homola 2008).

The key advantages of the SP when used as a sensor are its sensitivity, label free technique and that the measurement is done in real time.

Although SP sensor has been well studied, integrating the SP to a microscope is a relatively young field. A good example of where we require SP is if we want to image a sample with 1 nm coating sample.

Since the SP is surface wave, therefore it means that microscopy techniques to optimise the SP performance will require totally different techniques from the non-surface wave microscopy. The first attempt to develop SP microscope was by Yeatman and Ash in 1987 (Yeatman & Ash 1987), where they excited SP and took the images through a prism. They found that lateral resolution of images obtained from the SP microscope appears to conflict with contrast of the image.

Since the *k-vector* of the SP is greater than the *k-vector* of light in free space, it is necessary to increase the *k-vector* of the incident wave in order to get the SP excitation. This can be performed either by:

a) using an incident medium with high refractive index, such as prism coupling or

b) grating

Here in this study, a high refractive index couplant was used. These requirements normally made the excitation of SP difficult; especially for optical configuration when a glass prism was required. Thanks to the advert of oil immersion high NA objective lenses, it is possible to excite SP in a more convenient fashion for microscopy application compared to using a prism. This makes SP microscopy more attractive as we can apply special techniques to enhance the resolution and sensitivity of the microscope setup.

In 2000, Somekh *et al* (Somekh *et al* 2000a) demonstrated that the scanning heterodyne interferometer SP microscope can optimize the trade-off between sensitivity and resolution. Although there are some experiments provided by a number of authors to illustrate the use of SP for imaging, there is still the need for a theoretical framework for SP microscope in order to understand the SP microscope mechanisms and also to develop new techniques to improve them.

1.2 Current issues in SP microscopy

In this thesis, I will discuss two types of objective lens based SP microscopes, namely non-interferometric and interferometric SP microscopes.

It has been found that where in the non-interferometric SP microscope; there is a trade-off between resolution and contrast, which means that an image with a good resolution will have a poor contrast and image with a good contrast will have a poor lateral resolution. The interferometric system was then developed and showed that it can improve the trade-off by detecting the interference signal using a heterodyne interferometer. Although the heterodyne system works very well in terms of resolution and sensitivity, the complexity of the optical setup is challenging, particularly when aligning the system (Zhang *et al* 2012).

In addition, another challenge in the research was that from the computer simulation point of view. Although Fresnel equations can be used to accurately simulate SP effects for uniform metal and uniform sample case, before this study we did not have a good tool to calculate SP responses from a non-uniform sample. Therefore the challenge here was to develop a software package that can be used to accurately calculate SP responses for a non-uniform sample in order to quantify the resolution and sensitivity for different types of SPR microscope setups.

One of the main issues here is that most of the SP measurements reported were normally carried out using microfluidic channel arrays for a high throughput measurement without well separated channels. Crosstalk means that the measurement carried in a particular channel is also affected by the SP from other channels. As SP are a surface wave, it will propagate and transmit through the other channel as well as being reflected back into its own channel. In this thesis, I will show that the crosstalk plays a crucial role in the measurement of SP, where it determines the accuracy of the measurement as well as the resolution when imaging.

1.3 Research objectives

The objectives of the study are to solve the problems outlined above.

The first task is to develop a robust modelling tool that is capable of modelling accurately SP microscope responses for different microscope configurations. The details of the software package are described in chapter 3 and validated by comparison with experimental results and results obtained from other simulation techniques in the literature.

The second objective is to develop a theoretical framework to describe plasmon microscopes. The detail of different microscope responses can be found in chapter 4.

The third objective is to develop a quantitative measurement method to quantify the performance of the microscopes. The quantitative measurement does not only enable us to compare the performance of each SP microscope configuration, but also gives us some insight into the accuracy of the measurements; making it possible to make localized measurements on non-uniform samples. The detail of quantitative measurement is provided in chapter 5.

The fourth objective of this study is to reduce the complexity of the SP instrumentation. We (Zhang *et al* 2012) have recently demonstrated that the confocal microscope can be used to performance the same function as the heterodyne interferometer with a much simpler optical configuration. Results obtained by the confocal system are similar to the results obtained from the heterodyne system. These details are provided in chapter 6.

The fifth objective is to develop SP measurement techniques that can be used to reduce the need for mechanical scanning during data acquisition and also improve resolution. In this thesis, I will suggest a few techniques using either amplitude or phase spatial light modulators to measure SP response without mechanical scanning; this will, of course, improve the measurement speed and microphonic noise due to mechanical vibration. Details of this are provided in chapter 7.

1.4 Methodology

My main contribution in this research is on the computer model of the system. The simulation results are compared with experimental results that were obtained by my colleague Bei Zhang. More details about experimental setup will be available in her thesis.

There are two types of simulation software used in the study; firstly Fresnel equations, which were used for uniform sample simulation; and rigorous coupled wave analysis (RCWA), which was used for non-uniform sample. The details of the software are provided in chapter 3.

1.5 Thesis outline and purposes of each chapter

The thesis is structured and presented as the followings:

Chapter 1 Introduction: in this chapter I provide an overview of the project, explain the key problems and suggested some ways to solve them.

Chapter 2 Literature review and background: in chapter 2, I will review the literature and summarise the features and properties of SP. The theoretical background and detailed equations of SP are provided in the chapter. More importantly the current research activities for SP sensors and SP microscopes are summarized.

Chapter 3 Methodology and validation: in chapter 3, I will show how the microscope simulation software was implemented and compare the simulation techniques available in order to show the reason why we are of a firm view that RCWA is a suitable technique for the study. The microscope software was validated by comparing the simulation results with experimental results and other simulation techniques and results obtained from the literature.

Chapter 4 SP microscopes: in chapter 4, I will provide some general quantification for SP sensors and calculate their theoretical sensitivity. Some linescan images and back focal plane images are presented and compared with experimental results. The performance of the non-interferometric and interferometric systems will be compared. I will confirm that the resolution of the interferometric system is superior to the non-interferometric one. The chapter ends with description of techniques to improve the resolution and sensitivity of the two types of SP microscopes.

Chapter 5 Quantitative analysis of SP microscope performance using RCWA: in chapter 5, I will provide a quantitative measurement method to compare the performance of the non-interferometric and interferometric microscopes as they have totally different measurement processes, so we need a means to compare them. This quantitative measurement method also provides us with a new way of looking at the resolution problem, where we strongly believe that the most important factor that limits the accuracy of measurement and resolution for imaging is crosstalk in the interferometric system. We will also show that one of the main key advantages of the interferometric system is that we can measure the effect of SP in a confined region.

Chapter 6 Confocal SPR microscope: in this chapter, I will provide theory and experimental details for the confocal SP microscope, which can reduce the complexity of optical configuration compared to the heterodyne microscope. Although there is a difference between the heterodyne system and the confocal system which is the fact that the signal detected in the confocal system is much smaller than the signal obtained in the heterodyne system, we will show that by engineering the pupil function, we can enhance the amplitude of detected signal. Effects of confocal pinhole sizes will be also discussed. More importantly, it will be shown that the confocal system can give the same measurement performance as the heterodyne interferometer.

Chapter 7 SP imaging using spatial light modulator: in chapter 7, I will present several modulation techniques to extract key SP propagation parameters using SLM. The algorithm does not only enable us to measure the SP profile without mechanical scanning, but also enables us to measure the SP profile with less crosstalk from adjacent regions, which means it gives us better resolution.

Chapter 8 Discussion, conclusion and future works: in this chapter, I will summarise all the findings of this study and suggest some interesting future work that can be exploited using the simulation tool developed in the study.

Chapter 2

Background and literature review

In this chapter, I will provide background theory of surface plasmon resonance (SPR) and review literature in this field in order to summarise findings and development in this field of research. The principal aim of this chapter is to point out the current issues and key challenges in SP microscopy.

2.1 Surface Plasmon resonance theory

The first recorded observation of surface plasmon resonance was made by Wood (Wood 1902), who observed dark bands when light illuminated a metal grating. The subsequent explanation of these observations was attributed to surface plasmon resonance. Research in surface plasmons and plasmonics has grown massively, especially over the past few decades; scientists and engineers widely have made use of surface plasmon resonance to detect chemical or biological changes within a small layer and have also developed many advanced techniques to optimise its sensitivity (Roh *et al* 2011). Surface plasmons arise from electromagnetic wave coupling phenomenon on noble metals and the theoretical framework is very well established (Raether 1988); however there are still many potential uses of the SPR to be exploited such as subwavelength optics (Barnes *et al* 2003), data storage (Wan *et al* 2010) , microscopy (Somekh *et al* 2009), bio-photonics and quantum plasmonics (Barnes *et al* 2003; Bergman & Stockman 2003). In this chapter, I will review the literature that relates to this research and summarise the findings in this field that are particularly relevant to the present thesis.

2.1.1 What is SPR?

Surface Plasmon Resonance (SPR) or surface plamon polaritons (SPP) are optically excited surface plasmons. So I will shorten the title to simply surface plasmons. Surface plasmons (SP) are guided electromagnetic waves propagating along the surface of noble metals (Raether 1988). The propagation properties of the SP depend on the surrounding ambient medium, in other words, it is sensitive to the properties of the surfaces of the noble metals. In addition to metallic films, the effects can be found in metallic nano-sized particles (Grabar *et al* 1995; Mulvaney 1996). These are also plasmonic effects but are usually referred to as localized surface plasmons.

In this research, the main objective is to understand the effect of the SPR on the sensitivity and lateral resolution of microscopes.

2.1.2 Properties of SP

After Krestchmann (Kretschm.E & Raether 1968) and Otto (Otto 1968) clearly demonstrated that SPR could be excited by visible light using prism based configurations, many theoretical works to explain this phenomenon have been widely investigated (Cardona 1971; Chen *et al* 1976).

In 1988, an elegant theoretical approach to explain the SPR effect was published by Cardona (Cardona 1971) and described in Raether's book (Raether 1988). Here the SPR phenomenon is explained in terms of the Fresnel reflection coefficients.

Here we will start to explain SPR effect using the well know Fresnel's equations (Born & Wolf 1999), which can be derived from Maxwell's equations (Reitz 1993). The Fresnel's equations for two layers of uniform materials as depicted in figure 2.1 are given by:

$$t_p = \frac{\cos(\theta_i)\sin(\theta_t)}{\sin(\theta_i+\theta_t)cos(\theta_i-\theta_t)} \qquad [2.1]$$

$$t_s = \frac{2\cos(\theta_i)\sin(\theta_t)}{\sin(\theta_i+\theta_t)} \qquad [2.2]$$

$$r_p = \frac{\tan(\theta_i-\theta_t)}{\tan(\theta_i+\theta_t)} \qquad [2.3]$$

$$r_s = \frac{\sin(\theta_i-\theta_t)}{\sin(\theta_i+\theta_t)} \qquad [2.4]$$

where t_p, r_p are the reflected and transmitted electric fields of the *p-polarised* light respectively. t_s, r_s are the reflected and transmitted electric fields of the *s-polarised* light respectively. θ_i, θ_t are incident and refracted angles of the incident plane wave.

The *p-polarisation* (or transverse-magnetic : *TM*) means a linear polarised plane wave with its electric field pointing in the plane of incidence, whereas the *s-polarisation* (or transverse-electric : *TE*) means an linear polarised plane wave with its electric field pointing perpendicular to the plane of incidence.

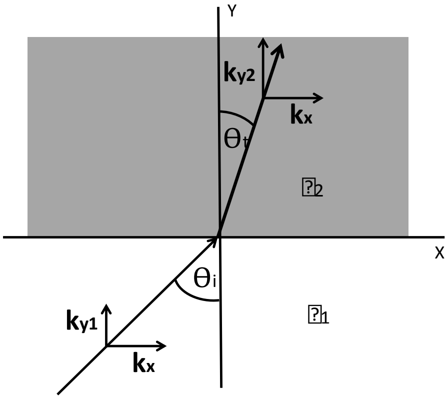


Figure 2.1 shows the two material layers system Fresnel’s equations

Firstly for the reflectance; if $\theta_i + \theta_t = \frac{\pi}{2}$, this makes the denominator of equation [2.3] r_p become infinity. This makes the r_p term equal to zero. This condition is the Brewster angle, where the reflectance becomes zero for the *p-polarised* light. On the other hand, this condition will not affect the r_s.

The second condition is where $\theta_i - \theta_t$ approaches $\frac{\pi}{2}$, this instead makes the r_p term become infinite or a pole in the reflection coefficient causing the resonance in r_p. If $\theta_i - \theta_t$ approaches $\frac{\pi}{2}$, then $\cos\theta_i \approx -\sin\theta_t$ and $\tan\theta_i \approx \frac{k_x}{ky1} \approx -\frac{n_2}{n_1}$. In contrast, this condition will not affect the r_s term, in other words, the SP cannot be excited by *s-polarised* wave; this becomes one of the key criteria to excite the SP that the SP can only be excited with *p-polarised* wave.

It is interesting to note that the condition $\theta_i - \theta_t = \frac{\pi}{2}$ cannot be met in the real physical scenario as the θ_t will become imaginary before satisfying the condition above. This can be explained by poles and zeros of reflection coefficients; the details can be found in (Tamir *et al* 1985). This also indicates that with the present optical configuration the SP cannot be excited with only a propagating wave component, but instead it will always come with the imaginary part, which provides attenuation constant.

We can then rewrite the wave vector *k* as:

$$k_1^2 = k_x^2 + k_{y1}^2 = k_x^2 + k_x^2\frac{\varepsilon_1}{\varepsilon_2} \text{ and } k_2^2 = k_x^2 + k_{y2}^2 = k_x^2 + k_x^2\frac{\varepsilon_2}{\varepsilon_1} \quad [2.5]$$

$$k_x = \frac{\omega}{c}\sqrt{\frac{\varepsilon_1\varepsilon_2}{\varepsilon_1+\varepsilon_2}} \quad [2.6]$$

$$k_{yi} = \frac{\omega}{c}\sqrt{\frac{{\varepsilon_i}^2}{\varepsilon_1+\varepsilon_2}} \quad [2.7]$$

where *k* is the wave vector $\frac{2\pi}{\lambda} = \frac{\omega}{v}$.

If the second material is a metal layer where it consists of an abundant number of free electrons, the relative permittivity ε_2 will have negative real part (Johnson & Christy 1972) and $Re(\varepsilon_2) > -\varepsilon_1$. For the metal case, the term k_{yi} for both medium becomes complex; in other words, they become evanescent, whereas k_x remains a real number. These wave vectors can be interpreted in a real physical situation as an electromagnetic which propagates along the *x* direction at the interface between the two media with exponentially decaying field strength in the y direction on both sides of the interface as schematically shown in figure 2.2. Note that the k_x value that makes k_{yi} become imaginary, is the so called wave number of surface plasmon k_{sp}. A physical description of SP that is easier to appreciate; they can be thought of as collective oscillations of free electrons on the metal transferring the excessive energy parallel to the metal interface.

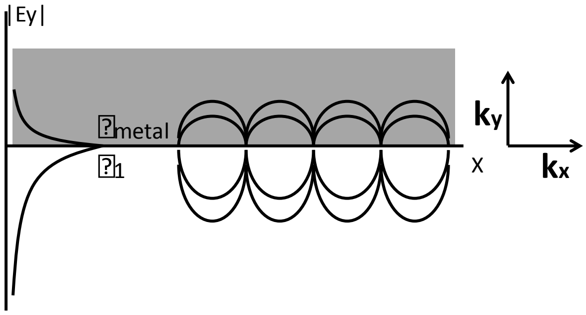


Figure 2.2 shows evanescent tails and electric fields for SP propagation

It is important to note that the SP waves can only be generated with an oblique angle that matches the incident k_x and k_{sp}. In other words, we cannot excite SP from normal incidence unless some other means to match the incident *k-vectors* in the *x* direction are provided, such as grating coupling structure (Ritchie *et al* 1968).

In addition, k_{sp} is greater than the free space *k-vector* (Kretschm.E & Raether 1968), therefore the other requirement is to have k_x greater than the free space *k-vector*. This implies that we cannot excite SP from air. These requirements actually lead to some challenges in optical experimental setup, which will be explained later.

Propagation depth

Propagation depth is an important property of SPR, which is defined as the distance for the magnitude of SP intensity decay along the normal direction of the interface plane by a factor of $1/e$.

From equation [2.7], we can calculate the propagation depth as:

$$Propagation\ depth = \frac{1}{k_{yi}} = \frac{c}{\omega}\sqrt{\frac{\varepsilon_1+\varepsilon_2}{\varepsilon_i^2}} \quad [2.8]$$

In order to get some feeling about the SPR technique, I will start by putting some numbers that depict the real scenario into equation [2.8].

$\lambda = 633\ nm$ so $\omega = 2.98 \times 10^{12} s^{-1}$ at a gold water interface with $\varepsilon_{water} = 1.77$ and $Real(\varepsilon_{gold}) = -11.5$ corresponding to the particular wavelength. We can obtain the penetration depths:

$$\frac{1}{k_{y\,water}} = 177nm$$

$$\frac{1}{k_{y\,gold}} = 27nm.$$

Surface sensitivity of SPR

The key reason why SP are very attractive for studying biological and measurement chemical changes is due to the fact that any changes at the surface of the metal layer within the propagation depth of the SP lead to the changes in propagating wave vector k_x. Therefore this technique is a useful candidate for label free experiments on biological, chemical or surface science experiments on the nano scale.

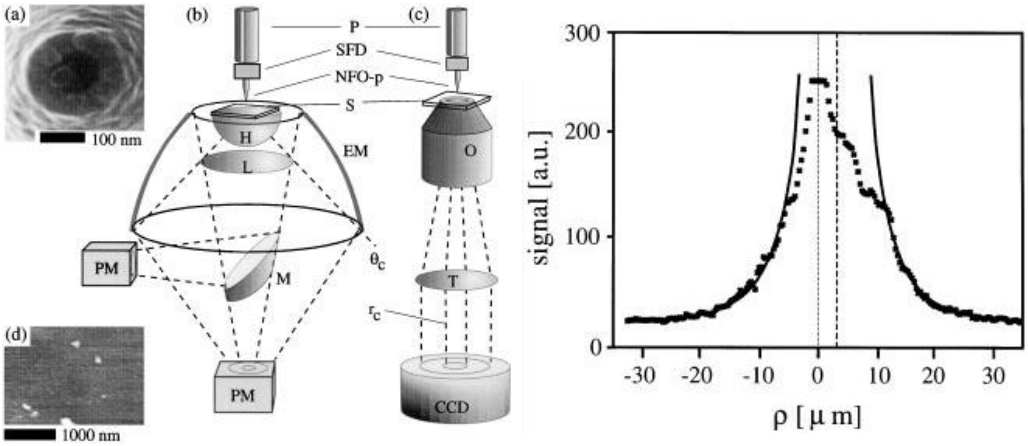
It is an interesting idea that if we can control the SP propagation depth by some technique such as confining the field strength, this will enable us to have a very sensitive sensor over a very thin layer such as cell membranes.

Propagation damping and plasmon dip

Having discussed that for metals, the real part of the complex permittivity is negative this leads to the propagation of SP, however real metals also have the imaginary part of the permittivity (Johnson & Christy 1972). In this case, the k_{sp}

now becomes a complex number $k_x = k_x' + ik_x''$. The k_x' is the propagating vector along the x axis, but the k_x'' term is the damping factor, which includes both absorption radiation and internal absorption losses. The later might be thought of an ohmic loss of the SP in the metal dissipating as heat (Bosenber.J 1971).

In near field optics, the heat loss and the SP propagation can be measured directly with a probe type experiment (Hecht *et al* 1996; Novotny *et al* 1997), this is where the near field is a more direct approach to measure the SP than the far field. However the drawback of near field optics is, of course, that it is not so convenient for biological experiments because the near field systems require an optical probe placed very close to the sample.

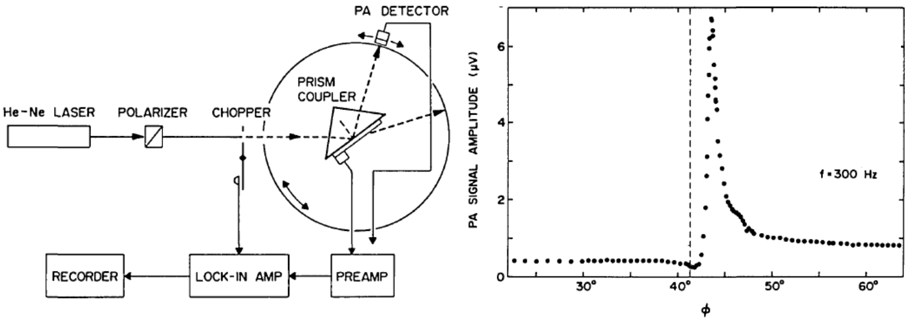


a) b)

Figure 2.3 shows a) near field optical setup b) SP field strength measured by the setup (Hecht *et al* 1996).

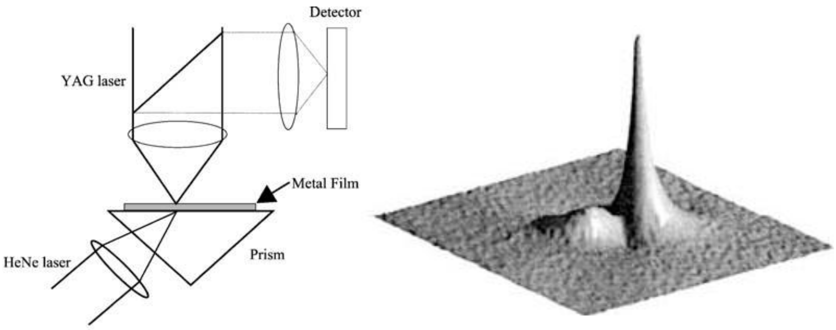
In far field optics, the heat can be measured experimentally by photo-acoustic surface modulation techniques; such as chopping an incoming light to periodically heat the metal film and measure the microphonic changes (Inagaki *et al* 1982; Rothenhäusler *et al* 1984; Talaat *et al* 1985).

Two visible light wavelengths can also demonstrate the same effect on a simple experimental setup. The SP were generated using a 633 nm HeNe laser. The gold film was heated periodically by a YAG green laser, where the laser beam was modulated to a convenient frequency of several kHz. The green laser altered the surface profile of the gold causing the SP propagation condition to change. This perturbation caused the SP to scatter back into the air (Velinov *et al* 1999). The image of SP propagation obtained using optical modulation method is shown in figure 2.5. There were two lobes obtained in the experiment; the left hand side lobe in the figure 2.5b was the SP wave.



a) b)

Figure 2.4 shows a) far field photo-acoustic surface modulation optical setup b) SP field strength measured by the setup (Inagaki *et al* 1982).

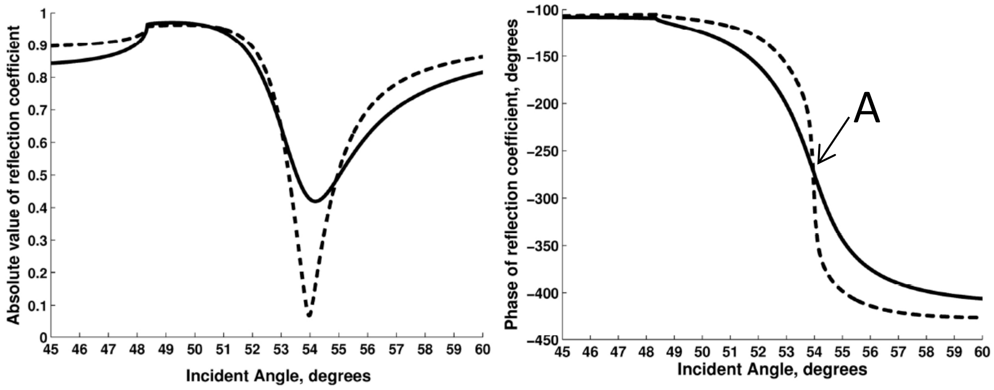


a) b)

Figure 2.5 shows a) far field 2 wavelengths surface modulation optical setup b) SP field strength measured by the setup (Velinov *et al* 1999).

The reflectance curves for “real” gold and “ideal” gold are different as can be seen by comparing figure 2.6 with figure 2.7. This arises because the permittivity of “ideal” gold is real and negative whereas there is an additional imaginary part with “real” gold. The real gold shows that the amplitude of reflectance dramatically dropped down at an angle, so called ‘plasmon dip’ and ‘plasmon angle θ_p’ respectively.

The variation of reflectance with incident angle from a metallic film can provide us with considerable insight of the SP phenomenon and its behaviour (Azzam 1977). The plasmonic dip indicates the angle where the strongest SP coupling occurs; this leads to the highest ohmic loss in the metal therefore it appears to be the minimum dip position on the reflectance curve. In other words, the dip can tell us how strongly the SP is excited; the lower plasmonic dip magnitude is the stronger SP excitation.

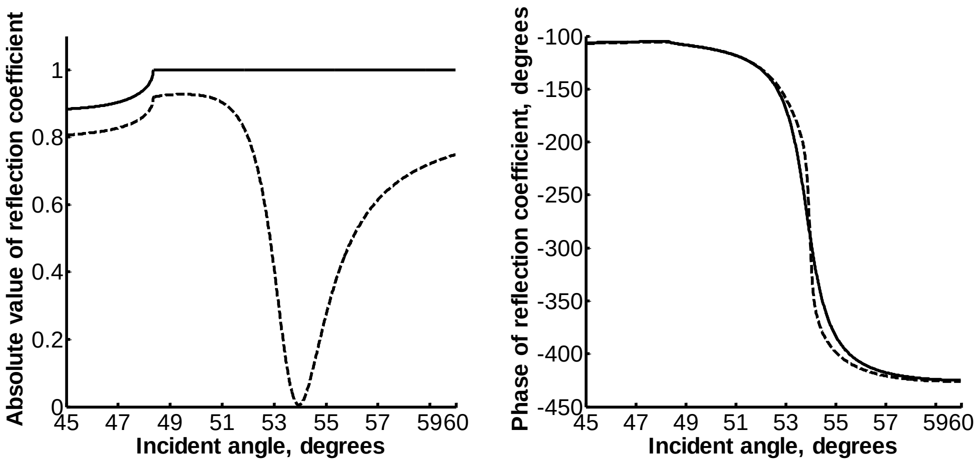


a) b)

Figure 2.6 shows a) Modulus of reflection coefficient from gold sandwiched between two half spaces. *Wavelength=633nm p-polarisation*, $n_0=1.78$, $n_{gold}=0.17+3.52i$, $n_{lower}=1.33$. Solid curve 36nm gold thickness, dashed curve 46nm gold thickness and b) Phase of reflection coefficient corresponding to case of a) (Pechprasarn & Somekh 2012)

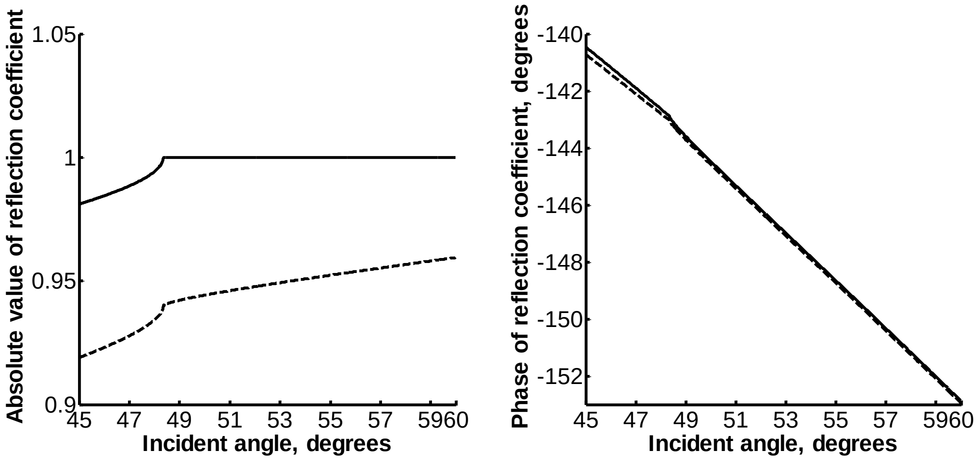
From figure 2.6a, it can be interpreted that the 46 nm has stronger SP excitation as the plasmonic dip is much lower than the 36 nm case. In addition, the plasmon angles for the two cases are at virtually the same position. This indicates that different thicknesses of metal cause only a change magnitude of SP excitation, not the plasmon angle.

The ideal gold does not show the dip pattern. This does not mean that there is no plasmonic excitation. The amplitude of the reflectance curve might not always be a good indicator of existence of the SP. The reduction in the modulus of reflection coefficient only shows us how much energy is dissipated as heat in the metal film. On the other hand, the phase can act as the real indicator to determine the existence of SP.



a) b)

Figure 2.7 shows a) Modulus of reflection coefficient from gold sandwiched between two half spaces. *Wavelength=633nm p-polarisation,* $n_0=1.78$, $n_{lower}=1.33$, 46 nm uniform layer of gold. Solid curve ideal gold with $n_{gold}=3.52i$, dashed curve real gold with $n_{gold}=0.17+3.52i$ and b) Phase of reflection coefficient corresponding to case of a).



a) b)

Figure 2.8 shows a) Modulus of reflection coefficient and b) phase of reflection coefficient from gold sandwiched between two half spaces. *Wavelength=633nm s-polarisation,* $n_0=1.78$, $n_{lower}=1.33$, 46 nm uniform layer of gold. Solid curve ideal gold with $n_{gold}=3.52i$, dashed curve real gold with $n_{gold}=0.17+3.52i$.

It is clear from figure 2.8 that there were no SP excited from *s-polarised* incidence. The reason that *s-polarisation,* where its magnetic field is parallel to the metal interface, cannot excite SP is because the electric field of the *s-polarisation* points perpendicular to the direction of incident plane; therefore its electric field cannot interact with the free electrons at the metal film. Recently, there has been an interesting idea to observe magnetic surface plasmons where it is, of course, excited by *s-polarised* wave (Ermolaev *et al* 2011; Shen *et al* 2011). This requires a large negative magnetic permeability material, which has recently been available.

Propagation length

Propagation length or coherence length is another property of SP defined by the distance for the magnitude of SP intensity decay along propagation direction by a factor of $1/e$.

$$Propagation\ length = \frac{1}{2k_x''} \qquad [2.10]$$

The propagation length is dependent upon many factors such as types of material and their shapes.

Another approach to approximate the propagation length is to calculate the gradient of the phase change around the critical angle, where it is directly proportional to the propagation length of the surface waves. This effect is essentially the Goos Hänchen shift (Born & Wolf 1999) analogous with the Schöch displacement (Brekhovskikh *et al* 1960) in ultrasonics. The propagation length can be mathematically described as it is proportional to $d\phi/\mathrm{d}k_x$ (McGuirk & Carniglia 1977), where ϕ is the phase of the complex reflection coefficient. In other words, the gradient of the phase indicates how far the SP wave can propagate.

From figure 2.6b, it can be clearly seen that SP in the 46 nm case have a longer propagation length than the 36 nm case since the gradient of the phase of reflection coefficient of the 46 nm case is steeper than the 36 nm case. The propagation dip can indicate the strength of the SP excitation; this also directly relates to how far it can propagate along the metal interface.

If one designs an SPR based sensor without taking the propagation length of the SP into account, the sensor will be affected by excitation from adjacent regions, this is the so called crosstalk. One of the main objectives of this thesis is to quantify the crosstalk (see chapter 5 for more details). For example, SP on gold are excited at $\lambda = 633\ nm$, the propagation length is around 7 microns. Therefore the requirement when designing this SPR device is to have the micro-fluid channel or binding pad considerably larger than 7 microns x 7 microns in order to reduce the crosstalk effect.

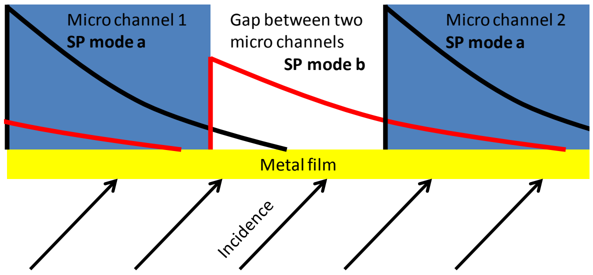


Figure 2.9 shows a sensor structure that has two SP modes of interfering with each other

Figure 2.9 above gives an illustration of the crosstalk if the detection channel is not properly separated wide enough, the measurement in both channels would be interact with the SP mode b, where we intend to measure only the responses from only SP mode a. This will, of course, lead to some errors in the measurement.

Phase change with incident angle around plasmonic angle

From figure 2.6b and 2.7b, we can see that the phase variation of the *p-polarisation* around the plasmon angle changes by approximately 2π radians. The position A on figure 2.6b where the phase varies much rapidly with incident k_x represents at the plasmon angle θ_p. This rapid phase change is a more fundamental feature of SP excitation than the plasmonic dip in the modulus of the reflection coefficient. As discussed, the phase variation, in fact, gives real evidence of the existence of the surface waves. The phase variation also allows us to have a sensor with better sensitivity and a new mode operating mode for SPR microscopes, which will be discussed in the later sections.

Field enhancement

In this section, the Fresnel's equations are extended from two layer case to three layer case. Therefore, the equations are, of course, more complicated than the two layer case. The equations of the three layer and four layer cases can be found in (Reitz 1993). However the idea of how SP are generated is the same. It is interesting to see what will happened if metal of different thickness is used.

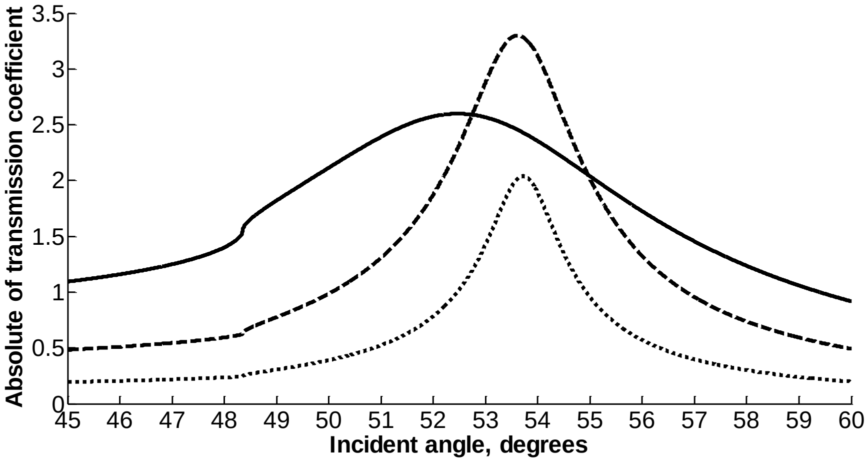
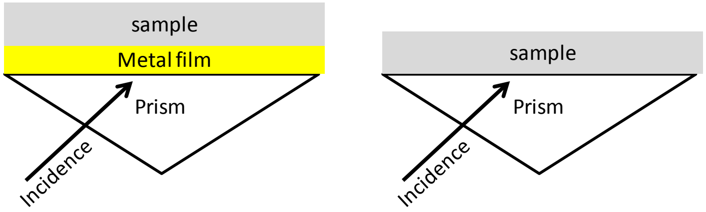


Figure 2.10 shows modulus of transmission coefficient from gold sandwiched between two half spaces. *Wavelength=633 nm p-polarisation*, $n_0=1.78$, $n_{lower}=1.33$, uniform layer of gold. Solid curve gold 25 nm thick, dashed curve gold 50 nm thick and dotted curve gold 75 nm thick.

For the *p-polarised* wave, there is a large field enhancement at the interface between the metal and the final dielectric next to it as shown in figure 2.10. The field enhancement arises primarily from the component of the evanescent field normal to the interface between the metal and dielectric. On the other hand, for the *s-polarised* wave, the field exponentially decays in the metal so the field at the interface is small because the field is continuous across the interface as its electric field vector is parallel to the interface plane.

The sensitivity of SP arises partly from this field enhancement. The SPR experiment essentially looks like an attenuated total internal reflection (ATR) experiment except a metal layer is sandwiched between the prism and sample.



a) b)

Figure 2.11 shows a) SPR experimental configuration and b) Other surface wave experimental configuration

Of course, the large field enhancement does not contradict conservation of energy because the Poynting vector normal to the surface is imaginary.

The magnitude of the field enhancement is dependent on the thickness of the metallic layer. For the layer that is too thin, all SP waves are strongly dampened by the small thickness and cannot be coupled to transmit their energy through the final metal-dielectric interface. On the other hand, for a layer that is too thick, the SP wave will decay excessively within the metal film, so the field strength coupling onto the other side will be lower.

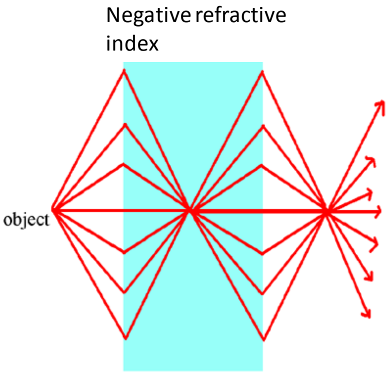


Figure 2.12 show negative refractive index Pendry lens : modified from (Pendry 2000)

An interesting idea to expand this concept was proposed by (Pendry 2000). He provided a theoretical prediction that it is possible to generate a perfect lens to generate an infinitely sharp focal spot by enhancing the evanescent wave components. His idea was based on negative refractive index material to couple both evanescent and propagating waves together at the focal point of the lens as shown in figure 2.12. The decay of the evanescent waves in the free space is exactly compensated by the gain in the material of negative refractive index, so the image contains high frequency evanescent wave components.

Reradiation and back scattering wave

If the evanescent field of an SP is perturbed, light will be scattered into propagating modes which can be detected with a conventional photo detector. This scattering and reradiation can be effected by a physical probe or by localized temperature changes (Velinov *et al* 1999).

Of course, similar scattering will occur when the surface over which the SP propagate are not perfectly smooth.

This property of SP is very important as it allows us to investigate direct responses from the sample. Yeatman and Ash explained that one key property of the SP to perform SPR imaging is the reradiation process; they explained that if the SPs are not reradiated, we would see virtually no contrast in far field detection. (Yeatman & Ash 1987)

2.1.3 Choice of metals

The key characteristics of metals in order to excite SP are:

a) $Re(\varepsilon_{metal}) < -\varepsilon_{first\ layer\ material}$ and
b) Low $Im(\varepsilon_{metal})$ to reduce the ohmic loss.

The suitable metals according to the criteria above are aluminium (Al), cadmium (Cd), copper (Cu), gold (Au), indium (In) , platinum (Pt) and silver (Ag) for the visible and infrared light spectrum. (de Bruijn *et al* 1992; Johnson & Christy 1972)

Alongside metal choice, we should also consider metal thickness as different thicknesses will have different field enhancement, propagation length and sensitivity

(Ekgasit *et al* 2005). Choices of metal with their thicknesses and incident wavelength for SP excitation are listed in appendix C.

The purpose of this thesis is to apply the science of SP to bioscience; therefore it is also important to understand how compatible the metals are when they interact with biological samples. It is important to design a bio-sensor so that it is bio-compatible to limit toxicity. This degrades the sensor, for instance, Ag_2S can be formed by Ag and sulphur compounds in the atmosphere, which can limit the usefulness of silver. This can be improved by adding a less toxic layer on the surface in contact with the analyte, for instance Ti on Ag (Anker *et al* 2008).

In practice, the noble metals such as gold and silver are excellent choices to excite the SP propagating along their smooth surfaces. The majority of SPR researchers used these metals (Eustis & El-Sayed 2006; Homola 2008). Aluminium also supports the SPR excitation; however the propagation length of the SP wave is quite short as it is strongly attenuated due to the large imaginary part of the Al permittivity. It has been shown that this short propagation length improves the lateral resolution of the images obtained from an Al based SPR microscope compared to a Au based SPR microscope because of the shorter propagation length (Giebel *et al* 1999).

2.1.4 Excitation of SP

Two ways to obtain a high incident *k-vector* are described below.

Kretschmann and Otto configurations

In 1968, Kretschmann (Kretschm.E & Raether 1968) and Otto (Otto 1968) clearly illustrated ways to excite SP using the visible light as shown in figure 2.13, so called Kretschmann's configuration and Otto's configuration respectively. Both methods employed a prism to satisfy the requirement that the incident *k-vector* is greater than that in free space.

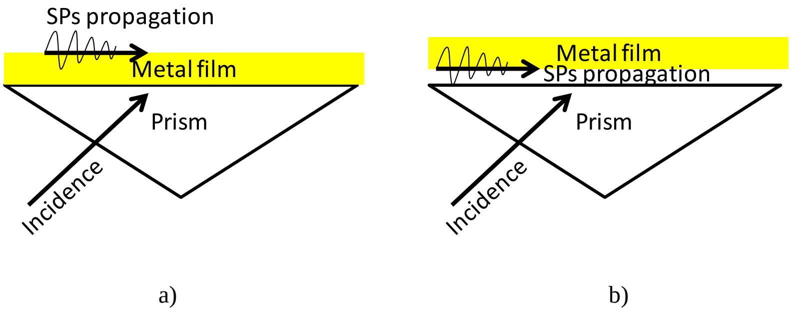


Figure 2.13 shows a) Kretschmann's configuration b) Otto's configuration

The differences between the two approaches are as follows. Firstly, prism position; for Kretschmann configuration, the metal film is attached to the prism and will, of course, be a feature of the prism, whereas the prism for Otto's configuration is placed above the metal film.

Secondly, they are designed for different purposes. The Kretschmann configuration is suitable for studying the SP effect above the metal, such as surface science, monitoring antibody-antigen reactions, cell attachment, pharmacology and toxicology (Cooper 2002; Liedberg *et al* 1993; Wink *et al* 1997). Therefore the majority of the SP experimental setups and sensors for biological and chemical experiments are Kretschmann prism based experiments.

In contrast, Otto configuration is suitable for studying the fundamental SP effects at the surface located in the evanescent field from the prism. The configuration can be used to study fundamental physics in bulk metal samples. The requirement to maintain a constant and small gap means that it is most commonly used in the infrared and to my knowledge it has not been used for biosensing. The configuration is well adapted so that the metal surface will not be damaged or touched by the prism. This configuration is important for studying (for example) surface phonon polaritons on a single crystal surfaces (Iwata & Mizutani 2010).

For Kretschmann configuration, the metal film, usually gold or silver, is deposited on the prism and the light is incident from the prism side; so that it hits the film through the prism at an angle close to plasmonic angle.

The prism based Kretschmann configuration is indeed very powerful from the sensor point of view, but from the microscope point of view, it does however, lack the spatial resolution required to image highly localised regions such as at a cellular level (Jamil *et al* 2008). Therefore, there have been a number of publications reporting the development of SP microscopes that have high lateral resolution and at the same time good surface sensitivity; this will be discussed in details in section 2.3.

Objective lens excitation

The main theme of the thesis is to study and build up a theoretical knowledge of objective lens excitation SP microscopes to explain imaging performance of such microscopes. Therefore it is worth discussing the relationship between the light incident on a microscope objective lens, the polarisation states and incident angles that illuminate the sample.

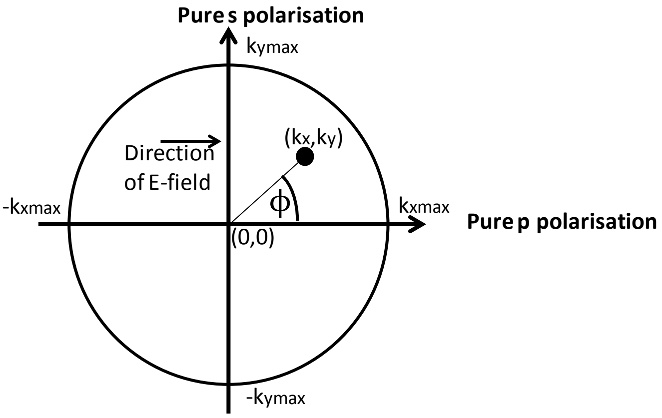


Figure 2.14 shows the bfp distribution of an objective

We will start with a back focal plane (bfp) diagram of an objective lens with a collimated beam fully covers the bfp as shown in figure 2.14. Each point on the bfp represents a plane wave and its k_x and k_y correspond to the incident angle and its azimuthal angle ϕ at the sample. If the polarisation of the incident beam on the back focal plane is linearly polarised, the polarisation state of the incident radiation varies from pure *p-polarised* to pure *s-polarised* at the sample surface corresponding to its azimuthal angle ϕ.

Definition of X,Y and R polarisations

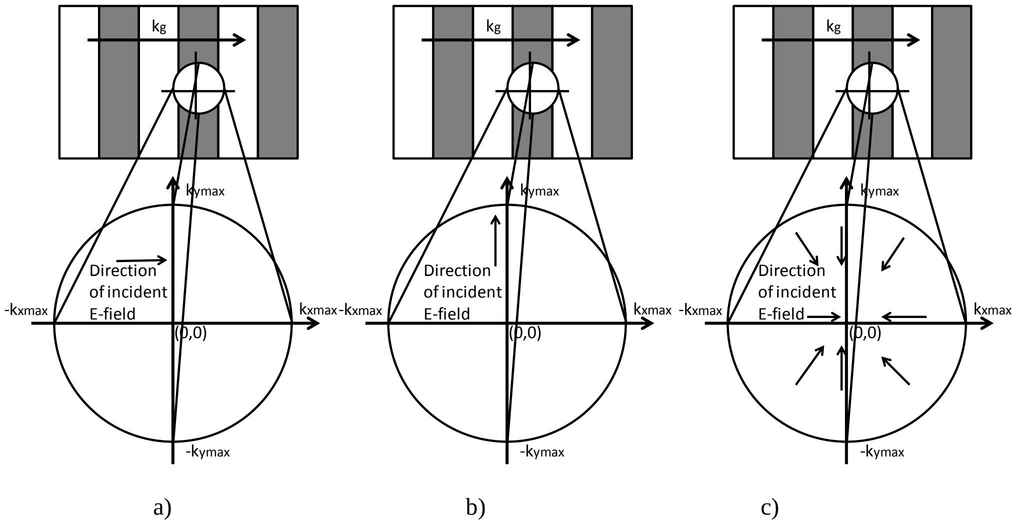
In this thesis, I will show that SP generated by the linearly polarised and radially polarised light interact differently with the sample structure including the sample orientation.

The sample modelled throughout this thesis is a grating assuming that the sample has its grating vector k_g pointing along the positive *x* direction as shown in figure 2.15.

For the case that all polarisations on the bfp are parallel with the direction of grating vector, from here on will be called *'x polarisation'*.

On the other hand, for the case that all polarisations on the bfp are normal to the grating vector will be called *'y polarisation'*.

For the case electric fields point along their azimuthal angles, in other words, purely *p polarisation* on the bfp, will be called *'r polarisation'*.



a) b) c)

Figure 2.15 shows a) *x polarisation* b) *y polarisation* and c) *radial polarisation*

Recently, a number of researchers (Moh *et al* 2008; Zhang *et al* 2010) have been interested in the radial polarisation; this is because it gives pure *p-polarisation* for all azimuthal angles, in other words, this polarisation will give the strongest excitation of SPR. The other reason is that devices to produce radial polarisation conveniently are now commercially available.

SP excitation mechanism using objective lens excitation

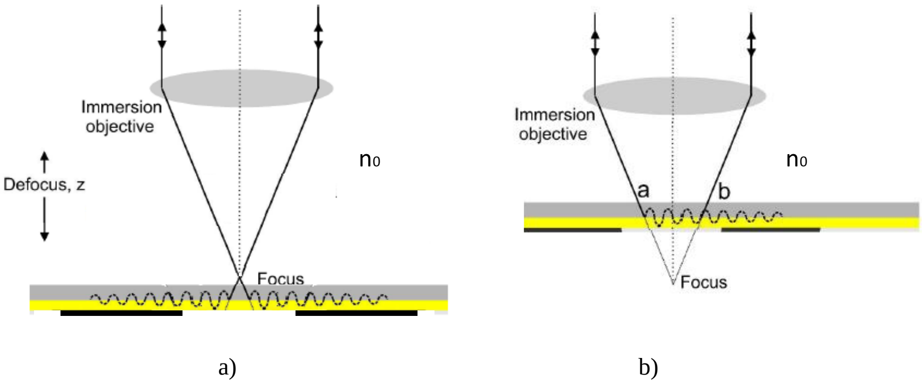


Figure 2.16 shows schematic diagram explaining how SP can be excited and detected a) when the sample is in focus and b) when the sample is defocused

Figure 2.16a shows SP being excited in directions on a sample surface due to incident waves at the surface plasmon angle. Figure 2.16b shows the situation where the SP appears to be excited at points ‘a’ and ‘b’ and reradiate continuously along the surface. I will show later in the thesis that with interferometric and confocal detectors we can limit the points of excitation and detection to ‘a’ and ‘b’.

Oil immersion objective lens

Oil immersion objective lenses are now commercially available for a wide range of products ranging from $NA=1.1$, 1.25, 1.49 to 1.65. For an air backing experiment, an oil immersion lens with a numerical aperture greater than approximately 1.1 with a fully illuminated objective lens will be sufficient to excite SP on the sample. However aqueous material based experiments, such as biological experiments, require a higher NA objective.

For air backing experiments, a high quality immersion objective with NA of 1.45 such as the Zeiss Plan Fluar 100x is a good choice offering high numerical aperture, also it requires conventional immersion oil $n = 1.518$ which is inexpensive and safe. However for the water backing experiment, it is becoming more troublesome where the plasmonic angle becomes too big even for the NA of the 1.45 objective lens; for example, 72 degrees at 633nm wavelength. Nikon have introduced an objective lens with an NA of 1.49, which uses conventional coverslips and immersion oil. This objective lens might be able to solve our water based experimental problem, especially if one operates at a longer wavelength where θ_p is smaller. The other approach to get around the problem is to use the highest commercially available NA objective with NA of 1.65 such as Olympus 100x, which requires high refractive index immersion fluid and coverslips. The immersion fluid $n=1.78$ is made of di-iodomethane, which is volatile, flammable and there are also some health risk issues such as it can cause permanent eye damage and is respiratory irritating (Aldric 2011). The coverslips are made of sapphire, which is, of course, very expensive, although recently there is a cheaper alternative coverslip made of SF11 glass. These are limiting factors that make long term quantitative experiments inconvenient.

Solid immersion objective lens

High NA objective lenses require a high refractive immersion media; solid immersion materials can give higher refractive index values than immersion liquids. The uses of solid immersion lenses (SILs) in SP imaging systems have been demonstrated and published by many authors (Mansfield & Kino 1990; Terris *et al* 1994) and it seems to be a very promising way to get a high NA objective. The SIL system is sometimes referred to the word ‘super-SILs’ as the glass ball can firstly increase the *k-vector* of the illuminations and reduce the wavelength by the high refractive index medium. Very high NA objectives greater than 2 are achievable using the SILs (Ichimura *et al* 1997; Yoshita *et al* 1998).

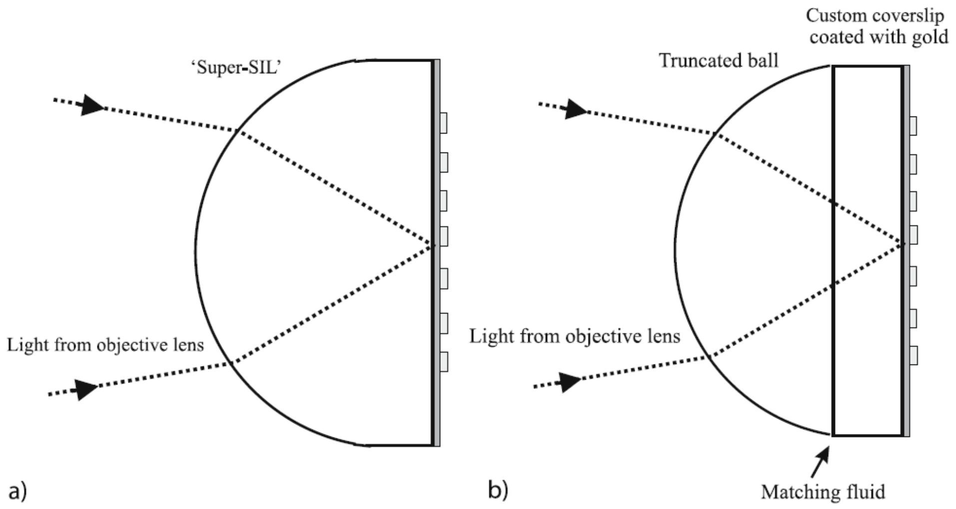


Figure 2.17 shows structure of a) Super-SILs and b) Truncated ball (Somekh 2007)

The technique has one problem which is that the single ball is not practical with scanning systems where the requirement is to move the sample relative to the optical and sample axes. A way to get around this problem is to split the ball into two pieces (Zhang *et al* 2007) as shown in figure 2.17b.

Sandwiched gold structure

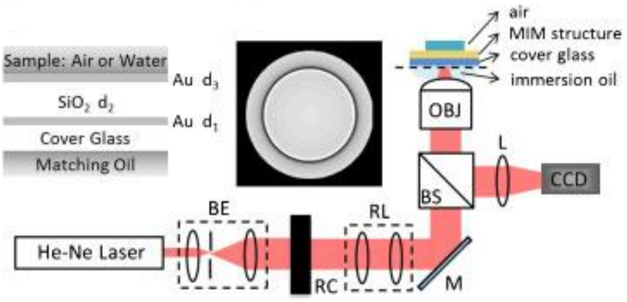
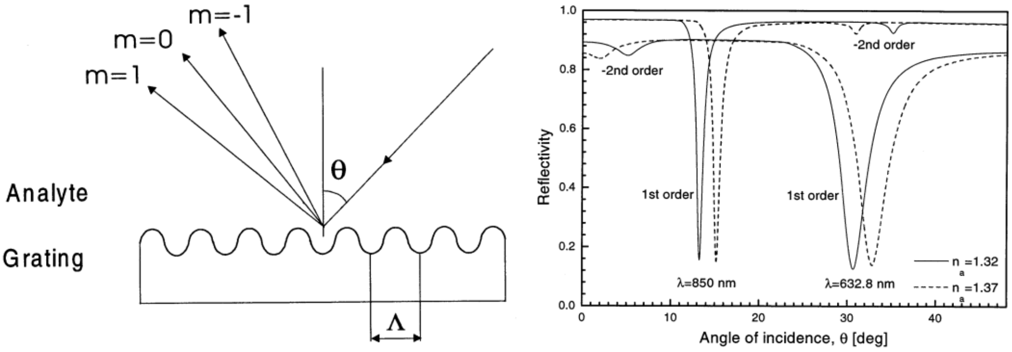


Figure 2.18 Optical configuration of objective-based SPR sensor, where BE: beam expander, RC: radial converter, RL: relay lens, M: mirror, BS: beam splitter, OBJ: objective lens, and CCD: charge-coupled device. Sandwiched gold structure (Lan *et al* 2011).

Recently, Lan *et al* (Lan *et al* 2011) demonstrated the use of double layers of sandwiched gold to move the plasmonic angle from around 60 degrees down to 30 degrees by coupling SP of the first gold layer onto the other gold layer. Although the width of the plasmonic dip widens, this is still a promising candidate for a sensor type experiment. However, the literature did not provide phase information of the system, and this information is needed to predict the propagation length of SP waves and also the lateral resolution if such the structure is used for imaging system.

Grating excitation



a) b)

Figure 2.19 shows a) grating structure and diffracted order definition
b) Reflection coefficient of 1st and 2nd orders (Homola *et al* 1999a)

It is worth pointing out that a grating can also provide suitable conditions for SP excitation even though this is not in the scope of this thesis. This is done by adding the incident *k-vector* up with integer numbers of grating vector k_g as shown in figure 2.19.

$$k_x = k_{vector} n_0 \sin(\theta_i) + m k_g \qquad [2.11]$$

where m is the m^{th} number of diffracted order.

It has been demonstrated that SP waves can be excited with normal incident by employing a grating (Porto *et al* 1999). This is very good for SP excitation, the reasons that it has not been employed in this study are firstly that the grating structure can generate a higher order mode of SP (see chapter 4 for more details and its potential is discussed in the future work section) and with a small period grating there will be more than one diffracted order that can excite SP as shown in figure 2.19b; another problem is when exciting with a lens there is a large range of azimuthal angles and the instrument will behave differently for each angle, whereas the aim of this thesis is to develop a theoretical background to understand the imaging performance of SPR microscopes.

2.2 SPR sensors

SPR sensors are a very successful story of dramatically growing from a laboratory scale experiment to a commercially available instrument. The most famous SPR instrument company is developed by the company Biacore. Biacore is a leading SPR instrument company, who has set a gold standard for SPR devices (Raether 1988). In fact, many review journals actually followed the step by step improvement of Biacore devices (Schasfoort & Tudos 2008). However, here I am not going to follow this path, but instead I will first introduce the SPR sensor principle; follow by presenting an overview of the differences between amplitude and phase measurement and end the section by a summarised table to review the current most interesting techniques to improve the sensitivity of SPR.

As discussed earlier, SPR is very attractive for biological and chemical experiments because of its properties (Flanagan & Pantell 1984; Green *et al* 1997). Therefore in this section I will briefly summarise the key features and current problems of SPR as a sensor.

Key features of SPR sensor and its capabilities

a) Very *sensitive*, the sensitivity can be employed to measure refractive index changes within a small nano scale layer of the sample

b) It is a *label free* technique. Any changes in the refractive index will alter the propagation constant of the SP without any labelling of the analyte.

c) The detection can be performed in *real time*.

Current problems of SPR sensor

a) *Lack of selectivity*: the issue is that SP is sensitive to refractive change and not specific to any analyte, therefore measures need to be taken to improve selectivity using, for instance antibody capture.

b) *Low throughput*; this means that it is hard to fabricate the SPR device so that it can measure many channels in parallel. A way to get around this problem might be to use an imaging array for parallel SPR detection, such as DNA hybridisation imaging (Brockman *et al* 1999; Nelson *et al* 2001). There is a Biacore device that uses grating to perform array sensors. Nevertheless the sensitivity for array is generally a lot poorer than four point measurements.

c) *Narrow range*: an SPR sensor is normally designed for a specific range of refractive index change such as *n=1.33* to *n=1.35*; the sensor will not be able to detect the changes beyond the designed range.

2.2.1 Principle of SPR sensor

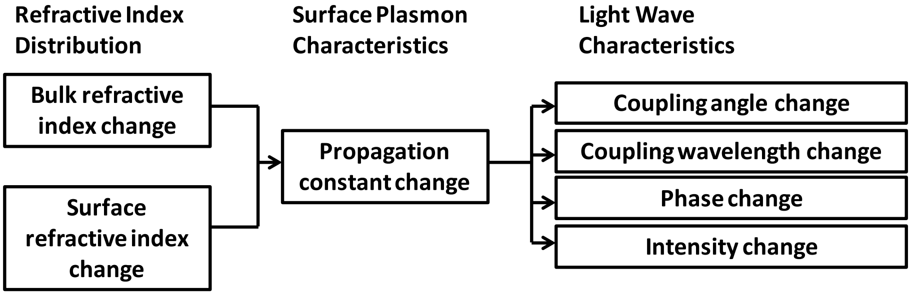


Figure 2.20 shows the concept of how SPR sensors work : modified from (Homola 2008)

In SPR sensors, surface plasmons are excited at the metal interface; changes in the refractive index will produce a change in the propagation constant of the SP, which will change the coupling condition of the SP and result in the change of SPR intensity, angular shift, coupling wavelength and phase as shown in figure 2.20.

2.2.2 Intensity and phase detection

It has been established that a phase measurement can provide an order magnitude better sensitivity than the modulus measurement (Huang *et al* 2012). The only drawback of the phase measurement is that a more complicated optical setup is required as it is necessary to build up an interferometer to measure the phase change.

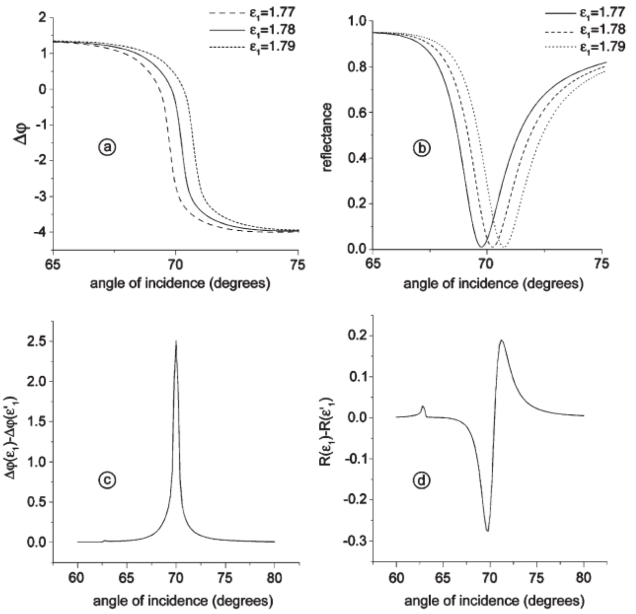


Figure 2.21 shows differences in intensity and phase detection (Raether 1988)

a) Phase responses as a function of incident angle to detect different materials
b) Intensity responses as a function of incident angle to detect different materials
c) Difference in phase detection
d) Difference in intensity detection

The reasons why the phase detection is more sensitive than the amplitude detection have been explained by (Kabashin *et al* 2009b), where they took detection limitations, such as noise and intensity of the detected signal into account. They explained that the phase detection is more sensitive than the amplitude detection because (i) The maximum phase change occurs at the strongest SP excitation position, whereas the maximum intensity change occurs at the angles smaller or bigger than the strongest SP excitation position as shown in figure 2.21c and 2.21d; (ii) under a proper design of a phase detection setup, the phase noises can be order of magnitude lower than the amplitude detection noises. In other words, the phase detection has a better SPR performance compared to the amplitude one; and (iii) the phase detection offers much better possibilities for signal averaging and filtering. In order to confirm these, I have provided a set of simulations to quantify and determine the smallest change in thickness that can be measured in the presence of noise. The minimum thicknesses that can be measured were 0.265 nm and 0.025 nm for amplitude detection and phase detection respectively. The system parameters used in the calculation can be found in appendix D.

2.2.3 Summary of the SPR sensors performances and techniques

I will summarise some of current interesting SPR sensor techniques and their performances in the table 2.1.

Metal structure	**Metal details**	**Optical configuration**	**range of refractive index**	**wavelength**	**Sensitivity**	**Ref**
Uniform metallic film	Au, Ag metal film	Kretschmann	1.33-1.34	400-800	100-300 deg/RIU	(Roh *et al* 2010)
Uniform metallic film	Ag film with low index prism	Kretschmann	1.328-1.332	1310	500 deg/RIU	(Huang *et al* 2010)
Uniform metallic film	Au metal film	Kretschmann	1.35	500-1000	7500 nm/RIU 10^{-8}RIU	(Nenninger *et al* 2002)
Mulit layer metallic films	Au and Si, ZrO2 thin film	Kretschmann	1.325-1.335	633	50-230 deg/RIU	(Lahav *et al* 2008)
Mulit layer metallic films	Ag-Au bimetallic layer	Kretschmann	1.33-1.34	633	$7.85\text{x}10^{-6}$RIU	(Lee *et al* 2010)
Nanostructured array	Au nano cylindrical array	Kretschmann	1.33-1.335	633	10^{-7}RIU	(Chen *et al* 2008)
Nanostructured array	Au nanorod metamaterial	Kretschmann	1.33	1200-1300	30000 nm/RIU	(Kabashin et al 2009a)
Nanostructured array	Metallic-dielectric mixed grating	Kretschmann	1.33-1.36	633	120 deg/RIU	

Metal structure	**Metal details**	**Optical configuration**	**range of refractive index**	**wavelength**	**Sensitivity**	**Ref**
Multichannel sensor	Dual channel, dielectric over-layer	Kretschmann	1.33-1.34	550-1150	5 ug/mL alpha-DNA	(Homola *et al* 1999b)
Multichannel sensor	Angled polishing prism	Kretschmann	1.444-1.450	500-900	2710nm/RIU 8500 nm/RIU	(Dostalek *et al* 2005)
Symmetrical cladding off	Au, Ag metal film	Fiber-optic	1.33-1.34	400-650	2000-4500nm/RIU	(Roh *et al* 2010)
Nano-structured sesor	Au metallic grating	Fiber-optic	1.33-1.34	900-1600	4000-9800 nm/RIU	(Roh *et al* 2010)
Grating-coupling	Au surface grating	Grating coupling	1.33-1.34	600	440 nm/RIU	(Roh *et al* 2010)
Grating-coupling	Al-Au bimetallic layer	Grating coupling	1.32-1.37	900	187.2 deg/RIU	(Hu 2010)
Metamaterail-like	Au nano-structed layer	Grating coupling	1.332-1.372	2000	588nm/RIU	(Liu *et al* 2010)
Nanostructured array	Single or double-square periordic array	Grating coupling	1.333-1.420	400-950	200-350 nm/RIU	(Nishiuma *et al* 2008)
Nanostructured array	Nanoparticle pair, disk pair	Grating coupling	1-1.5	500-900	172434 nm/RIU	(Jain *et al* 2008)
Nanostructured array	Unperiordic array	Grating coupling	-	300-700	165 nm/RIU	(Galush *et al* 2009)
Nanostructured array	Gold nano-ring array	Grating coupling	1-1.3	300-1800	637.3nm/RIU	(Jiang & Sabarinathan 2010)

Metal structure	**Metal details**	**Optical configuration**	**range of refractive index**	**wavelength**	**Sensitivity**	**Ref**
Nanohole	Square nanohole array	Grating coupling	1.33-1.34	600-1000	300nm/RIU	(Roh *et al* 2010)
Nanohole	Nanohole	Grating coupling	-	1520-1570	1110nm/RIU 1570nm/RIU	(Pang *et al* 2007)
Nanohole	Fluoropolymer substrates	Grating coupling	1.33-1.37	600	323 nm/RIU	(Kikuta *et al* 1995)
Uniform metallic film	Au, Ag metal film	Interferometer Mach-Zehnder	1.33	1550	250nm/RIU	(Nemova *et al* 2008)
Uniform metallic film	Au, Ag metal film	Interferometer Two slits	1.32-1.325	877.3	4547nm/RIU	(Wu *et al* 2009)
Ring resonator	Disk resonator		1.33-1.34	1460	600nm/RIU	(Kim *et al* 2007)
Ring resonator	Triangular resonator		-	1555		(Kim *et al* 2007)

Table 2.1 Performances of SPR based sensors (updated at the end of 2011)

2.3 SPR microscopy

Although SPR sensor is a well established field of research, to utilise these SP waves and integrate them with a microscope is a relative young field.

There have been many works reported for successful SP microscopy in near field (Bozhevolnyi 1996; Hecht *et al* 1996; Konopsky 2000). The near field SP microscopy can, of course, provide us with very high lateral resolution. However this technique is not very compatible with biological experiments as they require inconvenient probing geometries.

Before starting to mention about far field SPR microscopy, it is worth pointing out that there is a hybrid method between far field and near field configurations reported by Smolyaninov (Smolyaninov *et al* 2005a; Smolyaninov *et al* 2005b). This technique offers potentially outstanding lateral resolution; however it still requires an unusual optical configuration and special conditions.

The above techniques require inconvenient configurations and here we wish to incorporate, as far as possible, the SP microscope in the framework of a conventional microscope system.

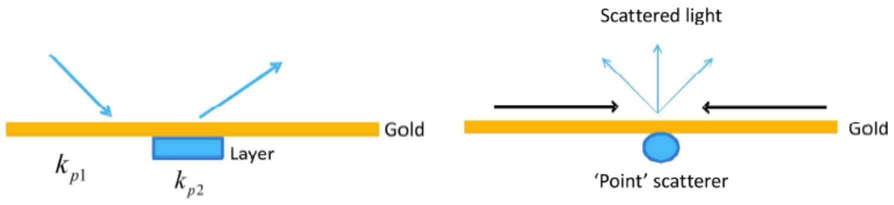
Having discussed that the physical properties of surface plasmons are so different from the unconfined light, therefore the image formation utilising the SP will, of course, be significantly different from the other microscopy techniques. This means that the considerations for optimising surface wave and surface plasmon microscopy differ from conventional microscopes.

For SPR microscopy, there are also two optical configurations that we might use; namely, ‘near field’ and ‘far field’.

2.3.1 Contrast mechanisms in SPR microscopy

The contrast mechanisms in SPR microscopy can be divided in two types as shown in figure 2.22.

1. A change in the local SP propagation constant due to a change in the local resonant conditions.
2. Scattering of a surface wave into a propagating wave which is collected by a microscope objective.



a) b)

Figure 2.22 shows two contrast mechanisms for SPR microscopy (Pechprasarn & Somekh 2012)

a) Mechanism 1 scattering back due to the change in the local refractive index region
b) Mechanism 2 scattering back of the excited SP, which is collected by an objective lens

Contrast mechanism 1

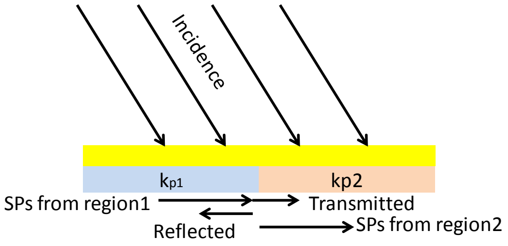


Figure 2.23 shows schematic diagram showing reflection and transmission of a source from a structured sample with different surface wave vectors

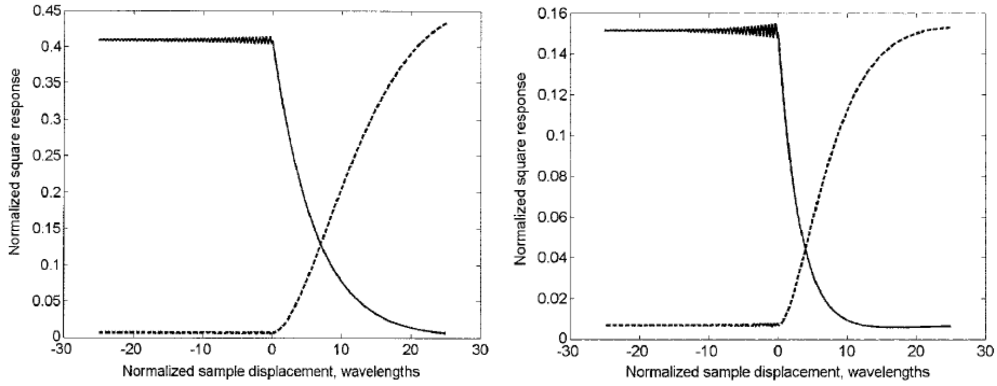
In this situation we wish to measure the value of k_p in a localised region. In order to illustrate this mechanism, let us consider the case where a single uniform plane wave with wave vector k_x. The plane wave excites SP and the SP then propagates from left to right as shown in figure 2.23. The output collected by the microscope objective is contributions from three reradiated fields namely.

1. a field generated in region 1 and reradiated in region 1
2. a field generated in region 1 and reradiated in region 2

3. a field generated in region 2 and reradiated in region 2

Zhang *et al* (Zhang *et al* 2006) have demonstrated that by calculating the field contribution for each region will allow us to calculate the field arising for a plane wave incident on an inhomogeneous sample.

The results published in (Zhang *et al* 2006) where the plane wave was slightly off resonance has provided us with a theoretical framework and a good illustration for a trade-off between resolution and sensitivity. The trade-off has also been reported by many research groups who commenced their works on SPR prism based microscopes, such as (Berger *et al* 1994; Yeatman & Ash 1987).



a) b)

Figure 2.24 show surface plasmons wave responses across an interface on and off resonance for strong and weak coupling for a single incident angle k_x calculated by using Green's function. One unit of wavelength is the surface wave wavelength in the first medium (Somekh 2007).
a) Weak coupling of SP in the left hand side region propagating from left to right as shown in figure 2.23
b) Strong coupling of SP in the left hand side region propagating from left to right.

From figure 2.24, we can summarise the key trade-off in SPR microscopy, which is the conflict between lateral resolution and sensitivity. The figure 2.24a shows that when the SP were weakly excited and propagated, we can see that the lateral resolution is poorer than the other case where the SPR is strongly excited shown in figure 2.24b. However, the sensitivity, which can be measured by difference in amplitude of the two on and off resonance regions, is lower for the strongly excited case than the other case. In other words, from a microscopy point of view we can get a very good lateral resolution when the contrast is rather low. We can see that the value of SP response can be accurately measured once the signal value reaches a constant level. The length of the transition region gives an indication of the localization that can be achieved with particular SP properties.

A number of papers (Berger *et al* 1994; Rothenhausler & Knoll 1988; Yeatman & Ash 1987; Yeatman 1996) have pointed out that the lateral resolution in the direction of SP propagation is poorer than the direction normal to the propagation direction.

Contrast mechanism 2

This mechanism applies for imaging a point object but does not lend itself so well to quantitative measurement of k_p. This scattering mechanism is based on point object type sample, which perturbs the SP propagation, then reradiates the SP back into the free space. The light reradiated from the SP is eventually collected by a microscope objective. The magnitude of the reradiated back SP will, of course, be quite weak compared to the first mechanism and the lateral resolution is now dependent on the optical detection and wavelength of the SP. The contrast of the image is now dependent on the intensity of the back scattered wave rather than only the strength of SP coupling. In other words, the trade-off between resolution and sensitivity is not a dominant factor in this case and resolution should be diffraction limited.

In addition, for the mechanism where contrast is dependent on field enhancements, other types of surface waves would be preferred to serve such a mechanism so ATR can produce images with better quality.

2.3.2 Far field SPR microscopy

The first attempt to develop SPR microscope in far field was reported by Yeatman in 1987 (Yeatman & Ash 1987), where a prism based Kretschmann’s configuration was used. However the prism based system has encountered many problems in term of microscope performance because the prism introduces considerable aberrations. Fortunately, high refractive index materials and high NA objective lenses are becoming more widely available. These enable us to construct SPR microscopes with a more convenient optical configuration and develop more advanced SPR microscopy techniques. In this section, I will review the Kretschmann based SPR microscopes and follow by the objective lens based microscopes.

Kretchmann based SPR microscopes

In 1987, Yeatman and Ash (Yeatman & Ash 1987) reported the idea on combining the SP with a microscope to get a good lateral resolution and sensitivity microscope. One year after that, Rothenhausler and Knoll (Rothenhausler & Knoll 1988) also reported the prism based SPR microscope configuration.

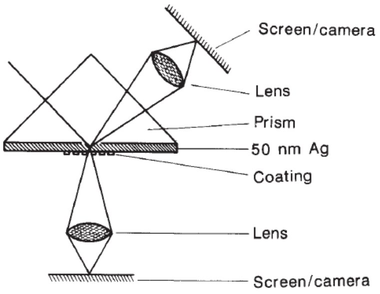


Figure 2.25 shows Rothenhausler and Knoll's SPR microscope setup (Rothenhausler & Knoll 1988).

Both works were done in 'bright field' configurations, where the light emitted from the sample was imaged by a camera. Local variations in the SP propagation constant altered the intensity of the reflected light. This enabled them to form an SPR image with intensity contrast mechanism. Rothenhausler and Knoll also suggested the possibility of capturing a 'dark field' image by taking the SP image at the far side of the prism as shown in figure 2.25.

In order to improve the lateral resolution, it requires a tightly focused beam with a high numerical aperture. In contrast, the SP can only be excited with a narrow range of angles therefore the useful numerical aperture is limited to the SP range. If one excites the SP using the full NA, the image obtained on the camera will be saturated, in other words, there will be virtually no contrast as the contribution due to SP is small compared to the other reflected waves.

In 1996, Yeatman (Yeatman 1996) showed that the lateral resolution can be improved at the expense of reduction in sensitivity, where he reported the propagation length effect on the contrast of the image; the shorter propagation gives a better lateral resolution image with a poor contrast, on the other hand the longer propagation length gives a poorer lateral resolution with a good contrast. For the ‘bright field’ wide field system, the lateral resolution is also limited by propagation length of SP as discussed.

Berger *et al* (Berger *et al* 1994) used prism based SPR microscope to experimentally demonstrate that propagation length of the SP determined the lateral resolution by varying the incident angles and capturing the images corresponding to each mode of SP coupling. They were also able to show that the lateral resolution can be improved by shortening the propagation length of SP. They did this by using a shorter incident wavelength as the shorter wavelength had a shorter SP propagation length. The best resolution reported was approximately 2 microns at 531 nm wavelength but the sensitivity was greater reduced.

As mentioned briefly in section 2.1.3, Giebel (Giebel 1999) suggested another method to improve the lateral resolution of the SPR microscope. He replaced the normal gold layer with an aluminium layer, which had a relatively large imaginary part of the complex relative permittivity. This will, of course, introduce more ohmic loss and shorten the propagation length of SP. As discussed earlier, the lateral resolution can be improved by using a shorter propagation length of SP at the expense of losing contrast in the image. Therefore by using the aluminium layer, he managed to get higher lateral resolution images to monitor cell behaviour but with a rather poor sensitivity typically by a factor of 3 compared to the image obtained by 50 nm of gold.

For the ‘dark field’ image where the image is obtained at the far side of prism by using a high NA objective, it is possible to obtain the lateral resolution better than the 'bright field' as at this side of detection the full NA can be used. In other words, the bright field is limited by the working distance between the prism and the lens, but this is not the case for the dark field where the high NA lens can be employed. Essentially dark field operates by mechanism 1 as discussed in section 2.3.1.Although this approach can improve the lateral resolution, it requires us to have access to the sample as all the detection objects are placed above the sample.

In 2000, Grigorenko *et al* (Grigorenko *et al* 2000) suggested a way to solve this problem by blocking the zero order reflection. This enabled them to obtain a dark field image at the reflection side on reflection mode.

There has been an interest in using phase detection in a prism based Kretschmann SPR microscope. Grigorenko *et al* (Grigorenko *et al* 1999) suggested using an appropriate thickness of metal (they used 47.7 nm of gold); so that the minimum of the reflection becomes zero. In this case, the gradient of the phase variation as a function of incident angle can increase without bound around θ_p. This idea has been experimentally confirmed by Notcovich *et al (*Notcovich *et al 2000a;* Notcovich *et al 2000b)*, who imaged flowing gas with a refractive index change of 10^{-6} RIU.

Although the Kretschmann configuration gives us an appropriate way of building a sensor, this configuration is rather problematic for SPR microscope configuration. The difficulty of the configuration can be summarised as

1. The prism is not compatible with a conventional microscope configuration as it requires imaging through a prism.
2. The lateral resolution obtained from the configuration is rather poor compared to the conventional optical microscope.
3. The configuration suffered from severe aberrations as it requires an oblique angle illumination and imaging.

Having discussed the properties of SPs, it is important to note that these properties are very important and attractive to biological research. Therefore it is necessary to develop techniques that can overcome the limitations and utilise the properties of SPR.

Objective lens excitation SPR microscopes

In section 2.1.4, I have discussed the development of high refractive index material and high NA objective lens. In this section, I will review the objective lens based microscopes. The objective lens based SP microscopes not only overcome the limitations in Kretschmann configuration, but also enable us to develop more advanced techniques to overcome or at least optimise the trade-off between resolution and sensitivity.

Objective lens based SPR microscopes can be constructed in three ways, namely non-interferometric SPR microscope, interferometric SPR microscope and confocal SPR microscope. The interferometric SPR microscope and the confocal SPR microscope presented in this thesis are employed to detect relative phase change of the SPs.

Non-interferometric SPR microscopes

The non-interferometric microscopes can be constructed in two ways; a scanning microscope and a wide field microscope. The drawbacks of the scanning microscope are the time required to scan over the sample and also that it is hard to access the sample from the far side of the optical configuration. There are also some interesting theoretical concepts on the resolution and sensitivity reported.

Therefore in this section, I will provide some discussion of the theoretical ideas arising from the non-interferometric microscopes.

Wide field non-interferometric SPR microscopes

The system shown in figure 2.26 requires a rotating diffuser to ensure that the laser light becomes spatially incoherent as the coherent light can introduce speckle noise. The reradiation can be seen on the back focal plane and in the SPR image (Tan 2011).

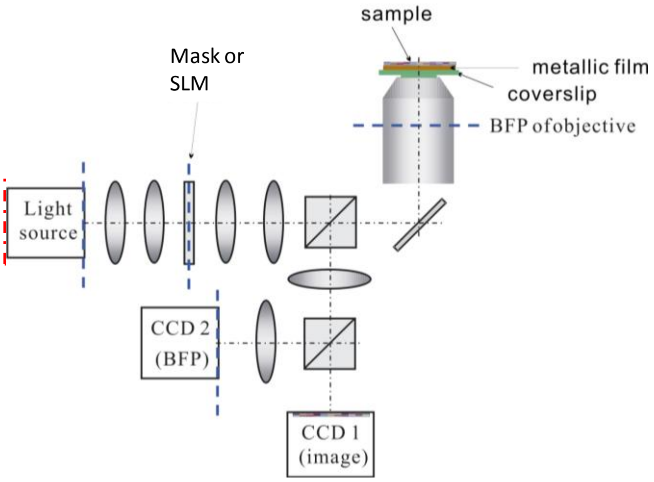


Figure 2.26 shows a wide field non-interferometric SPR microscope setup (Somekh & Pechprasarn 2010)

The system employed two CCDs to capture the image plane and also the back focal plane. Stabler *et al* (Stabler et al 2004) set up this experiment and obtained the experimental results shown in figure 2.27, where they used 45 nm thick gold layer and a Zeiss Plan Fluar 100x objective with an NA of 1.45.

From figure 2.27a, we can see the characteristic plasmonic dip corresponding to the theoretical calculation in figure 2.27b, this is an evidence of the existence of SPs, which were excited through the objective. These results shown in figure 2.27 are different from the results obtained by Tanaka and Yamamoto (Tanaka & Yamamoto 2003). They used prism Kretschmann based configuration to excite SPs with four detectors to detect different azimuthal angles; the other difference is they scan over the individual point on the back focal plane of objective by using a single diffraction limited spot rather than the full aperture at the same time. The back focal plane image can be thought of multiple detectors but it is much faster and captures all the full azimuthal angles profile with only one physical CCD detector.

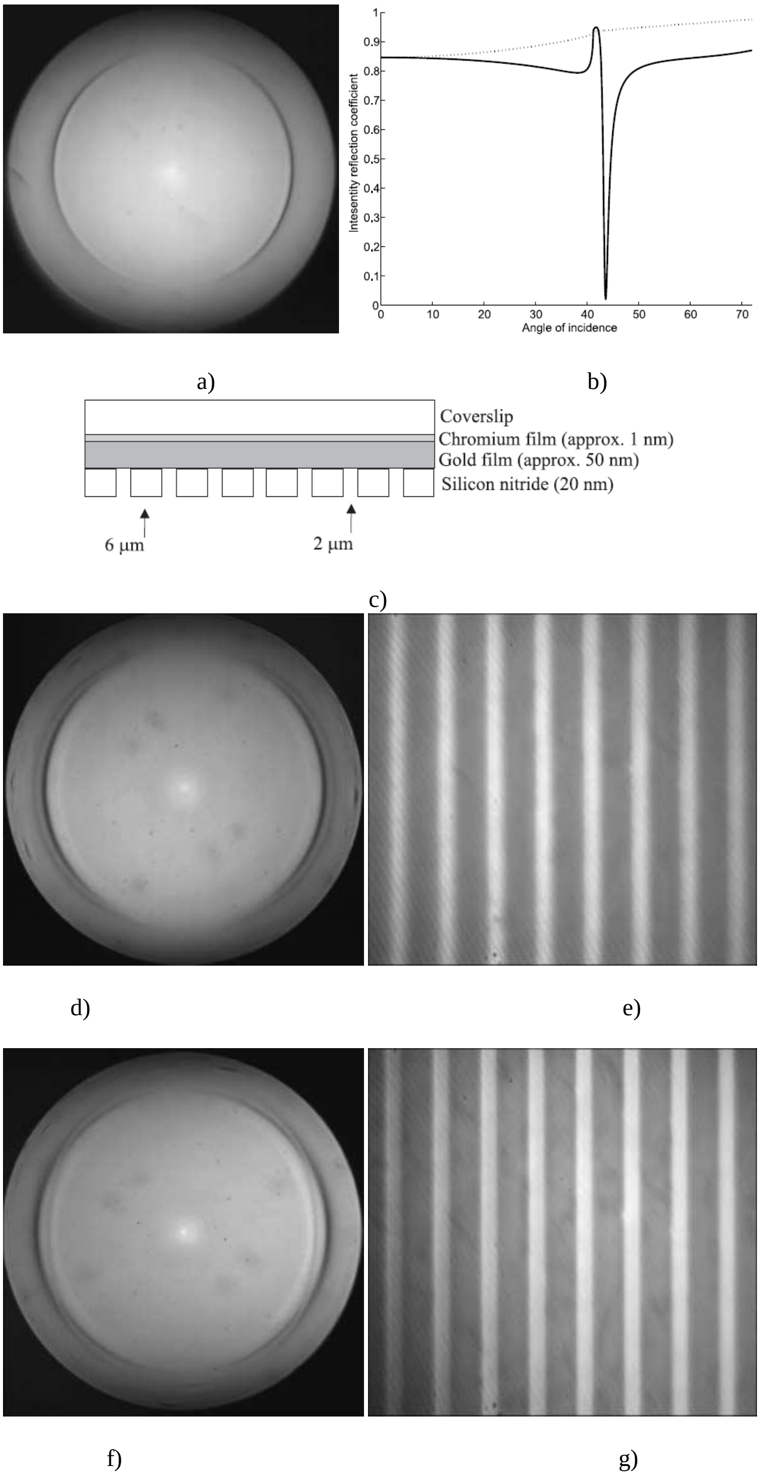


Figure 2.27 shows results published in Stabler's paper (Stabler et al 2004)
a) Back focal plane distribution on a uniform gold sample. The horizontal direction corresponds to *p*-incidence and the vertical to *s*-incidence.
b) Intensity reflection coefficients for *p*-(solid) and *s*-(dashed) *polarisations*.
c) Schematic diagram of the sample.
d) Back focal plane distribution obtained for input polarisation direction in the back focal plane parallel to the grating vector.
e) Image of grating structure captured when input polarisation direction in the back focal plane parallel to the grating vector.
f) Back focal plane distribution obtained for input polarisation direction in the back focal plane normal to the grating vector.
g) Image of grating structure captured when input polarisation direction in the back focal plane normal to the grating vector.

The reason that the mask or the amplitude spatial light modulator (A-SLM) is required is that if the image is obtained without them; the image will have virtually no contrast. This is due to the strong reflected background intensity level compared to the small amount of reradiated SPs. The annular mask and the A-SLM are used to block the light that does not contribute to the image contrast. The results in figure 2.27e and g clearly show that the lateral resolution is dependent on the propagation direction. The lateral resolution reported using this setup was approximately 1.3 μm and 0.93 μm where the electric field was parallel (*x polarisation*) and normal (*y polarisation*) to the grating vector respectively.

It has been shown that the contrast in wide field SP microscope can be controlled by varying the incident angles using A-SLM and also reversal contrast image can be obtained with different incident angle (Tan 2011).

Scanning non-interferometric SPR microscopes

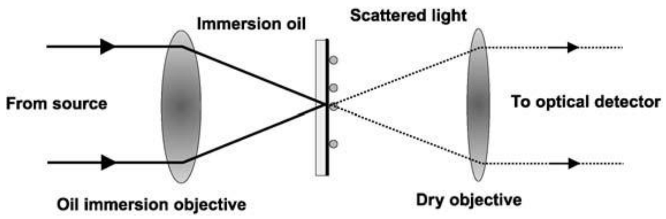


Figure 2.28 shows scanning non-interferometric SPR microscope used by Kano (Somekh 2007)

In 1998, Kano *et al* (Kano & Knoll 1998; Kano *et al* 1998) introduced an objective lens based scanning non-interferometric SPR microscope. This idea is actually similar to the dark field prism based configuration where the SPs image is captured at the sample side. The reradiated SPs on the sample side are detected by a dry objective lens. This experiment has clearly demonstrated the capability of using an objective lens to excite SPs and showed that the lateral resolution obtained is in fact comparable to the diffracted limit spot size. This arrangement is also limited by the arrangement so that the sample is not accessible in practical biological measurements.

Measuring the back focal plane distribution is similar to the idea of back focal plane ellipsometry, where the sample reflectance is measured as a function of incident angle and input polarisation. This technique measures the reflectivity by scanning a focal point over the back focal plane or a line of corresponding azimuthal angle on the back focal plane. This allows the SPs wave to be measured in a highly localised region. This technique has been widely used and reported by several authors to measure film thickness in semiconductors (Fanton *et al* 1993; See *et al* 1996). See *et al* (See *et al* 1996) suggested a technique to compensate the material variations in the sample, such as microphonic variation. In 1995, Shatalin *et al* (Shatalin *et al* 1995) has extended Fanton's idea to extract ellipsometric properties over a confined local region of a sample.

Kano and Knoll (Kano & Knoll 1998) extended this ellipsometry concept to measure the local thickness of Langmuir–Blodgett films. They managed to obtain a reasonable result for 4 Langmuir–Blodgett mono layer films by measuring the plasmonic dip on the back focal plane. In 2000, they also extended this technique to make a scanning microscope by scanning a ring excitation over the sample and obtaining the image (Kano & Knoll 2000). The best lateral resolution obtained with the technique was approximately about 1.5 microns. The difficulty that limited them from achieving a better lateral resolution was noise and interference in the back focal plane.

Interferometric SPR microscopes

Interferometric SPR microscopy is a more advanced technique and there is not much information reported on this technique. In this thesis, one of the main objectives is to explore and confirm that the interferometric microscope can provide us with much better lateral resolution and new defocus operating points.

In this section, I will show that an interferometer not only enables us to measure the phase of the signal, but also allows us to obtain far better spatial resolution compared to non-interferometric methods.

Scanning interferometric SPR microscopes

In order to simplify this SPR microscopy technique, let us consider a scanning heterodyne interferometric system in figure 2.29. The interference signal is produced by two beams interfering each other, namely the reference beam and signal beam. The signal beam is excited by the objective lens and reflected back from the sample, whereas the reference beam is provided by the reflection from the mirror.

This system is quite similar to the confocal microscope in term of their characteristics. They both actually have the same transfer functions (Holmes & Somekh 1994; Somekh 1992) with only one difference, which is the amplitude of the outputs obtained from the two systems. The output of the interferometric system will be much stronger than the confocal system. This is due to the fact that the output of the interferometric is directly proportional to the amplitude of the incoming field, but the confocal output is a square of the amplitude.

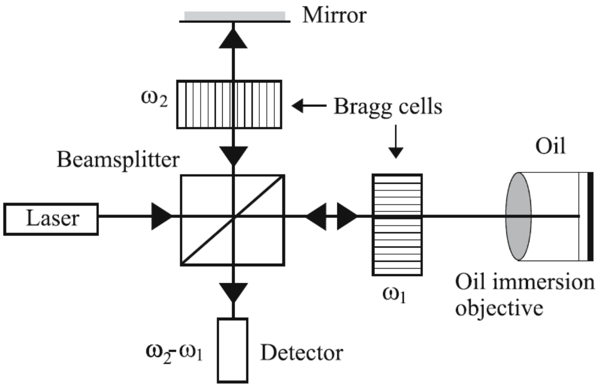


Figure 2.29 shows a scanning interferometric SPR microscope setup (Somekh 2007)

When the sample is in focus $z = 0$ the reference and the reflected beams have the identical flat wave front; the two beam can then interfere and give the higher magnitude of the interference, whereas the phase of the signal beam will become curved when the sample is defocused and there will be a reduction in the interference signal as schematically shown in figure 2.30. This is similar to a confocal pin hole where only the in phase signal is detected at the detector. The reference beam can be seen as a virtual pinhole in confocal system.

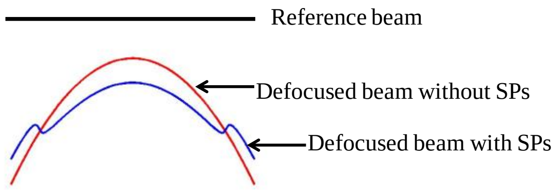


Figure 2.30 shows the phase of the reference beam and the signal beam

The heterodyne interferometric microscope is illuminated with 633 nm wavelength and the beam is then split into two parts. The first beam will be directly reflected by a mirror; this provides the reference beam. The second beam will be focused through an oil immersion objective; the reflection from this beam is the signal beam. There are two Bragg cells operating at two different frequencies (the difference between the two frequencies is typically about 10 kHz). The Bragg cells frequency difference provides us with a convenient operating frequency to detect the inherent signal at the beat frequency and also reduces the multiple reflection effect in the Bragg cells (Migdall *et al* 1990). The sample sits on a 3D piezo-scanning stage which is computer controlled. The interference signal is then detected using a lock in amplifier.

Now let us consider the case where there are SPs waves propagating along the metal interface. The SPs will alter the phase of the signal beam by 2π as shown in figure 2.30. This allows the signal beam and the reference beam to interfere. This effect has been explained for the case of the scanning acoustic microscope by Atalar (Atalar 1979).

The contrast mechanism of the interferometric SP microscope can be explained by *V(z)* equation. The *V(z)* is a measurement of the output response from the interferometer *V* as a function of defocus position *z*. The *V(z)* effect has been widely used to evaluate pupil functions of objective lenses and to determine the aberrations in an optical system (Zhou & Sheppard 1997). The effect is also an important tool in acoustic wave microscopy, especially when lenses with large aperture angles are used (Atalar 1978; Parmon & Bertoni 1979).

The *V(z)* curves provide us with a method to accurately measure wave velocities and a powerful contrast mechanism for imaging in acoustic microscope (Ilett *et al* 1984; Somekh *et al* 1985). Therefore for the SP interferometric microscope, the *V(z)* is now applied with SPs, which is similar to the surface acoustic waves. One difference between acoustic waves and light is, of course, that the light properties are dependent on its polarisation.

In order to calculate the *V(z)* for SP interferometric microscope, it is required to know the reflected field output from the sample on the back focal plane of the objective lens. Let us consider the case where linearly polarised light is illuminating the back focal plane of the microscope objective. Let the field amplitude reflected from the sample and turned by the objective lens on the back focal plane radially as a function of *P(r)*, where *P(r)* is the objective lens pupil function. This pupil function *P(r)* may account for amplitude-phase variations and aberrations. We can then express the reflected field as (Somekh 2007):

$$E_{reflected} = P^2(sin\theta)[R_p(sin\theta)cos^2\emptyset + R_s(sin\theta)sin^2\emptyset]e^{jn_0kcos\theta z} \qquad [2.12]$$

The pupil function here is squared because the beam passes through the objective twice. The terms R_p and R_s are the Fresnel reflection coefficients for *p-polarisation* and *s-polarisation* respectively. θ and $\emptyset$ are incident angle and azimuthal angle respectively. The exponential term accounts for the phase change when defocus *z* is changed. *k* is the wave number in free space ($2\pi/\lambda_{free}$), where λ_{free} is the free space wavelength. Negative *z* refers to movement of the sample towards the objective.

The interference signal results from the signal beam and the reference beam E_0. It is required to resolve the $E_{reflected}$ term again with respect to the reference beam polarisation direction. Therefore the interference signal detected by this heterodyne system is given by:

$$V(z) = 2Re(E_{reflected}E_o^*) \qquad [2.13]$$

The *V(z)* term can also be rewritten as the summation of amplitude and phase between the reflected signal and uniform reference signal over the full aperture of the objective, which can be expressed as:

$$V(z) = |\iint P^2(sin\theta)\left[R_p(sin\theta)cos^2\emptyset + R_s(sin\theta)sin^2\emptyset\right]e^{jn_0kcos\theta z}\, dsin\theta d\emptyset| \quad [2.14]$$

The $V(z)$ result from the equation [2.14] is the modulus of the interference signal.

As a part of this thesis, I will introduce a few robust ways to extract the amplitude and phase of the signal beam from the interference signal and show that they enable us to have another operating mode for sensor type and microscopy type applications.

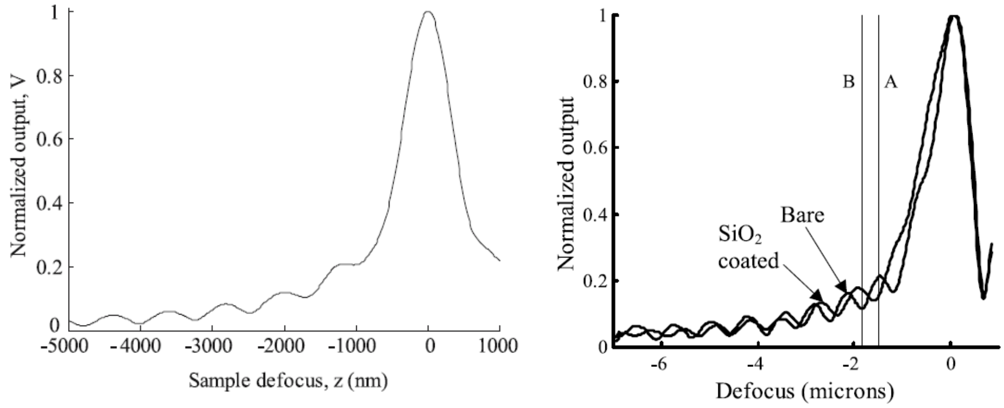


Figure 2.31 shows a) Theoretical *V(z)* curves obtained on bare gold 50 nm thick
b) Experimental *V*(*z*) curves obtained on 50 nm thick of gold (*bare*) and the gold layer coated with an additional 20 nm layer of SiO2 (*coated*) (Somekh 2007)

The additional layer of dielectric will, of course, change the propagation constant of SPs, so that θ_p changes. This change in the plasmon propagation vector can be seen on the *V(z)* curves by a change in the period of the ripples as shown in figure 2.31b.

The ray paths A and B cannot interfere with each other directly because they are spatially separate. However both ray paths can interfere with a common reference beam (Offside *et al* 1989). This means that the strength of the interference signals will oscillate periodically, going through a complete cycle when the phase difference between the two ray paths changes by 2π. It can be shown that (Somekh 2007):

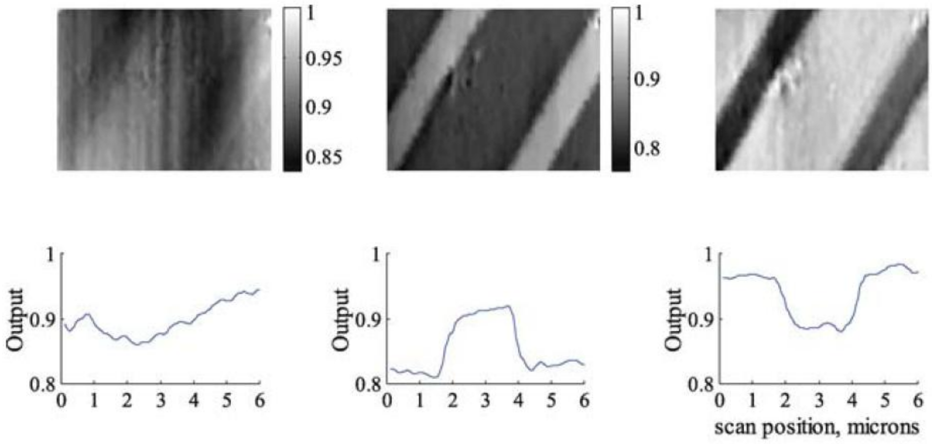
$$\Delta z = \frac{\lambda_{free\,space}}{2n_0(1-cos\theta_p)} \qquad [2.15]$$

Where Δz is ripple period. This is discussed in more detail in chapter 5.

The ripples period measured for the bare gold and the coated cases were 761 nm and 676 respectively. These values were very close to the theoretical values of 758 nm and 672 nm respectively. The presence of the dielectric layer moves the θ_p from 43.6 degrees to 46.4 degrees.

The *V*(*z*) curve can be related to imaging with the heterodyne interferometer. The *V(z)* curve indicates that if we scan the sample with the sample in focus there will be virtually no contrast or very low contrast in the image obtained from the scan. However if the sample is scanned at different z defocus position, the contrast can be improved. Let us assume that the results shown in figure 2.31b were obtained from the two on and off centres of a grating sample; if one images the sample at point A and B (see figure 2.31b), the images obtained will be contrast reversed as shown in figure 2.32. The ability to form images with different contrast and with lateral resolution maintained is a significant advantage of interferometric plasmon microscopy.

The images obtained in focus in figure 2.32a shows very low contrast. On the other hand, when the sample is defocused towards the objective lens as shown in figure 2.32b and c; contrast of the images improved. It is interesting to note that the two images have reversal contrast. The resolution of the images is less than 0.5 microns, which is the best result obtained from the far field SP microscope so far.



a) b) c)

Figure 2.32 shows images experimentally obtained from different defocuses *z:* picture width 10 microns (Somekh et al 2000b).
a) In focus-little contrast.
b) Defocus −1.5 μm good contrast bare region appears bright.
c) Defocus−1.85 μm good contrast coated region appears bright.

The SPs are excited at a particular radius on the back focal plane, so the interferometric SP might be thought of as a 2π (radians) microscope, which is analogous to 4π (stereoradians) 3D confocal microscope (Hell *et al* 1994).

Recently, a research group in France (Argoul *et al* 2010; Berguiga *et al* 2010; Berguiga *et al* 2007; Elezgaray *et al* 2010; Roland *et al* 2010) adopted this technique and has shown that the resolution obtained from this technique is very good. They imaged a nano-sized sample using uniform gold layer. They have also pointed out that the ripples amplitude can be enhanced by using a thinner layer of gold, e.g. 25 nm.

Wide field Interferometric SPR microscopes

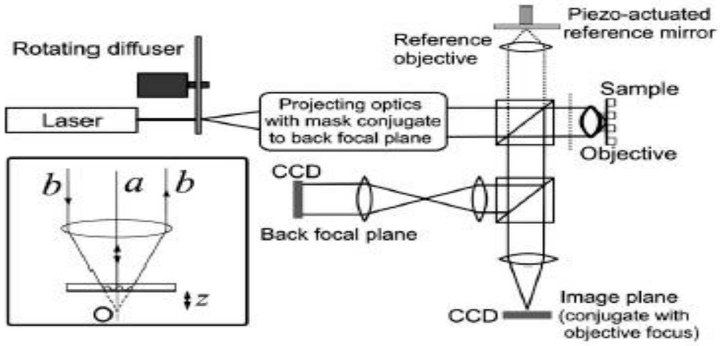


Figure 2.33 shows wide field interferometric SPR microscopes (Somekh et al 2009)

I have shown in the previous section that the interferometer has given us a new technique to overcome the resolution and sensitivity trade-off. The only problem with the scanning interferometer is its scan time. Therefore one potential way to get around this problem is to design a wide field interferometric microscope. It has been shown that confocal transfer functions can be achieved in a wide field configuration using a speckle interferometer (Somekh *et al* 2000c) as shown in figure 2.33. This microscope requires Kö hler illumination reflected from the surface of a diffuser. The speckle pattern is then imaged onto the back focal plane of the sample objective lens and the reference arm objective lens. The reflected distribution pattern is then projected onto a detector so that it can interfere with the reference pattern. The interference signal detected by phase stepping is a measure of the correlation between the reference beam and the reflection from the sample. Rotating the diffuser averages out the speckle noise so a wide field system with an effective confocal transfer function is obtained.

The result shown in figure 2.34 has shown virtually the same concept and pattern as shown in the figure 2.31b; however the two materials have different reflected background levels.

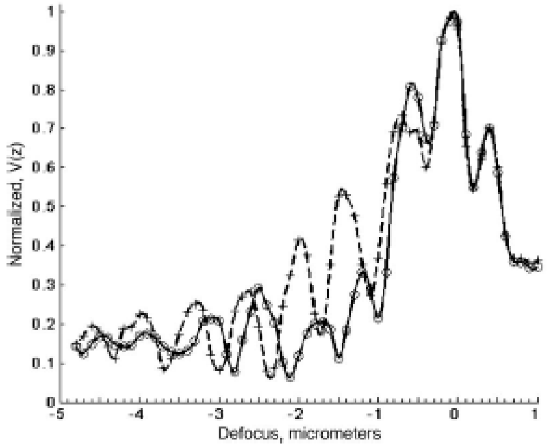


Figure 2.34 shows *V(z)* responses from the wide field interferometric system. Solid line is bare gold, dotted line for coated gold (Somekh *et al* 2009).

a) b) c)

Figure 2.35 shows images experimentally obtained from different defocuses *z*: picture width 10 microns (Somekh *et al* 2009) a) *z=+0.5mm* b) *z=-0.4mm* c) *z=-1.9mm*

Confocal Interferometric SPR microscopes

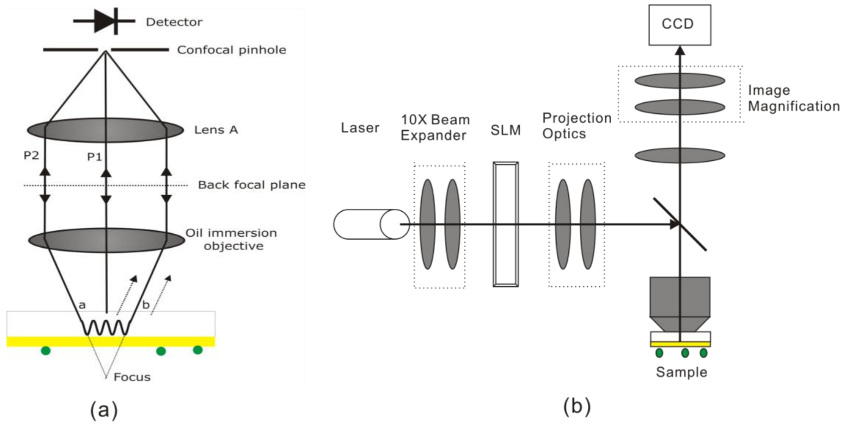
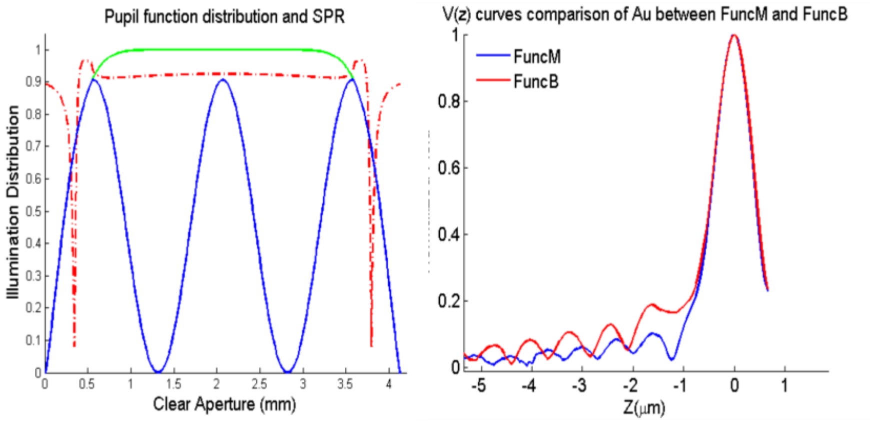


Figure 2.36 shows a) conceptual diagram of SP confocal microscope; b) simplified schematic diagram of the experiment setup (Zhang *et al* 2012)

Having explained that the interferometer actually has the same transfer function as a confocal microscope with the only difference being in the output signal, the output signal of the interferometer is $V(z)$ whereas the output of the confocal microscope is $|V(z)|^2$. The ripples response from the confocal microscope will be, of course, much smaller than the interferometric system. Figure 2.36a shows a common path interferometer and it is true when the system is confocal or non-confocal or even with an interferometric setup. The advantage of the confocal microscope is that the confocal microscope is a common path interferometric system, which requires only one arm in the optical setup. In other words, complexity in optical alignment can be greatly reduced using this configuration however with an expense of the reduction in output signal.



a) b)

Figure 2.37 shows a) pupil function distribution on the bfp. Green is function M denoted in figure b, blue is function B denoted in figure b and red is reflection coefficient when $\phi=0$ degree.
b) $V(z)$ responses of different pupil functions M and B (Zhang *et al* 2012)

As a part of this research, my colleagues and I (Zhang *et al* 2012) have shown that the amplitude output can be enhanced by using an appropriate pupil function. The pupil function can be controlled by an amplitude spatial light modulator. The concept of designing the pupil function is to avoid light that does not contribute to the output signal and to ensure that the light at the SPs angles are sufficiently passed through the objective as shown in figure 2.37. This is discussed in more detail in chapter 6.

Chapter 3

Methodology and validation

In order to explain SPR microscopy performance, we need to have an accurate computer simulation package to quantify the SPs effects for both uniform and structured samples. There are many computer simulation techniques that are capable of calculating the SPR responses.

In this chapter, an overview and background of electromagnetic modelling techniques that can be used and comparison between the pros and cons of each are provided. I will also provide a validation to confirm that such computer simulation packages give us reasonable results for the SPR effects by comparing them with experimental results reported in the literature.

More importantly, I will explain how the simulation techniques can be combined to calculate the SPR microscope responses for both interferometric and non-interferometric modes.

3.1 The need of electromagnetic modelling for SPR microscopy

As discussed in the previous chapter, SPR microscopy is a relatively young field. Therefore there was no software package designed for this application. Although researchers are able to obtain accurate SPR calculations for uniform layers by the Fresnel’s equations, the equations are not sufficient to investigate a more complex problem such as grating sample, metallic grating and other periodic structure. Many attempts have been made to obtain accurate simulation results or at least approximate such a complex structure. A good example for this case is Green’s function approximation, when it was used to approximate the lateral resolution and sensitivity of edge response (Somekh 2007). In other words, we were lacking a simulation tool to give us an exact solution of the Maxwell’s equation, which is fast enough to enable us to calculate the output responses from every incident angles and azimuthal angles on the back focal plane.

Let us firstly summarise the key features of the software package needed for this research:

1. It has to be able to accurately calculate the SPR effects for any particular input polarisation, incident angle and azimuthal angle in order to obtain microscope responses where input is the full aperture of an objective. It has to provide us with the way to determine whether we have reached the correct results for a given boundary condition.
2. The software should offer us the capability to get the accurate solution of Maxwell’s equations rather than an approximation. In other words, it has to be a rigorous computing algorithm.
3. It should be fast enough to run the simulation covering the full back focal plane aperture.
4. It has to be capable of calculating realistic structures. In practice, many methods require the structure to be periodic. We will show that this restriction allows us to examine the structures encounter in most experiments. In the future work section, I will list some research activities that have been initiated by the software implemented in this project.

3.2 Choices of modelling techniques, pros and cons of each

From the requirements listed in the previous section, there are two types of simulation techniques that are capable of doing the job; namely, frequency domain and time domain methods. For each of the methods, there are also many techniques available.

Therefore in this section, I will summarise what is available, compare them and more importantly choose some that are suitable for the job.

3.2.1 Time domain method

Time domain method is an electromagnetic simulation technique which solves solutions of the Maxwell's equations as a function of time. This technique requires the time interval step to be very small to satisfy the differential equation in the Maxwell's equations. Some examples of well known simulation techniques based on the time domain method are finite difference time domain method (FDTD) and transmission line method (TLM).

The advantage of this technique is that it is capable of solving any arbitrary structure and any boundary conditions. This method generally provides a simulation for an impulse response or an excitation of finite duration, which is quite time consuming as it is necessary to calculate the values for each time step. If we need the simulation for a wide range of frequencies, this may be obtained simply by simple Fourier transformation. In general, the method is quite time consuming for monochromatic sources but for polychromatic excitation it can be quite efficient.

Therefore this method is not suitable for the job described above as it takes very long run-time for an input incidence on the back focal plane and also we need to process the data by applying Fourier transform theory to work out amplitude and phase for each particular angles of incidence.

3.2.2 Frequency domain method

Frequency domain method (or Fourier modal method) is a basic electromagnetic modelling technique, where it transforms the time domain Maxwell's equations into frequency domain equations and conditions them with suitable boundary conditions. The equations arising from this method will be in a form of complex gradient and curl products, which can be solve in terms of Eigen-modes. Some examples of this Fourier domain techniques are Plane wave expansion (PWE), Rigorous wave coupled analysis (RCWA), C-method and Finite element method (FEM).

The disadvantage of this method is complicated Eigen-modes solving process which requires matrix inversion and sometimes the matrix can be ill conditioned. Some additional algorithm might be required in order to enhance the stability for the matrix inversion such as Legendre expansion RCWA (Khavasi *et al* 2007), transmission matrix algorithm (Moharam *et al* 1995b).

For single frequency sources, these methods provide stable and reliable solutions.

Plane wave expansion (PWE) and finite element method (FEM)

PWE is a very successful electromagnetic simulation technique and is widely used in designing photonic crystal devices. This algorithm can be applied with what we would like to achieve, however, the method is similar to FEM, where it is required to solve the solutions based on the consecutive mesh points. Therefore this is actually not fully utilising the periodic boundary condition of the structure. Therefore to solve one single input on the back focal plane it is required to solve a full set of spatial domain of the sample, which obviously slows down our calculations.

Finite element method (FEM) is one of the most successful electromagnetic wave simulation techniques and also very successful commercial packages are available such as FEMLAB-COMSOL; it is also capable of simulating any periodic and non-periodic boundary conditions. The technique is required to set a spatial domain grid on the sample, so the accuracy of the solution is dependent on fineness of the grid. This technique has a similar problem with the PWE technique, where it will slow down the calculation required as it has to solve all the mesh points.

Rigorous wave coupled analysis (RCWA) and C-method

These techniques are only suitable for periodic structure where they calculate the solution by matching the tangential electric and magnetic fields at its boundary for each diffracted orders. These techniques are based on the diffraction theory calculation, in order words the accuracy of the solution is more dependent on the number of diffracted orders used in the calculation.

For RCWA, the real advantage is that it is capable of calculating a multilayer structure by slicing the structure into layers so that each layer is homogeneous in the direction of propagation. However, if the sample shape is more complex, one can use it with small changes in dimensions so that more complex profiles can be mapped with piecewise addition of layers. These shapes might be approximated by a rectangular blocks of multiple layers but while more layers approximates complicate shapes more accurately. This comes at a cost of increased computing time and memory.

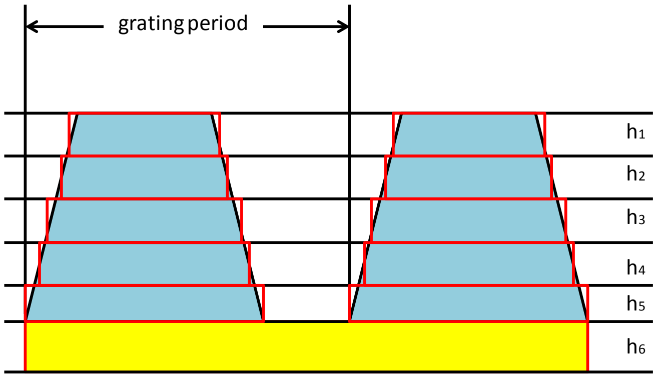
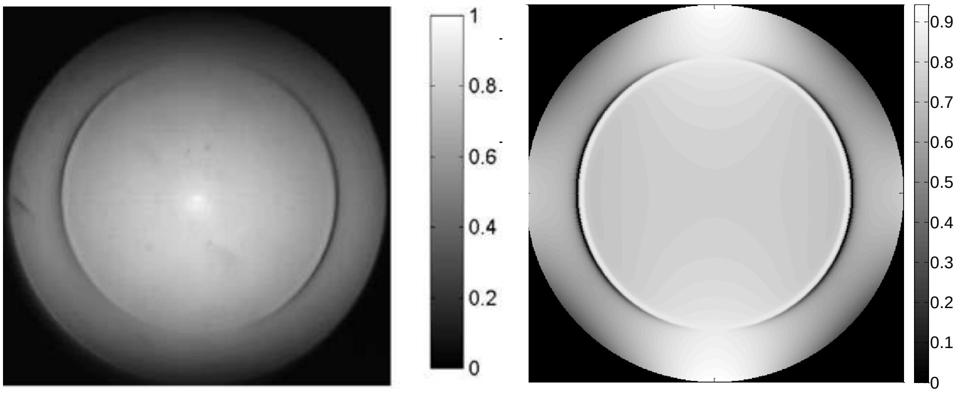


Figure 3.1 shows rectangular grid approximation on a triangular based grating

The C-method is actually the extended version of RCWA, where it has been improved to overcome the rectangular grid approximation required by RCWA. This is done by applying a coordinate transform to the structure so that it can give an accurate result without approximation. In other words, the real advantage of the C-method compared to RCWA is that it is capable of doing such geometries that are hard to achieve with RCWA. However the C-method is not suitable for a binary or rectangular grating as the algorithm requires a Fourier transform of the structure; this, of course, needs infinite number of harmonic orders to obtain accurate results.

For this project, the samples used in the simulation were binary grating samples, as the objective of the thesis is also to work out the theoretical limit of the lateral resolution that can be obtained from different SPR microscopy configurations; therefore we need an extreme case of the edge response. Therefore the obvious choice is RCWA.

3.3 Fresnel's equations



a) b)

Figure 3.2 shows a) experimental back focal plane intensity (Stabler et al 2004)
b) theoretical back focal intensity calculated using Fresnel's equations: $n_0=1.52$ (oil immersion refractive index), $n_1=3.1348+3.3103i$ (Chromium refractive index) with thickness of 1 nm, $n_2=0.1894+3.2994i$ (Gold refractive index) with thickness of 50 nm (Johnson & Christy 1972) in air: $NA=1.45$ and *wavelength=633 nm.*

I mentioned in the previous chapter that the Fresnel equations are actually an exact solution of Maxwell's equation where the structure simulated is a uniform layer of material. Therefore there is no error and approximation for this calculation. However it is necessary to check whether the equations can be used to reproduce the same results for some uniform SPR microscope cases published in order to validate the approach.

From figure 3.2b, we can see from the Fresnel's equation result that the strongest plasmonic dip is in pure *p-polarisation* direction, whereas there is no plasmonic dip in the other direction s-polarisation; these are correct according to the theory explained in the previous chapter. However, the result obtained from the experiment carried out by Stabler *et al* is different from the simulation result, this is due to the fact that the objective lens was not perfectly uniform distributed over the full aperture and also the system, of course, had some aberrations. It was also possible that the gold layer used in the experiment was not of perfectly uniform thickness. It is possible to take those discrepancies to account by adding a pupil function into the simulation.

For the SPR effect, the most important feature is the position of the plasmonic angle and its width as it indicates the properties of the SPs. From figure 3.3, we can see that the shape of the simulated plasmonic dips is similar to the experimental results; however with some discrepancies. The plasmonic angles for both cases are at 53 degrees and they both have similar width as shown in figure 3.3.

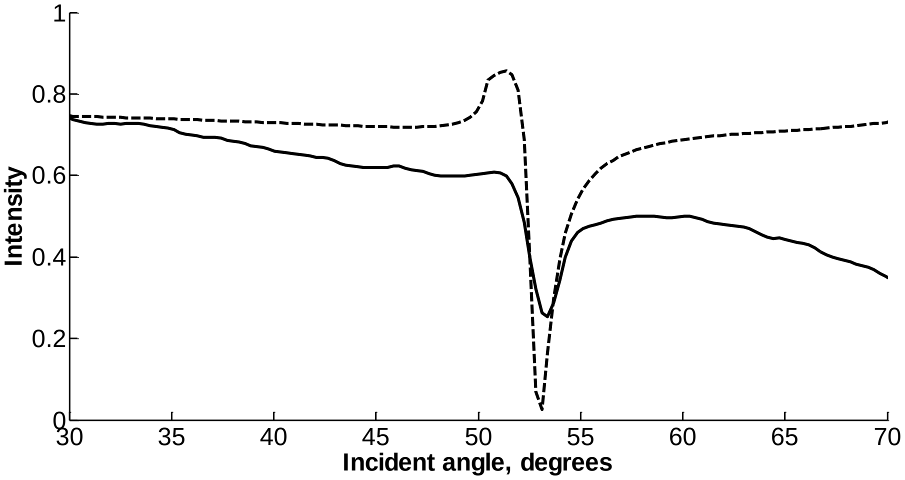


Figure 3.3 shows comparison between experimental and simulation plasmonic dips.
Solid curve is experimental result (Stabler *et al* 2004) and dashed line is simulation result.

3.4 Rigorous wave coupled analysis (RCWA)

The rigorous coupled analysis (RCWA) was employed to calculate the grating samples cases in this study. The RCWA software used in this thesis is called Rigorous Optical Diffractive Software (RODIS) written in C with python complier interface. The software has been developed by Ghent University. A frontend interface has been built and developed by the author to fully utilize multi-core computing technology to perform multi tasking calculations in parallel. The frontend has been written so that it has a user friendly interface under a Matlab environment. The software has also been validated with my own RCWA software written in Matlab; both programmes gave same results. However the C version is faster as it fully utilizes complied lower level C code, so it runs faster compare to the Matlab version.

In this section, the theory of RCWA 1D and 2D are provided. Some calculations using RCWA was done to show that the stable solution can be obtained with sufficient number of diffracted orders in the calculations and was also compared with some experimental results in the literature and simulation results from other simulation techniques.

3.4.1 RCWA theory for 1D and 2D

The concept of RCWA is to solve Maxwell's equations in the frequency domain. Therefore the accuracy of this calculation is dependent on the number of harmonic frequencies or number of diffracted orders included in the calculation. The calculation is done by solving the Maxwell's equations using Fourier series of electric and magnetic unknown coefficients, and then a set of matrix equations can be formed, so that the matrix can be solved as an eigenvalue problem. The eigenvalues obtained are then conditioned by making use of the fact that the sample is periodic. This enables us to solve the electric and magnetic fields normal to the plane of diffraction plane. The detailed equations for solving the 1D and 2D cases are provided in appendix A.

For the dielectric grating case, a practical limit to the number of diffracted order can be estimated as:

$$(k_0 n_I)^2 - (k_0(n_I sin\theta\ cos\phi - N\frac{\lambda_0}{\Lambda}))^2 - (k_0 n_I sin\theta\ sin\phi)^2 = 0 \qquad [3.1]$$

Where k_0 is the free space wave vector, n_I is the first homogenous layer refractive index. θ is angle of incidence. ϕ is azimuthal angle. λ_0 is free space wavelength. Λ is grating period. N is the number of diffracted order obtained from the above equation.

This is simply the condition that above this value of N both positive and negative diffracted orders become evanescent, so it is very important to use more than the number of diffracted order calculated from equation [3.1].

It is worth pointing out that for the uniform structure case RCWA matrix calculation is actually the Fresnel's equations (see details in appendix A).

3.4.2 RCWA 1D validation

This is the main tool used in this thesis, therefore it is important to validate the RCWA 1D software for all the cases presented in this thesis. However, this section will only show the process of validation by determining the amount of error for a few cases and compare them with a literature. All the validations for all the other cases were also done to ensure that the results obtained are correct and they are presented in appendix B.

Uniform sample case validation

As mentioned earlier the RCWA method for the uniform case is the Fresnel equation. Therefore the validation is done simply by using the smallest number of diffracted orders, which *N* order input of 1 corresponds to the total number of diffracted order of 3, as in theory there will be no diffracted order for the uniform sample case anyway.

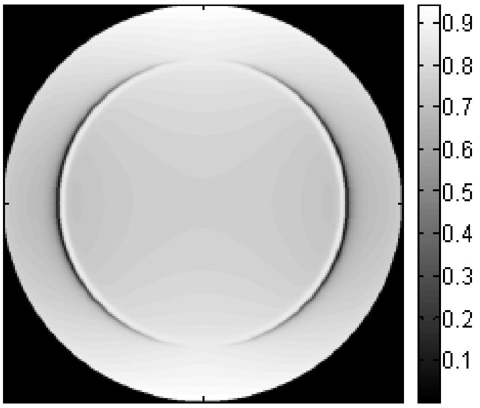


Figure 3.4 shows theoretical back focal intensity calculated using RCWA: $n_0=1.52$ (oil immersion refractive index), $n_1=3.1348+3.3103i$ (Chromium refractive index) with thickness of 1 nm, $n_2=0.1894+3.2994i$ (Gold refractive index) with thickness of 50 nm (Johnson & Christy 1972) in air: $NA=1.45$ and $wavelength=633\ nm$

Figure 3.2b agrees with figure 3.4 for all input angles. Therefore it is reasonable to use the RCWA technique with N order input of 1, which corresponds to the total number of diffracted order of 3 for the uniform sample case to calculate accurate results.

1D Grating cases

The most important features are the capabilities to accurately calculate the grating cases, therefore it is important to test convergence for all the conditions used in this thesis. All the conditions used in this thesis are listed in appendix B. In addition, in this section I will show the validation for a few interesting cases and compare the results with the literature.

The literature I compared the simulations results with was (Stabler *et al* 2004) for the cases of *X* and *Y polarisations*, where the simulation parameters are:

$n_0=1.52$ (immersion oil refractive index), $n_1=3.1348+3.3103i$ (Chromium refractive index) with thickness of 1 nm, $n_2=0.1894+3.2994i$ (Gold refractive index) with thickness of 50 nm (Johnson & Christy 1972) in air: $NA=1.45$, $wavelength=633\ nm$ and 6 microns of $n_{grating}=1.5$ in 8 microns grating period.

For any particular case, it is necessary to check all input (k_x,k_y) values to show that the solutions we obtained on the bfp converge to the stable solutions. Especially for grating cases, one (k_x,k_y) input will give us diffracted fields corresponding to the grating period. The number of orders was varied until there was little variation in the propagating diffracted orders.

The way that the cases have been validated was to calculate the difference in bfp between two consecutive numbers of diffracted values; this is done to ensure that all the input angles give us the convergent results. The way I resolved the electric and magnetic fields responses on the bfp is described in microscope simulation in section

3.5. In principle the absolute correct results can only be obtained with an infinite number of diffracted orders, which is, of course, impossible in practice. Therefore what we need are the solutions that are good enough to accurately represent the problems. Figure 3.5 shows how the RCWA converges as the number of diffracted orders increases. The value of N on the x-axis represents the difference between N and N-1 diffracted orders. We note that convergence at the plasmonic angle is somewhat slower than other angles since the diffraction efficiency is increased at this angle. The case given as example in figure 3.5 was the *X polarisation* grating sample case in (Stabler *et al* 2004).

The maximum difference in bfps between two consecutive N values allowed in the simulations is 0.0001 in intensity, which suggests that all the results obtained are reliable up to the fourth decimal point.

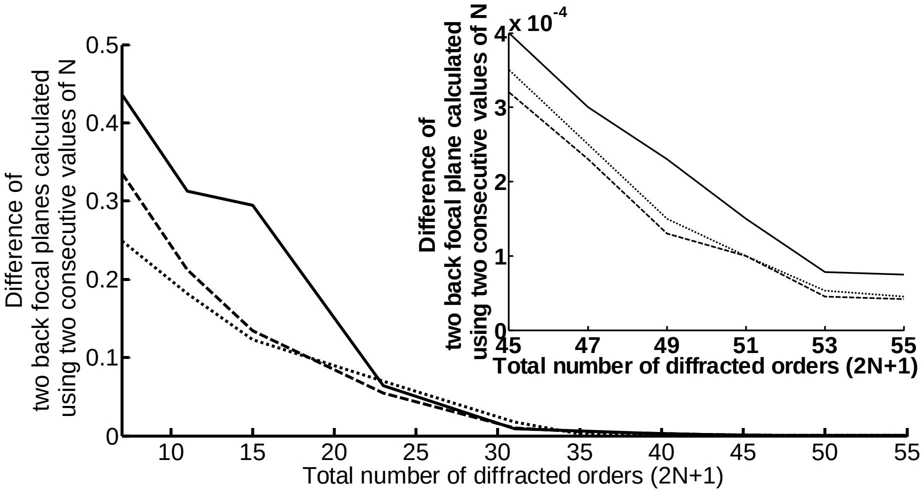
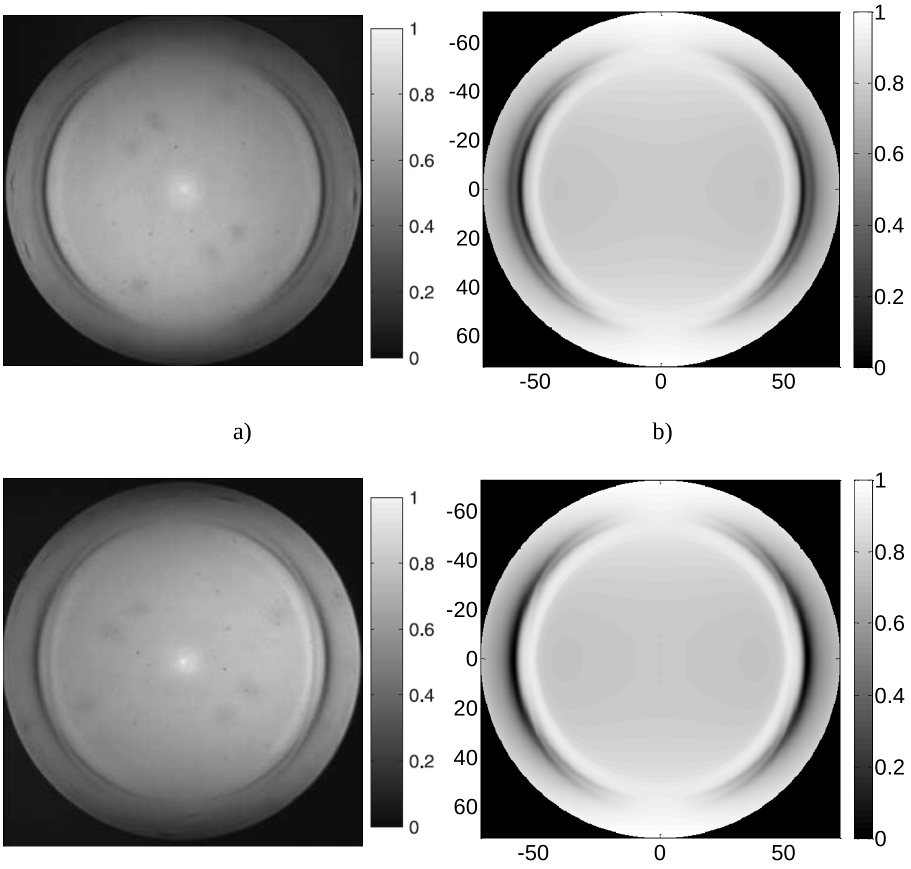


Figure 3.5 shows the difference between three *(k*$_x$*,k*$_y$*)* points on the back focal plane as the number of diffracted orders is increased. The x-axis represents the modulus of the difference between $2N+1$ prders and $2(N-1)+1$ orders. Solid line for $(k_x,k_y)=(13.3650,0)$ at the plasmonic dip , dashed line for $(k_x,k_y)=(0,0)$ and dotted line for $(k_x,k_y)= (8.1751,-9.448)$. $n_0=1.52$ (oil immersion refractive index), $n_1=3.1348+3.3103i$ (Chromium refractive index) with thickness of 1 nm, $n_2=0.1894+3.2994i$ (Gold refractive index) with thickness of 50 nm (Johnson & Christy 1972) in air: $NA=1.45$, *wavelength=633 nm* and 6 microns of $n_{grating}=1.5$ in 8 microns grating period.

The number of array points for all the calculations was set to 501x501, which corresponds to the k_{vector} step size of 0.058, this number was sufficient to recover plasmonic dip information, where the steepest dip was represented by 6 array points on the back focal point. This enables us to calculate the data between the discrete points by interpolation and this also give use an accurate value for the plasmonic dip amplitude and plasmonic angle.

The bfp results shown in figure 3.6 were calculated by using the total number of diffracted orders of 53, which gave the maximum difference on the back focal plane of 0.000078 in intensity relative to the maximum intensity on the bfp of 0.89.

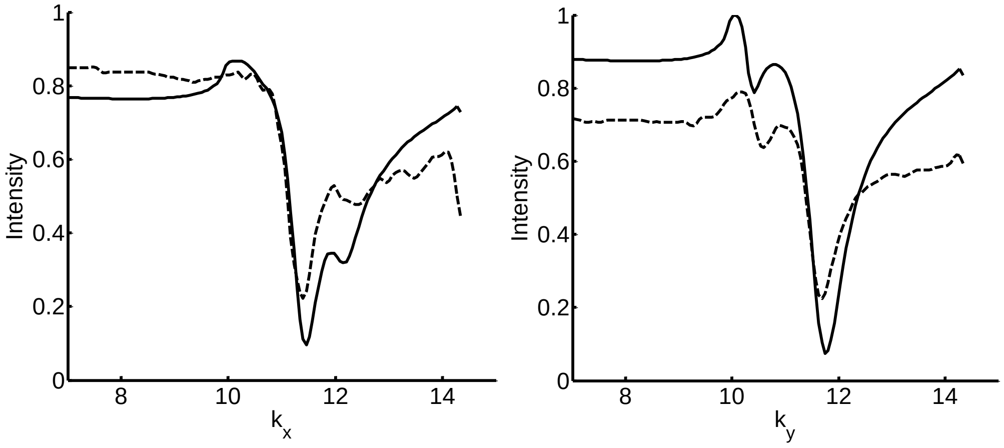


a) b)

c) d)

Figure 3.6 shows a) experimental bfp result for *X polarisation* (Stabler *et al* 2004) b) simulation bfp result for *X polarisation* calculated using RCWA with 53 numbers of diffracted orders c) experimental bfp result for *Y polarisation* (Stabler *et al* 2004) d) simulation bfp result for *Y polarisation* calculated using RCWA with 53 numbers of diffracted orders

As mentioned earlier that there are some discrepancies between the simulation and experimental results, these caused by the pupil function of the objective lens, aberrations, the sample itself and the noise in the experimental results. The position of the plasmonic dips has been checked to ensure that the experimental results agreed with the simulation results as shown in figure 3.7.



a) b)

Figure 3.7 shows the plasmonic dips of experimental results (Stabler *et al* 2004) compared with simulation results for a) *X polarisation* b) *Y polarisation*. Solid curves for simulation results and dashed curves for experimental results

In order to run a set of full bfp microscope calculations, it is necessary to call the RCWA function approximately 200,000 times for 501x501 points on the back focal plane; therefore it is sensible to be able to approximate and roughly determine how long the calculation will take. The computing run time required to calculate one RCWA calculation for different number of diffracted orders is shown in figure 3.8.

It is important to note that the run-time is not only dependent on the total number of orders, but also the number of layers used in the calculation. The ten layers case requires two times more of the computation time than the five layers case as it has to perform matrix calculation for two times more (Moharam *et al* 1995b).

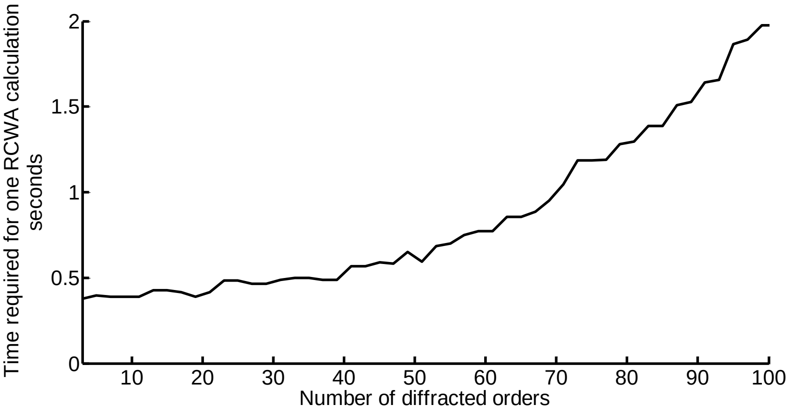
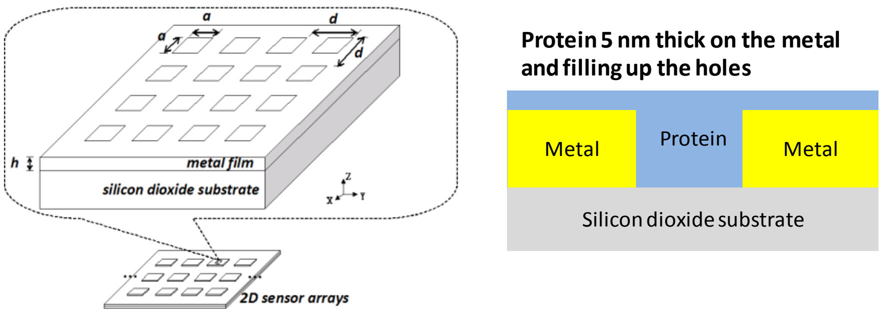


Figure 3.8 shows RCWA 1D run time for different number of diffracted orders: this run time was obtained using Intel i7 machine running at a clock speed of 3.06 GHz with 14 GB of memory. The multi-threading calculation was disabled during this calculation. This was calculated with 5 structural layers.

3.4.3 RCWA 2D validation

RCWA for 2D grating has not been used in this thesis, however my research group has extended the capabilities of the RCWA 2D to design some nano-optical structures for multi focal points for parallel confocal scanning. Therefore it is sensible to also validate the 2D grating system. The way I validated it was by comparing RCWA 2D simulation results with a literature (Yang & Ho 2009), where they used finite different time domain (FDTD) technique to study the effect of rectangular nanohole array in transmission mode. They showed that the phase change of the light transmitted through the hole array is relatively large due to multiple constructive interference modes. The maximum phase change of 52 degrees occurs at 427.6 nm square hole period with gold thickness of 108 nm for refractive index change from n=1.33 to 1.5. The *p-polarised* light was coming in at normal incidence with its electric field along the x axis corresponding to the coordinate system described in figure 3.9a.

Simulations parameters: $n_{gold} = 0.1894+3.2994i$ and $n_{silver} = 0.0595+4.2004i$ at 633 nm wavelength (Johnson & Christy 1972) rectangular nanohole varying the hole period 50% on-off ratio, with the metal thickness of 108 nm and 5 nm of protein layer sitting on and within the holes. The *p-polarised* light is coming at normal incidence from the water side.



a) b)

Figure 3.9 shows a) Schematic of nanohole array (*a* = hole width, *d* = hole period, *h* = hole depth) (Yang & Ho 2009) b) The side view of the nanohole device

The nature of the 2D grating is that it has 2 directional grating periods therefore, for the 2D case, the RCWA calculation is done by mapping the tangential fields in 2D field system (see detailed equations in appendix A). The software requires us to specify two numbers of diffracted orders, namely number of diffracted order in *x* axis and number of diffracted order in *y* axis. The actual number of orders calculated is $(2N_x+1)(2N_y+1)$; the similar idea with 1D RCWA is also applied here where the higher the number of diffracted orders used is implying that the more accurate results can be obtained.

Even though the total number of diffracted orders for 1D is the same as the number of diffracted order for the 2D case, the total number of tangential electric and magnetic fields mapping, size of matrices calculations and memory required are different, in addition the 2D case requires much more computing power than the 1D case.

On the 14 GB of memory machine I cannot perform the calculation with more than 1681 diffracted orders, which correspond to 20x20 orders for N_x and N_y. Although the maximum number of diffracted orders was used to repeat the results shown in figure 3.11, convergent results still cannot be reached.

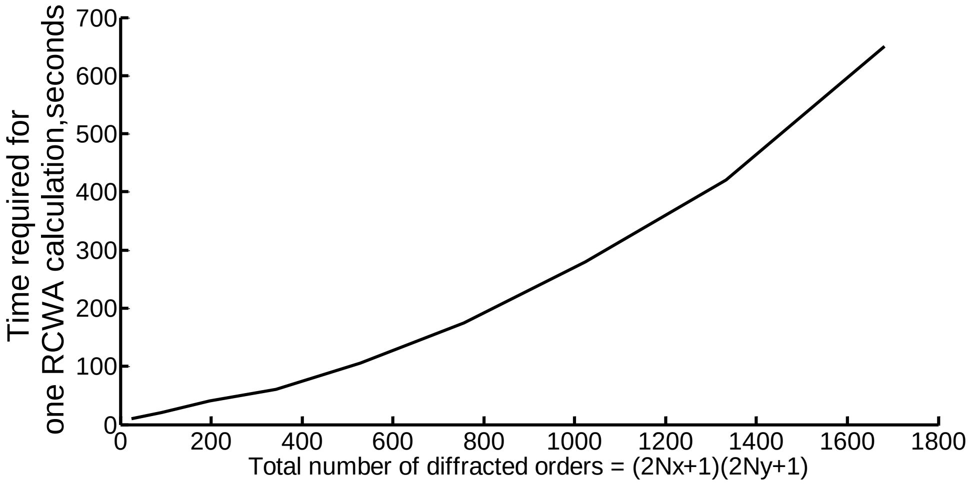


Figure 3.10 shows the run time for 2D case as a function of total number of diffracted order; the run-time was calculated for the case that N_x=N_y. This was calculated by 5 layers of structures.

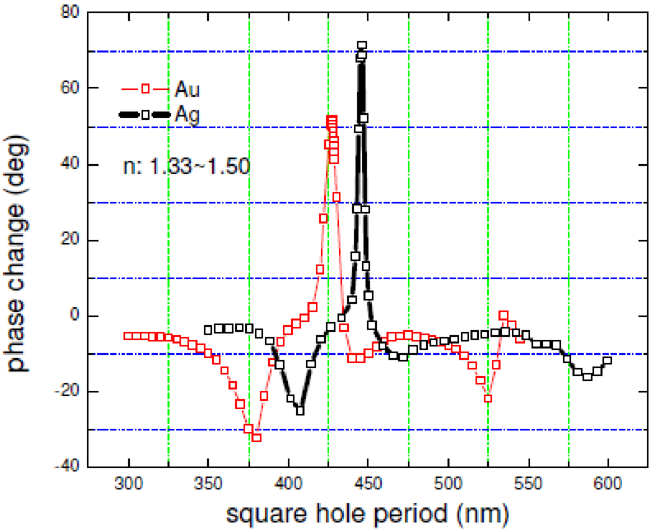


Figure 3.11 shows square hole period versus phase change calculated by FDTD (red line for gold and black line for silver) (Yang & Ho 2009)

Due to the fact that SPs can only be excited by *p-polarisation*, in other words, the *x-axis* actually contributes more to the SPR phase change effect compared to the *y-axis*; I have tried to get around this problem by setting the number of diffracted orders to 31x7 orders for N_xN_y and varying only the N_x to determine whether the solution approaches its convergence value. This approach worked for the case and gave us similar results to the figure 3.11, where the convergent results were obtained with 41x7 for gold and 43x7 for silver respectively as shown in figure 3.12.

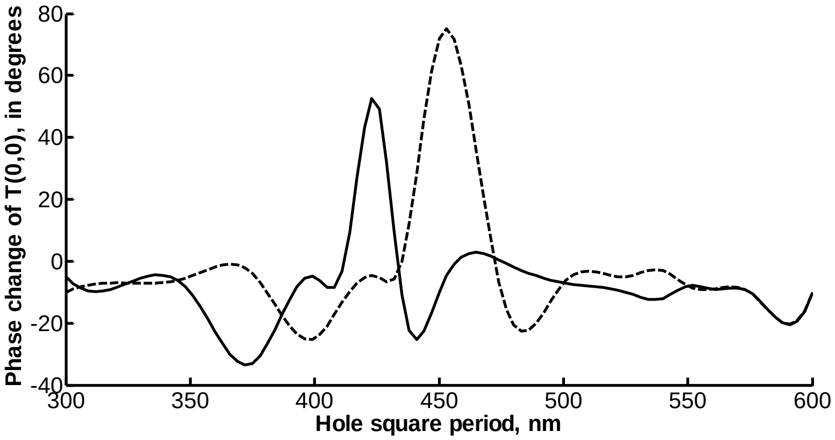


Figure 3.12 shows square hole period versus phase change calculated by RCWA. Solid line for gold and dashed line for silver.

The results shown in figure 3.12 are very similar to those shown in figure 3.11, the shapes are however different especially the size of the peaks; the FDTD seemed to show narrower peaks compared to RCWA.

3.5 Microscope simulation

We examine the microscope simulation algorithm for a 1D grating sample. However the same idea can be also applied with transmission mode microscope calculations and also microscopes other than the SPR microscopes.

In order to calculate SP microscope responses, it is required to calculate either the Fresnel function for uniform layer case or RCWA for grating case. Therefore in this section, I will explain how to calculate microscope responses in detail.

3.5.1 Step 1 Back focal plane array and microscope parameter setups

The number of array points is specified so that the minimum number of discrete spatial incident angles can be used without distorting the distribution in the bfp. In this study, I used 501x501 points on the back focal plane array which corresponds to 0.007 radian around the plasmonic angle. I then used parameters of the microscopes to calculate k_{vector}, k_x,k_y,k_z, incident angle θ and azimuthal angle ϕ for every point on the back focal plane by:

$$k_{vector} = \frac{2\pi n_0}{\lambda_{free\,space}} \qquad [3.1]$$

$$k_{xmax} = k_{ymax} = 2\pi NA/\lambda_{free\,space} \qquad [3.2]$$

The bfp array is then assigned its (k_x, k_y) values for every point on the bfp corresponding to their positions on the bfp. k_z is used to specify the circular area on the bfp.

$$k_z = \sqrt{k_{vector}^2 - k_x^2 - k_y^2} \qquad [3.3]$$

$$\theta = sin^{-1}\left(\frac{\sqrt{k_x^2+k_y^2}}{k_{vector}}\right) \qquad [3.4]$$

$$\phi = tan^{-1}\left(\frac{k_y}{k_x}\right) \qquad [3.5]$$

For the grating case we also need to specify the direction of the electric field with respect to the plane of incidence ψ as show in figure 3.13; as mentioned in chapter 2 there are three *polarisations X Y* and *R* respectively. The electric field direction with respect to phi can be calculated as:

For *X polarisation*: $\psi = \pi + \phi$ [3.6]

For *Y polarisation*: $\psi = \pi/2 + \phi$ [3.7] and

For *R polarisation*: $\psi = 0$ [3.8]

We can then call either Fresnel or RCWA calculations for every input $(k_x, k_y, \theta, \phi, \psi)$ on the back focal plane.

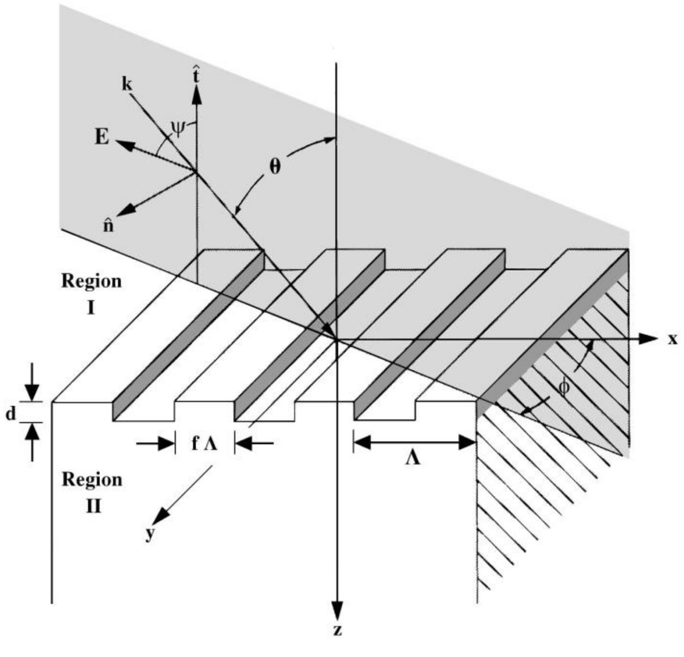


Figure 3.13 shows θ,ϕ,ψ angles and the coordinate system used in this study (Moharam *et al* 1995a)

3.5.2 Step 2 Calculate electric field outputs on the back focal plane

There are two options here; Fresnel function and RCWA function. The Fresnel function is for uniform layer cases, and RCWA is for both uniform and grating structure cases. Therefore in this section, I will explain the way of resolving the electric field outputs for both functions.

Step 2.1 Calculate electric field outputs using Fresnel's equations

For uniform sample cases, there will be no diffracted order which means the field will be reflected at the same angle. Therefore we can construct arrays to collect r_p and r_s outputs from the Fresnel function and keep them at the corresponding input array position. If necessary, we can then convert the r_p and r_s arrays into E_x and E_y arrays by:

$$E_x = r_p \cos\theta \cos\phi + r_s \cos\theta \sin\phi \qquad [3.9]$$

$$E_y = r_p \cos\theta \sin\phi - r_s \cos\theta \cos\phi \qquad [3.10]$$

Where E_x is an array consisting of the complex electric field amplitude pointing to the x axis and E_y is an array consisting of the complex electric field amplitude pointing to the y axis. A vector diagram to illustrate the field resolving directions is shown in figure 3.14.

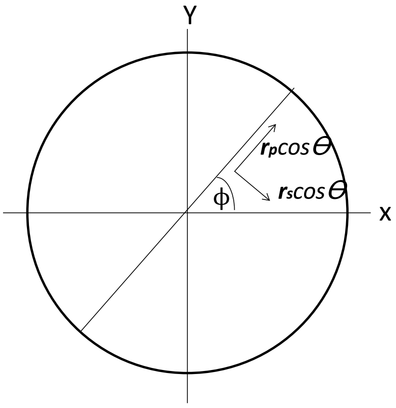


Figure 3.14 shows r_p and r_s resolving to the x and y directions

Step 2.2 Calculate electric field outputs using RCWA

For the grating case, one input on the back focal plane will give us *(2N+1)* pairs of complex electric and magnetic fields of diffracted orders where all the fields are pointing at the direction perpendicular to the plane of incidence, where N is the number of diffracted order specified in the programme. Here we need to allocate sufficient memory to store the output fields for each diffracted order. Then magnetic fields are then converted to the fields parallel to the plane of incident using Maxwell's equation as described in (Moharam *et al* 1995a).

It is necessary to ensure that all the electric fields collected are propagating wave by calculating k_{xm},k_{zm} and θ_m for each diffracted order outputs by:

$$k_{xm} = k_x - mk_g \qquad [3.11]$$

Note that we use a negative sign to be consistent with (Moharam *et al* 1995a).

Where m is the m^{th} diffracted order and k_g is sample grating vector, which is given by:

$$k_g = \frac{2\pi}{\Lambda} \qquad [3.12]$$

Where Λ is the sample grating period.

$$k_{zm} = \sqrt{k_{vector}^2 - k_{xm}^2 - k_y^2} \qquad [3.13]$$

$$\theta_m = sin^{-1}\left(\frac{\sqrt{k_{xm}^2 + k_y^2}}{k_{vector}}\right) \qquad [3.14]$$

The k_{zm} term allows us to exclude the evanescent field outputs. The θ_m term also enables us to determine which reflected diffracted order can be collected by the objective as we know the maximum angle that can be captured by the objective lens.

In some calculations, aberrations and pupil functions are required to be included in the calculations, such as amplitude SLM masking, annulus excitation for the non-interferometric method and *V(z)* calculation, where the output is also dependent on the pupil function as discussed in chapter 2. If a bfp mask is required to be included in the calculation, the masking process can be done by multiplying the mask values to the diffracted field outputs on the bfp during this process of calculation.

As we know the corresponding positions of all the electric field diffracted orders on the bfp, we can then resolve them into their correct positions on the back focal plane according to the (k_{xi},k_y) coordinates calculated.

It is important to note here that a proper field averaging is required as some of the fields might not be able to be perfectly fitted in with (k_x,k_y) bfp array. For example, a diffracted order with $(k_{xi},k_y)=(1.09,0)$ however the available positions around this region are $(k_x,k_y)=(1.1,0)$ and $(k_x,k_y)=(1.08,0)$ as shown in figure 3.15. The algorithm used to average the field is weighted average.

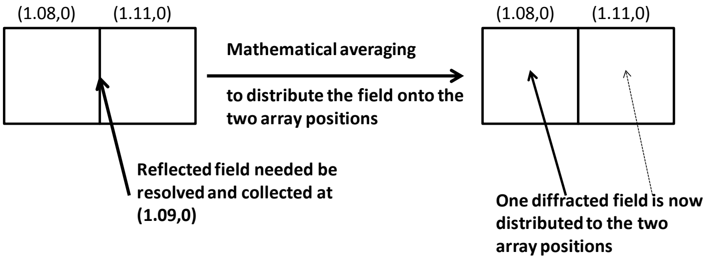


Figure 3.15 shows an averaging algorithm required to resolve some of the diffracted orders onto the bfp.

3.5.3 Algorithm for V(x) calculation

Once the back focal plane distribution is calculated at a specific lateral position x_0 and defocus z, the distribution for any other position can be calculated by appropriate phase shifts. For the lateral position x_1, the phase shift term can be calculated as:

$$Phase\ shift = \exp\left(i(k_x - k_{xm})x_1\right) \quad [3.15]$$

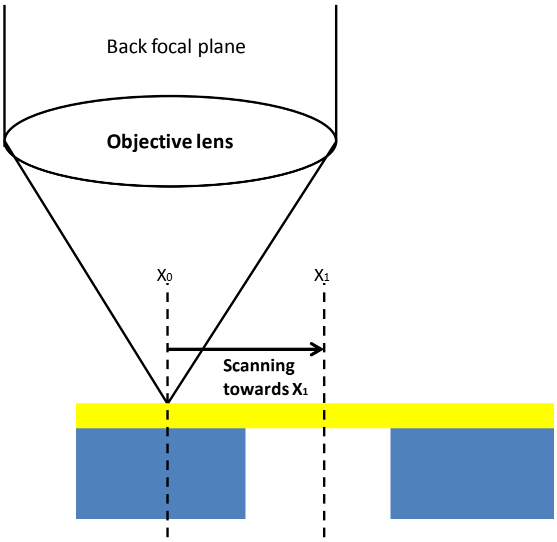


Figure 3.16 shows sample and phase shift required scanning along the sample axis

3.5.4 V(z) calculation for interferometric SPR microscopes

And for the z position the additional phase shift is given by:

$$Phase\ shift = \exp\left(i(k_z + k_{zm})z\right) \qquad [3.16]$$

Where *z* is the defocus distance; negative *z* means the sample is moving towards the objective.

The summation of the phase factors comes from the fact that the propagating waves travel a distance z from the focus both on incidence and return.

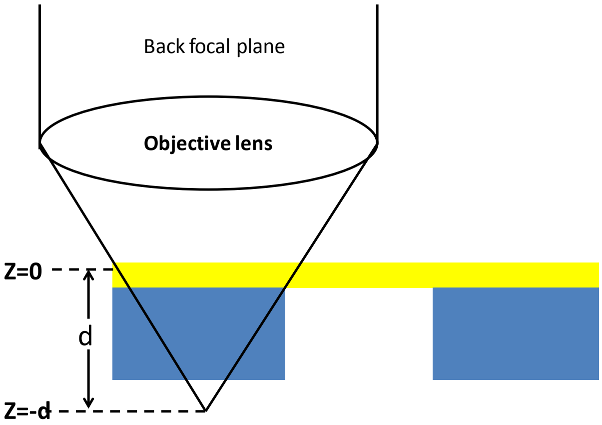


Figure 3.17 shows sample and phase shift required scanning along the optical axis

The details of the *V(x)* and *V(z)* calculations using the back focal plane distributions are described in chapter 4.

3.5.5 Diagram of computing process

All the calculation processes can be schematically described by the diagram shown in figure 3.18.

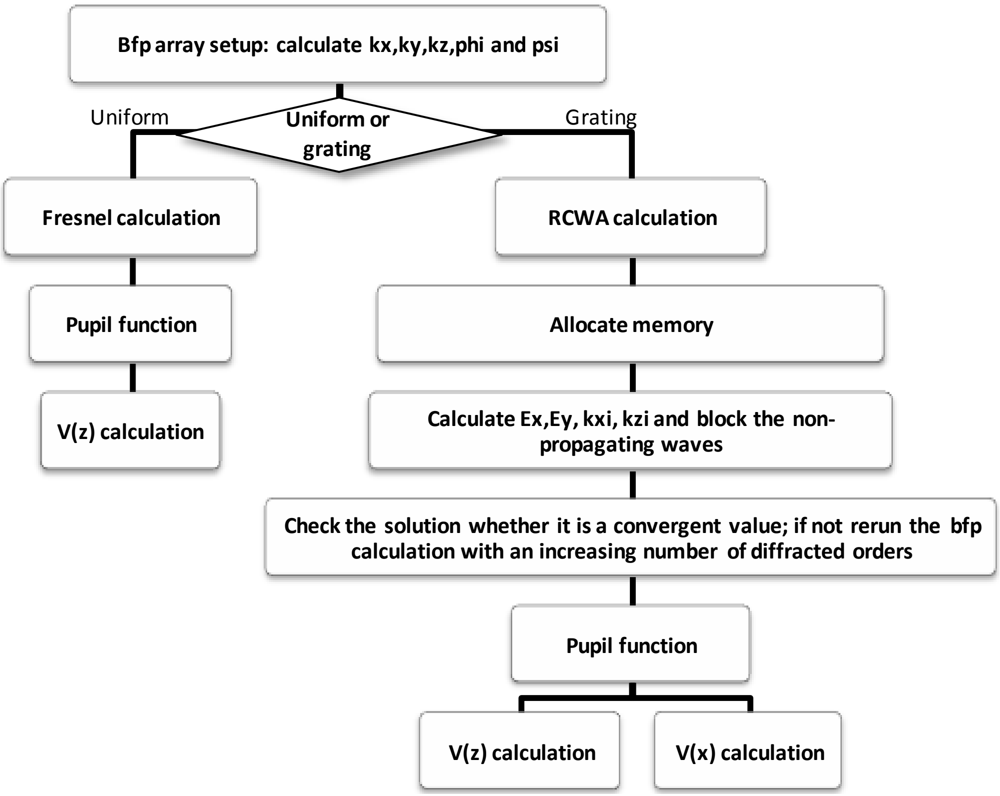


Figure 3.18 shows process to calculate microscope responses

3.5.6 Parallel computing choices and techniques for speeding up calculation

Let us consider the calculation time for one full set of bfp calculations; the number of points on the bfp is 501x501=250,001 points, where there are about 200,000 points inside the objective lens aperture. The number of diffracted orders used in the thesis was at least 20 orders (41 diffracted orders in total) which requires about 1 second for one RCWA call and for additional processes in Matlab such as, file reading and memory allocation. The run time required for these calculations sum up to 200,000 seconds or 2.32 days. The required computational time can be improved by employing the following computing techniques. Using the multiple core techniques described below enabled us to complete a full bfp calculation within 6 hours.

Parallel computing choices

Thanks to multi-core technology, the microscope responses code can be written so that it can fully utilise all the computer cpu cores. The computer used for this study was equipped with Intel i7 running at 3.06 GHz and 14 GB memory. The cpu unit comes with 4 cpu cores and each core has 2 computing threads, so in total one can use 8 parallel computations. This allows us to speed up the computing time by a factor of 8. This is done by separating the back focal plane into 8 areas having the same size and then writing 8 sets of input files to Python and then 8 sessions of Python can be run in parallel.

Calculating RCWA in python without calling Matlab

Matlab normally required a number of seconds to boot the software up; therefore some of the computing time might be spent to call Matlab. This can then be speeded up by avoiding calling Matlab and instead running all RCWA calculations in Python, which runs under DOS. Matlab is then only required twice in the process of back focal plane calculation, which writes input files to Python allowing different tasks (associated with different input incident angles) to be allocated to separate cores. Once these processes are completed, Python sends the output from each core back to Matlab which collates the data as a full map in the back focal plane. This is shown in the figure 3.19.

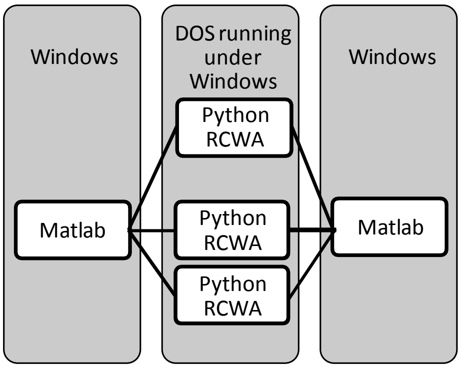


Figure 3.19 shows parallel computing steps used in microscope software

Pre-just in time compilation

For a lengthly calculation involving a large number of diffracted orders, the computational time can be speeded up by making use of Just In Time technology. Just In Time (JIT) is a computer architecture that stores the previously called instructions and if the same instruction is called again with a new input; the new input will be computed at faster speed as the command lines are not needed to be compiled into machine language again. In order to give an idea about how pre-JITed might be used in the microscope calculation let look at the table below.

Number of orders and calculation conditions	**Time to calculate each process**
Run 3 diffracted orders and run 200 diffracted orders separately (1D RCWA)	0.38 seconds for 3 diffracted orders and 5.76 seconds for 200 diffracted orders
Run 3 diffracted orders and run 200 diffracted orders continuously after completing the first calculation (1D RCWA)	0.38 seconds for 3 diffracted orders and 4.26 seconds for 200 diffracted orders

Table 3.1 shows comparison of computing with and without pre-JITed

From table 3.1, we can see that pre-JITed compilation is faster than the other case, where the large number of diffracted order was run separately. Therefore for the large number of diffracted orders, the pre-JITed algorithm is employed by performing the same calculation with very low number of diffracted orders first before running the large number of diffracted order right after the low number of diffracted orders case.

Chapter 4

SPR microscopy

In this chapter, we will explore the imaging performance of the non-interferometric and interferometric SPR setups. The aim of this chapter is to firstly develop a theoretical framework to understand SPR microscopy performance, secondly to determine the theoretical limits of such systems and also to be able to address the limits of each system. The other objective of this chapter is to compare the performance advantages and disadvantages of the two types of configurations.

In order to calculate microscopy performance, we need a sample that has a sharp edge in order to see the edge response effect and hence determine the lateral resolution. In this study a silicon nitride Si_3N_4 1D rectangular grating with thickness of 20 nm, refractive index of 2.023, 8 micron period, 3:1 aspect ratio (6 microns Si_3N_4 and 2 microns of free space) sitting on a uniform gold layer of 50 nm thick as depicted in figure 4.1. This structure enables us to compare the experimental results in (Stabler *et al* 2004) with simulation results. The microscope was setup with $NA=1.49$ with $n_0=1.52$ and 633 nm wavelength in air ambient.

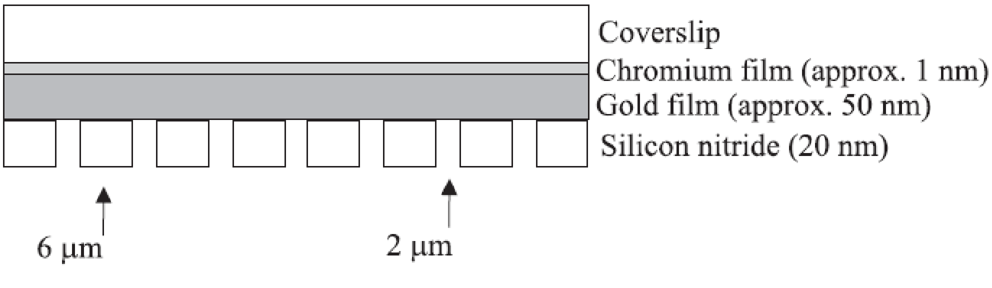


Figure 4.1 shows schematic diagram of the sample used in (Stabler *et al* 2004)

The chapter will begin with definition of terms used throughout this chapter and follow by non-interferometric and interferometric setups.

4.1 Imaging performance key terms and their definitions

4.1.1 Contrast ratio

Contrast ratio is defined as the ratio of the two central point intensities of on and off regions of the grating (points ‘a’ and ‘b’ on figure 4.2). The contrast ratio also indicates the sensitivity of the system; here we can tell how well we can distinguish between two refractive index regions in intensity. The reason why I used contrast ratio rather than actual contrast difference in intensity is to compare the response between the interferometric and non-interferometric microscopes as images obtained from both systems have different background levels and scaling factors.

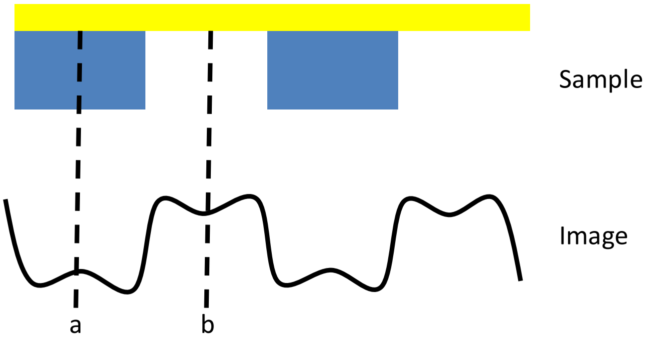


Figure 4.2 shows two points mechanism

From the figure 4.2, the contrast ratio can be calculated as:

$$contrast\,ratio = \left|\frac{I(a)-I(b)}{I(a)+I(b)}\right| \qquad [4.1]$$

4.1.2 10%-90% Lateral resolution

Lateral resolution is an important feature in imaging systems, where it indicates how well we can resolve small objects. In this study, the lateral resolution is defined as the 10%-90% edge response, where the lateral resolution is measured from the distance that the edge response changes its intensity from 10% to 90% or 90% to 10% of its maximum value as shown in figure 4.3.

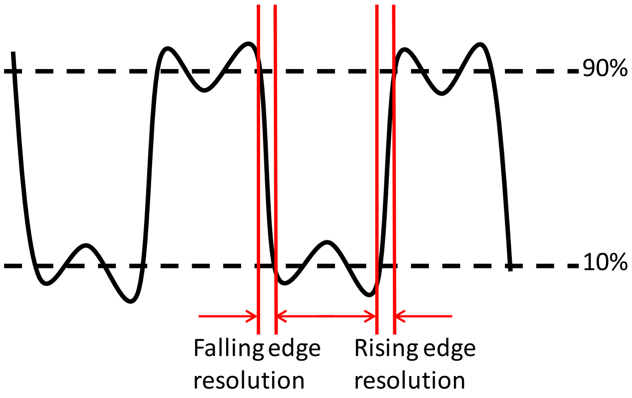


Figure 4.3 shows mechanism to determine lateral resolution

Note that this resolution definition might not be valid for all the cases as some images obtained from both experiments and simulations sometimes show funny edge responses; these will not be included in this study.

4.1.3 Aspect ratio

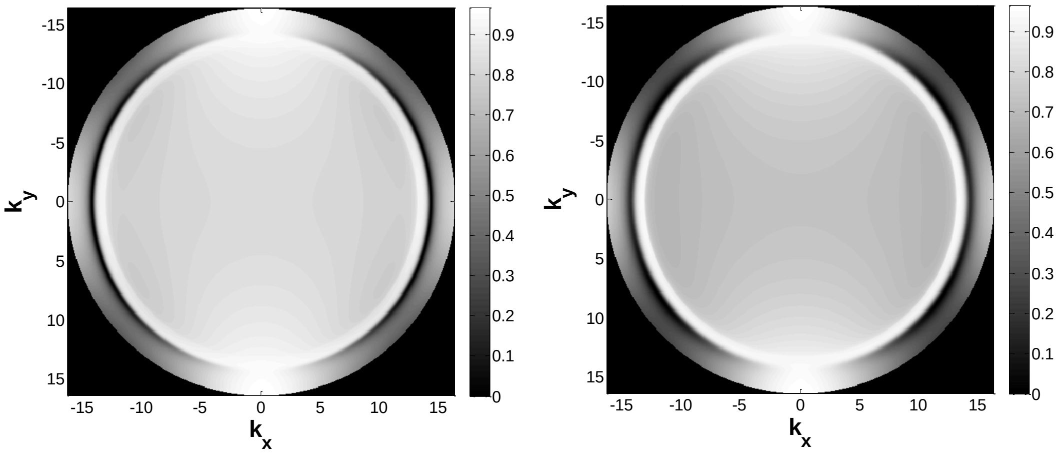
Aspect ratio in this study is defined as a ratio between the length of silicon nitride region and bare gold region measured from grating image at the reference point of 50% intensity ($50\%\ intensity = 0.5(I_a + I_b)$).

4.2 Non-interferometric SPR setup

Let us consider the wide field system as shown in figure 2.26 with a uniform sample in order to understand the SP effect on the setup.

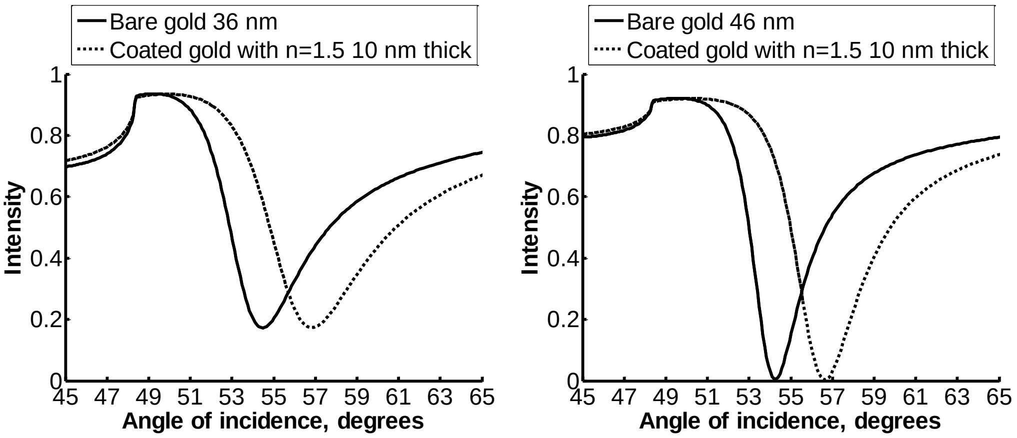
As mentioned earlier, the advantage of using Kretschmann based objective lens excitation is that all the azimuthal angles can be captured using one physical camera. Therefore it is clear to see that the SPs can only be excited with *p-polarisation* from figure 4.4.

Although the images in figure 4.4 look similar, from figure 4.5, we can see that the plasmon dips for the 46 nm cases are deeper than the 36 nm cases, however, the positions of the dips are almost at the same position for both the 36 nm and 46 nm cases. In other words, the thickness of the metal layer only affects the strength of SP coupling, not the coupling angle. This is potentially good in terms of optical setups as a different thickness of metal can be tested without having to realign the optical components. The diagram also illustrates the key feature of the SPs, which are the SPs coupling change corresponding to the change in the refractive index of the sample. For the 36 nm case, the plasmonic angle for the bare gold case is 54.46 degrees and 55.63 degrees for the 10 nm coated layer. For the 46 nm case, the plasmonic angle for the bare gold case is 54.27 degrees and 55.42 degrees for the 10 nm coated layer.



a) b)

Figure 4.4 shows a) back focal plane for 46 nm gold with *n=1.33*
and b) back focal plane for 36 nm gold with *n=1.33*
$NA=1.65$, $n_0=1.78$, $n_1= 0.1894 + 3.2994i$ linear polarisation with *633 nm* wavelength



a) b)

Figure 4.5 shows a) plasmonic dips for *36 nm* gold with *n=1.33 and n=1.5*
b) plasmonic dips for *46 nm* gold with *n=1.33 and n=1.5*
$NA=1.65$, $n0=1.78$, $n1= 0.1894 + 3.2994i$ *p-polarisation* with *633 nm* wavelength

The displacement in plasmonic dip position provides the sensitivity of intensity based SPR instrument. It is interesting to quantify the sensitivity of the SPR sensor and also determine the theoretical sensitivity limit. The details of sensitivity quantification will be provided in the next chapter.

4.2.1 Back focal plane response at the presence of grating sample

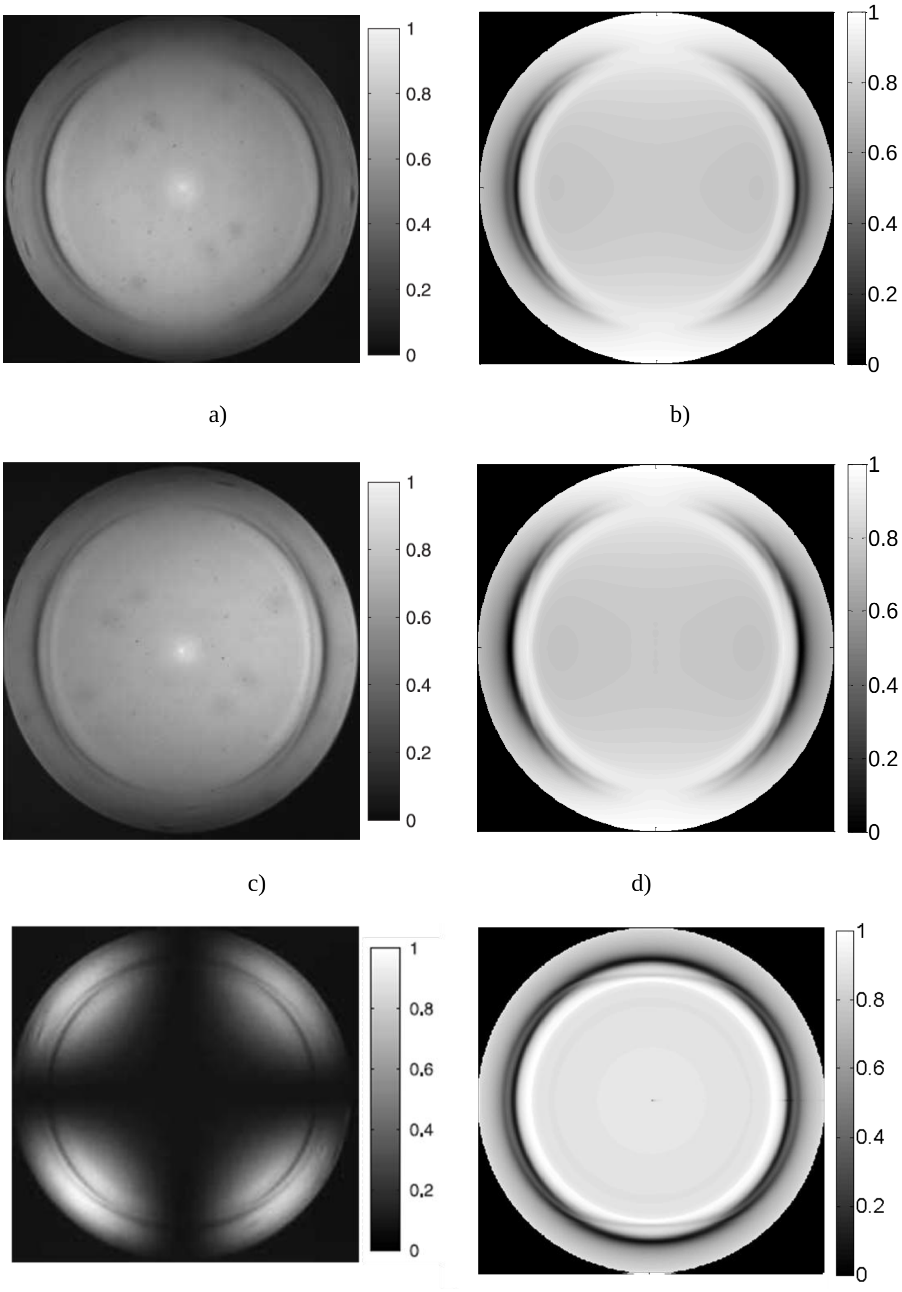
It is important to understand the effect of grating and plasmonic dip behaviour on the back focal plane when there is a grating. The grating will, of course, scatter the diffracted light on the bfp. Let us consider the back focal plane images of a wide field SPR microscopy when there is a grating as shown in 4.6. This was formed by summing the bfps from a scanning system over a grating cycle.

Figure 4.6 shows that the experimental back focal plane images agree well with the simulation results with some discrepancies in the intensity distribution on the back focal plane. This might be because firstly the actual thickness of gold in the experiment was, of course, not perfectly 50 nm. A thicker or thinner layer would, of course, increase the intensity of the SP dips. Secondly, it might come from the pupil function of the optical system; a uniform pupil function was used in the simulation.

It is interesting to note that for the grating cases where there are two SP dips, whereas there is only one plasmonic dip on the back focal plane as shown and discussed in the previous section.

For the *x polarisation*, we can see from figures 4.6a and 4.6b that there are two SP dips; the main dip was at lower angle on the back focal plane and the other small dip appeared at a slightly bigger angle on the back focal plane.

For the *y polarisation*, the dips had opposite trend to the *x polarisation*, on the other hand, the main dip was outside and the small dip was inside. Moreover, for the *r polarisation*, the effects of both *x* and *y polarisations* appear on the *r polarisation* back focal plane as shown in figure 4.6f. The main dip for the *x polarisation* was at 49.41 degrees and 51.32 degrees for the *y polarisation*.



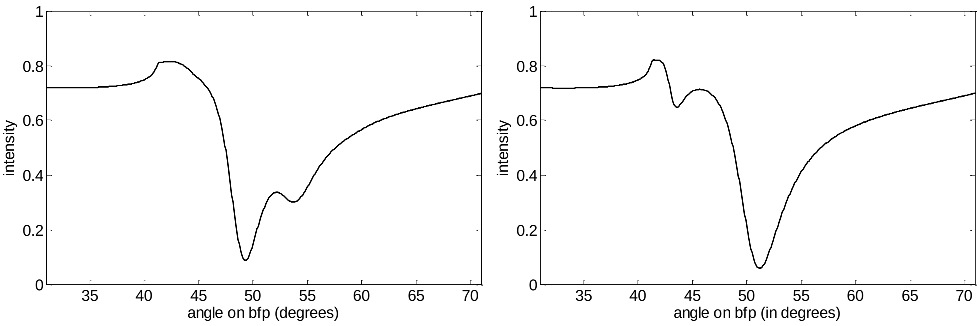
a) b)

c) d)

e) f)

Figure 4.6 shows plasmonic dips : all the experimental results were taken from (Stabler *et al* 2004)
a) Experimental result for wide field back focal plane image *x polarisation*
b) Simulation result for wide field back focal plane image *x polarisation*
c) Experimental result for wide field back focal plane image *y polarisation*
d) Simulation result for wide field back focal plane image *y polarisation*
e) Experimental result for wide field back focal plane image *cross polar x-y polarisation*
f) Simulation result for wide field back focal plane image *r polarisation*

Secondly the positions of the main SP dip for the *x* and *y polarisations* were different and we can clearly see this effect in figure 4.6f, where the main dip of *x polarisation* (along the x axis of figure 4.6f) was at a lower angle on the back focal plane compared with the *y polarisation* (along the y axis of figure 4.6f). Therefore it is interesting to understand what the grating does to the back focal plane. The positions of the main SP dip for *x* and *y polarisations* are clearly illustrated in figure 4.7.

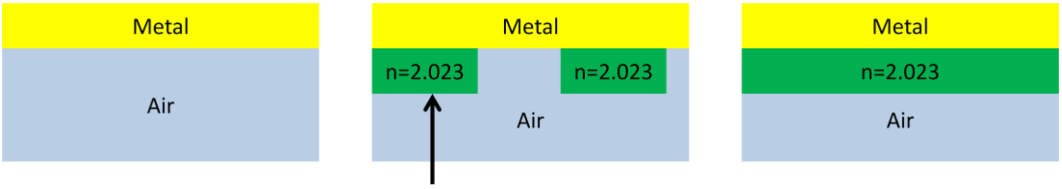


a) b)

Figure 4.7 shows plasmonic dips a) for *x polarisation* and b) for *y polarisation*

The two SP dips, of course, come from the effect of grating, which will diffract the incoming light and scatter the diffracted light beams back onto the back focal plane. Number of diffracted orders, of course, depends on the grating period and the magnitude of the diffracted orders depends on the shape and the height of grating.

Therefore let us investigate the effect of the grating on the back focal plane by firstly varying the silicon nitride width from 0 microns to 8 microns over a fixed grating period of 8 microns, which means we investigate the effect of grating by trying to find the relationship of the grating effect by comparing how the plasmonic angle changes with the bare gold and uniform sample cases as shown in figure 4.8.



Vary the period of grating from 0 microns to 8 microns over a fixed period of 8 microns

Figure 4.8 shows simulation method to study the effect of grating on the back focal plane

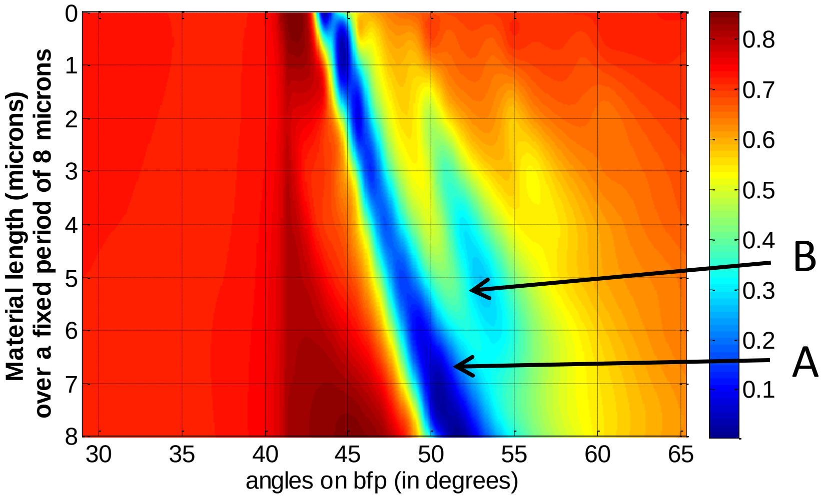


Figure 4.9 shows intensity as a function of bfp plasmonic dips and silicon nitride width *x polarisation*. The *y* axis shows the length of the coated region in the 8 micron period.

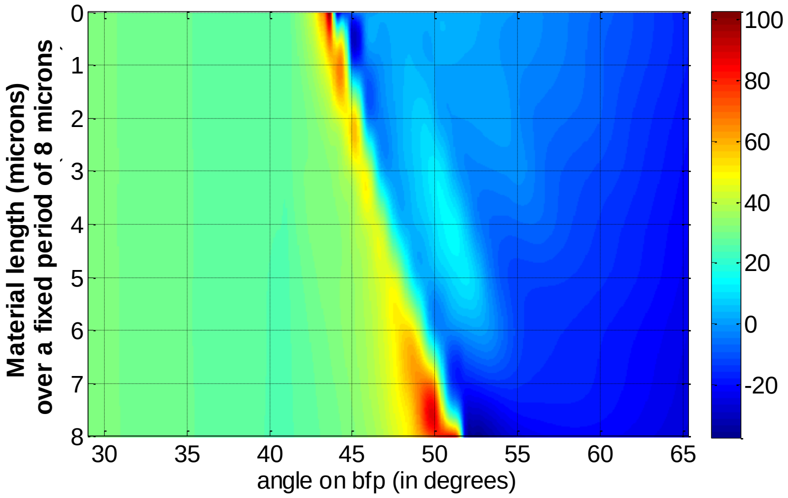


Figure 4.10 shows phase as a function of bfp plasmonic dips and silicon nitride width *x polarisation*. The *y* axis shows the length of the coated region in the 8 micron period.

From figure 4.9, we can see that for the *x polarisation* from the point of view of the zeroth order what the grating does to the bfp is average the two uniform cases (bare gold and coated gold) with silicon nitride width as a shifting factor. The reflection coefficient curves for the grating case can be expressed as "weighted average" of 2 uniform sample reflection coefficient curves as explained later. In other words, *x polarisation* is an average effect between the 2 plasmonic modes. The small dip at outer angle will appear when there is a grating; the position of the small dip corresponds to the 1st diffracted order ($kp_{grating} + kg$).

Let us consider the k_p values for the coated gold, bare gold and grating case:

$kp_{bare} = 1.01k_0$, which corresponds to 44.13 degrees.

$kp_{coated} = 1.14k_0$, which corresponds to 52.24 degrees.

$kp_{grating} = 1.10$ for *x polarisation*, which corresponds to 49.41 degrees.

$$kg = \frac{2\pi}{grating\ period} = 0.08k_0$$

$kp_{grating} + kg = 1.18k_0$ which corresponds to 54.47 degrees (2nd dip position in figure 4.9 for 6 microns of silicon nitride case).

$kp_{grating} + 2kg = 1.26k_0$ which corresponds to 60.33 degrees.

Figure 4.10 clearly tells us that the main plasmonic dip is an SP dip, which has a rapid 2π phase change at the plasmonic dip position. On the other hand, the small dip from the 1st diffracted order does show a phase shift but the magnitude of the phase shift is smaller because of the reduced amplitude of the diffracted order. We can therefore conclude that the main dip is generated due to the SPR but the small dip is generated by the first order. Also if we look carefully at figure 4.9 we can see another dip with very small amplitude due to the 2nd diffracted order ($kp_{grating} + 2kg$) at 62 degrees when silicon nitride width was 1 micron.

It is interesting to see how close we can reconstruct the grating bfp from the uniform cases by calculating the two uniform plasmonic angle shift corresponding to the silicon nitride width as shown in figure 4.11. The algorithm to reconstruct the grating bfp can be mathematically expressed as:

$$W_{air} = grating\ period - W_{silicon\ nitride} \qquad [4.2]$$

$$\Delta\theta_p = \theta_{p,coated\ gold} - \theta_{p,bare\ gold} \qquad [4.3]$$

$$\theta_{p,shift\ bare} = \frac{W_{silicon\ nitride} \times \Delta\theta_p}{grating\ period} \qquad [4.4]$$

$$\theta_{p,shift\ coated} = \frac{w_{air} \times \Delta\theta_p}{grating\ period} \qquad [4.5]$$

Where w_{air} and $w_{silicon\ nitride}$ are the width of air in the grating layer and width of silicon nitride in the grating layer. $\Delta\theta_p$ is the difference in plasmonic angles; $\theta_{p,coated\ gold}$ and $\theta_{p,bare\ gold}$ are plasmonic angles for the coated gold and uniform gold cases. $\theta_{p,shift\ bare}$ and $\theta_{p,shift\ uniform}$ are the angle shifts required for bare and uniform gold to calculate the grating bfp.

Then shift the position of the uniform and bare gold cases by the amount of $\theta_{p,shift\ coated}$ and $\theta_{p,shift\ uniform}$ respectively without changing the shape of the bfps and average the sum of the two shifted bfps.

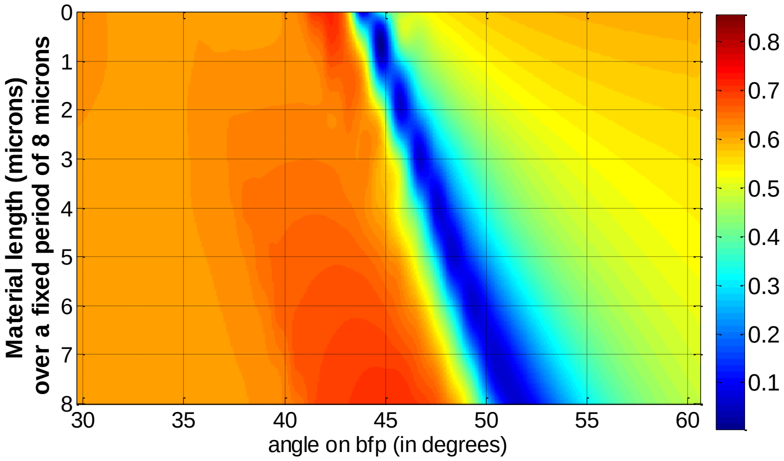


Figure 4.11 shows intensity as a function of bfp plasmonic dips and silicon nitride width *x polarisation* generated by weighted average of the bare gold and uniform sample cases.

From figure 4.11 we can see clearly that the averaging algorithm gives us a good approximation of the main dip position, however there is of course no small dip next to the main dip as the mathematical model does not take the other diffracted orders into account.

Figure 4.12 shows an interesting feature on the bfp where the bfp has two dips corresponding to the plasmonic dips of the uniform and bare gold cases without averaging and interfering each other. Figure 4.13 provides us with an even stronger evidence of the existence of the two plasmonic modes as the both dips have the rapid

2π phase transition around the SP dips. Therefore the SPs exhibit both SP modes at the same time; each of the modes propagates along its own material track.

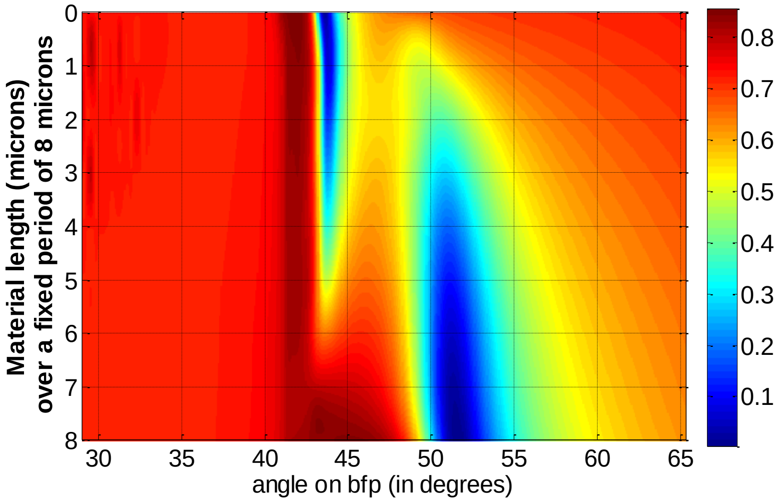


Figure 4.12 shows intensity as a function of bfp plasmonic dips and silicon nitride width *y polarisation*. The y axis shows the length of the coated region in the 8 micron period. The *y* axis shows the length of the coated region in the 8 micron period.

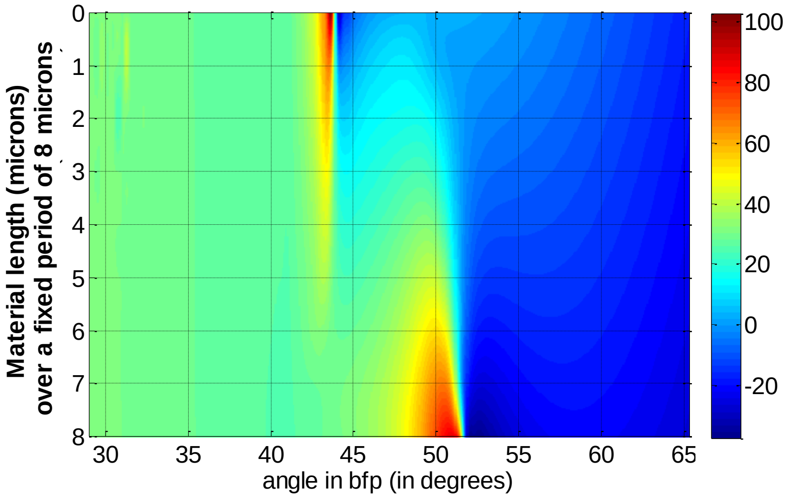


Figure 4.13 shows phase as a function of bfp plasmonic dips and silicon nitride width *y polarisation*. The *y* axis shows the length of the coated region in the 8 micron period.

A way to reconstruct the bfp of grating from the two bfp of uniform and bare gold cases can be done by 'weighted average' as shown in figure 4.14, which can be mathematically described as:

$$BFP_{grating} = \frac{(grating\ period - W_{silicon\ nitride})BFP_{silicon\ nitride} + (W_{silicon\ nitride})BFP_{bare\ gold}}{grating\ period} \qquad [4.7]$$

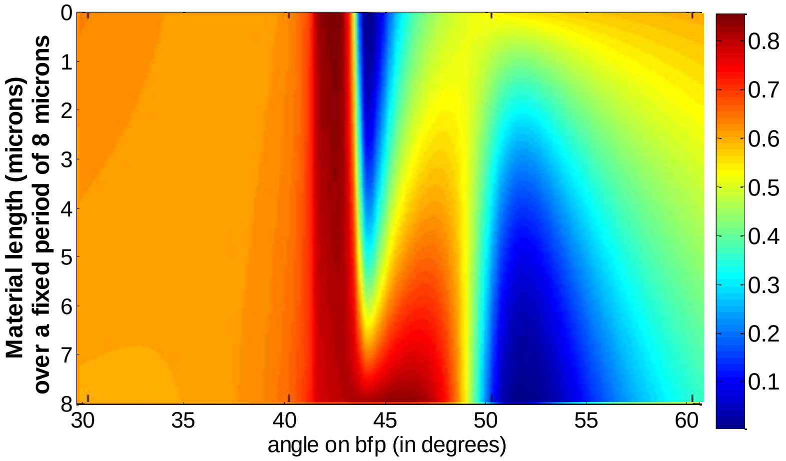
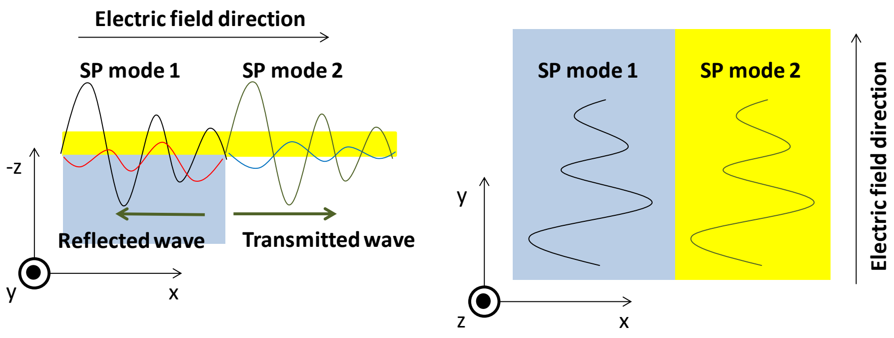


Figure 4.14 shows intensity as a function of bfp plasmonic dips and silicon nitride width *y polarisation* generated by weighted average of the bare gold and uniform sample cases. The y axis shows the length of the coated region in the 8 micron period.

In conclusion, for the *x polarisation* there are two dips on the bfp; the main dip is from the SPR and the second dip at outer angle is from the first diffracted order. The physical explanation of the main SP dip ('A' in figure 4.9) that it changes its position when the on-off ratio of grating changed; this is because the electric field is along the grating vector so that the two of SP modes propagate and interfere with each other as depicted in figure 4.15a. In figure 4.9 we see that the diffracted order ('B' in figure 4.9) detected occurs at angles greater than the angle to excite the SPs on a uniform sample arising from incidence at the plasmon angle and scattering by the grating vector. The thickness of the grating will also affect the amount of light in each diffracted order; the thicker grating will have more light intensity in the diffracted order than the thinner one. In other words, for the thicker grating thickness, it is more likely to see the second dip.

For the *y polarisation*, there are also two SP dips on the bfp caused by the two plasmonic modes of the on and off parts of the grating. Both dips correspond to surface plasmons in the appropriate region. The clear evidence of existence of the surface plasmons are, of course, their phase profile as shown in figure 4.13. The two dips are located at the position of the two uniform cases, which are bare gold and coated gold. What the grating does to the bfp plane is simply averages the two uniform back focal planes by the ratio of how much amount of that particular material is present in the grating. The two plasmonic dips will be easy to observe when the refractive indices of the two material regions have a high refractive index contrast because they are further apart. If the refractive index contrast ratio is low the two dips will be very close to each other and, we might not be able to distinguish them. The physical reason why there are two SPR modes for the *y polarisation* is that the two modes are generated on their own material path and propagate along the grating track as shown in figure 4.15b.

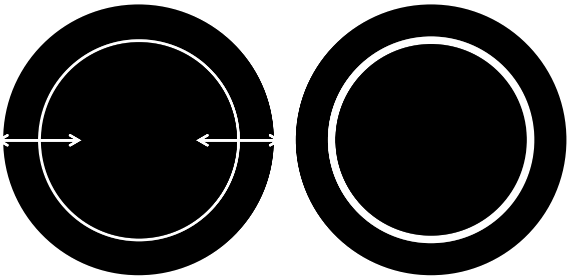


a) b)

Figure 4.15 shows surface plasmons propagation for a) *x polarisation* and b) *y polarisation*

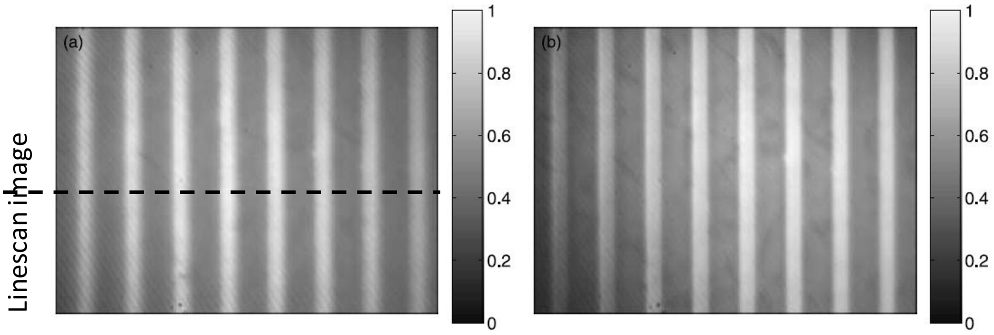
4.2.2 Imaging performance

As mentioned earlier in chapter 2, if we sum the intensity of the bfp to get the image; the image will show virtually no contrast as the contrast from the SPs will be swamped by the background. In other words, if we want to detect a small change in a large background signal it is necessary to block some spatial frequencies in the back focal plane that will not contribute to the image contrast. The way to get around this problem is to use an amplitude spatial light modulator (A-SLM) to generate a ring mask around the plasmonic angles as shown in figure 4.16. The imaging performance of the non-interferometric system is here controlled by three parameters which are firstly polarisation, secondly position of the ring mask and thirdly width of the mask.



a) b)

Figure 4.16 shows annulus ring excitation using A-SLM a) move the position of the ring and b) vary the size of the ring



a) b)

Figure 4.17 shows images and linescans of grating structure. a) *x polarisation* and b) *y polarisation*; the images were obtained using a mask which allows 45 degrees to 52 degrees of incidence (Stabler *et al* 2004)

The experimental results shown in figure 4.17 are images of the grating sample. They were obtained with a solid mask that allows a range of incidence from 45 degrees to 52 degrees (the centre of the mask was at 48.5 degrees with the mask width of 7 degrees). It is interesting to note that the experimental results have a wrong grating aspect ratio. The actual aspect ratio of the grating is 3:1, however the images here have 2:1 aspect ratio. The lateral resolution for the *y polarisation* was a lot better than the *x polarisation*, the lateral resolution was 1.3 and 0.93 microns for *x* and *y polarisations* respectively.

Although the experimental results were only for *x* and *y polarisations*, we will consider the radial polarisation as well. Figure 4.18 shows a series of simulated linescan images taken at different position of the mask position where the mask width was 5 degrees in width. The *y* and *r polarisations* show very clear features of the edge response, on the other hand the *x polarisation* had the worst edge response.

Note that the coordinate system is defined as shown in figure 4.15. Series of $I(x,y)$ taken at different mask positions where the mask width was 5 degrees are presented in figure 4.19. Figure 4.18 and 4.19 tell us some interesting stories; firstly resolution and contrast depend on the incident polarisation. Secondly the position of the mask does affect resolution, sensitivity and aspect ratio of the images. Moreover different mask positions give us opposite contrast reversal images such as shown at the 42 and 52 degrees in figures 4.18 and 4.19, where the air was dark and the silicon nitride was bright at 42 degrees become bright for the air and dark for the silicon nitride at 52 degrees. The contrast reversal due to the mask position has been experimentally confirmed by (Tan 2011).

The mechanism of reversal contrast for the non-interferometric system is different from the interferometric system. The interferometric microscope will be explained later. For the non-interferometric system, images with reversal contrast can be observed with scanning mask; this is due to the position of the mask ring on the back focal plane; for example if the sample is a grating with θ_{p1} and θ_{p2} corresponding to the two different material regions on the grating, if the mask is centred around the θ_{p1} the first material region will appear darker as we are at the dip in intensity and the second material region will appear brighter, whereas if the mask is located around the θ_{p2}, the first material region will now appear brighter and the other region will appear darker instead. However if the mask is at the centre of the two plasmons angles, image will have low contrast.

We can see from the figure 4.19 that for the cases where the centre of mask was around the plasmonic dip of the air period (2 microns), which was 44 degrees; the images seem to have the wrong aspect ratio. On the other hand, when the mask was around the plasmonic dip of the silicon nitride material (6 microns) which was 52 degrees; the images seem to have a better aspect ratio. We believe that the 2 microns region will suffer, of course, from crosstalk more than the 6 microns region.

The crosstalk here means that the SPs generated in one particular part of grating are affected by the adjacent regions as the surface waves propagate from one region to another. More quantitative details about crosstalk will be discussed in the next chapter. We believe that the appearance of the "wrong" aspect ratio arises from this crosstalk effect.

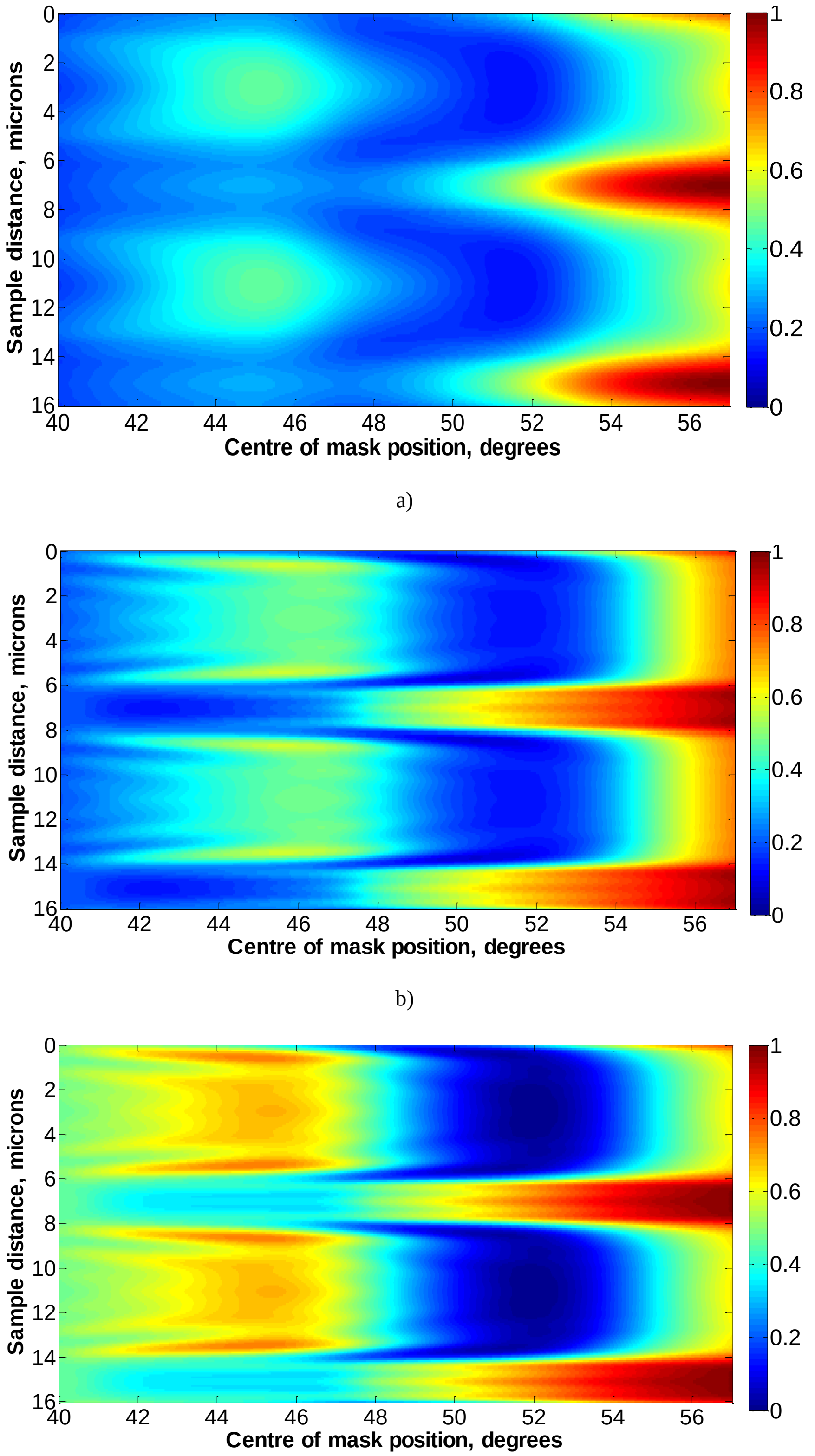
The simulations suggest that more faithful images are produced if the radius of the mask associated with the angle of incidence corresponds to the angle associated with the majority material system, in this case the coated region as shown in table 4.1.

		0.5 degrees mask width			5 degrees mask width		
Polarisation	Parameters	Centre of mask at 44 degrees	Centre of mask at 47.5 degrees	Centre of mask at 52 degrees	Centre of mask at 44 degrees	Centre of mask at 47.5 degrees	Centre of mask at 52 degrees
X polarisation	Resolution	2.37 µm	1.39 µm	2.77 µm	2.42 µm	1.45 µm	2.78 µm
	Aspect ratio	≈2:1	≈2:1	≈3:1	≈2:1	≈2:1	≈3:1
	Contrast ratio	0.04	0.07	0.13	0.04	0.05	0.08
Y polarisation	Resolution	1.06 µm	1.06 µm	0.82 µm	1.31 µm	0.86 µm	0.94 µm
	Aspect ratio	≈2:1	≈2:1	≈3:1	≈2:1	≈2:1	≈3:1
	Contrast ratio	0.09	0.27	0.03	0.12	0.13	0.04
R polarisation	Resolution	1.11 µm	1.15 µm	2.39 µm	1.34 µm	1.12 µm	2.07 µm
	Aspect ratio	≈2:1	≈2:1	≈3:1	≈2:1	≈2:1	≈3:1
	Contrast ratio	0.18	0.85	0.13	0.23	0.35	0.11

Table 4.1 shows resolution, aspect ratio and contrast ratio of different mask positions and mask widths for *x,y* and *r polarisations*

It is also interesting to see the effect of mask width, so let us look at the same set of linescan images taken at the same position however with the smaller mask width of

0.5 degrees as shown in figure 4.20. From the figure we can see that firstly the reversed contrast images can be observed however over a narrower range of mask positions compared to 5 degrees mask shown in figure 4.19. Secondly, *y polarisation* still had the best lateral resolution, followed by *r* and *x polarisations* respectively. The table 4.1 shows quantitative comparison for different mask widths and polarisations. From table 4.1, we can conclude that the contrast ratio depends on the mask width; a smaller mask width gave us a better contrast ratio; *r polarisation* had the best contrast ratio followed by *y* and *x polarisations* respectively. Secondly, resolution performance is not affected much by the mask width; however the mask position does affect the lateral resolution performance. Thirdly the aspect ratio; for the case where the centre of the mask was around 44 degrees, the contrast ratio was around 2:1, whereas the contrast ratio when the centre of the mask position was around 52 degrees gave us a correct aspect ratio image of 3:1. This is evidence of the crosstalk effect where the dominant material system will have less crosstalk effect, in this case the longer material in grating is silicon nitride 6 microns over the grating period of 8 microns. Therefore if we operate at the plasmonic angle around the plasmonic angle of silicon nitride we should get a correct aspect ratio. Our simulations demonstrate that if the experiments were carried out with 52 degrees mask position rather than 47.5 degrees; we would be able to get images with the correct aspect ratio.

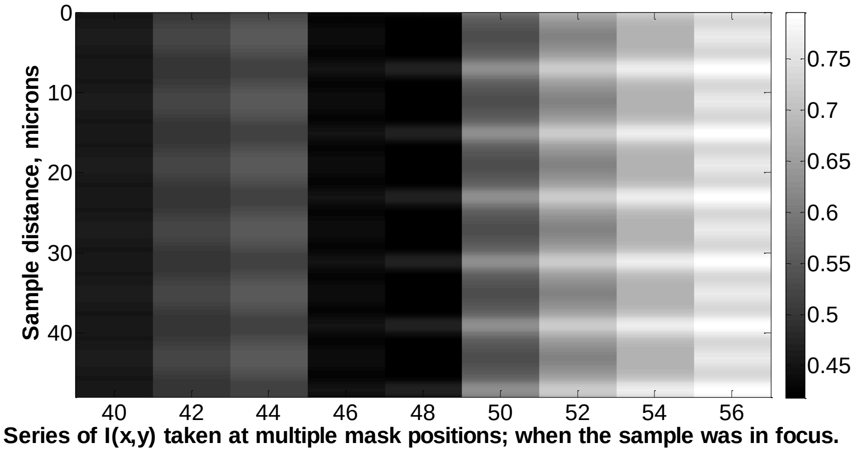


a)

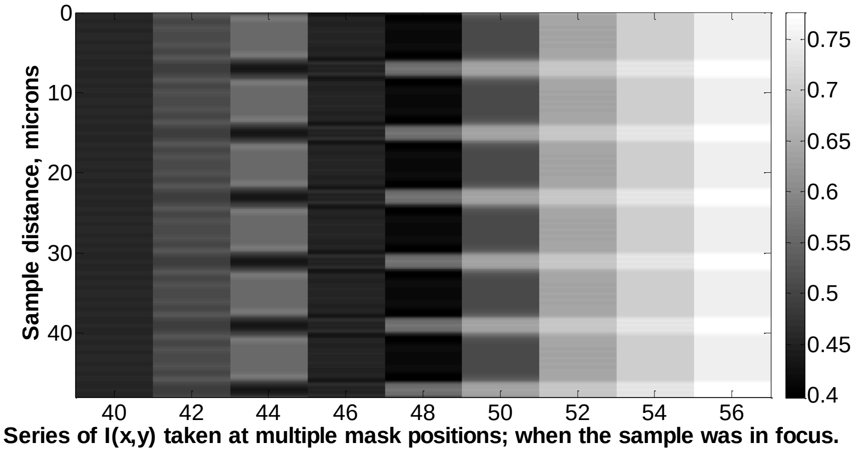
b)

c)

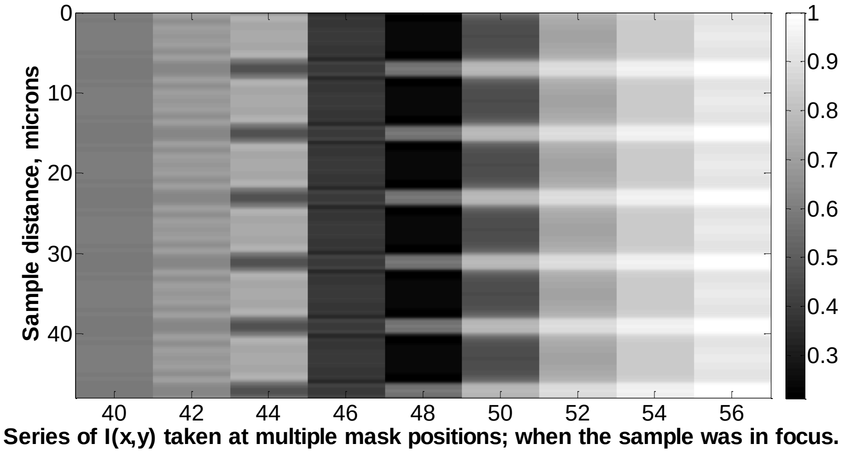
Figure 4.18 shows normalized linescan intensity as a function of mask central position when the mask width is 5 degrees for a) *x polarisation* b) *y polarisation* and c) *r polarisation*



a)

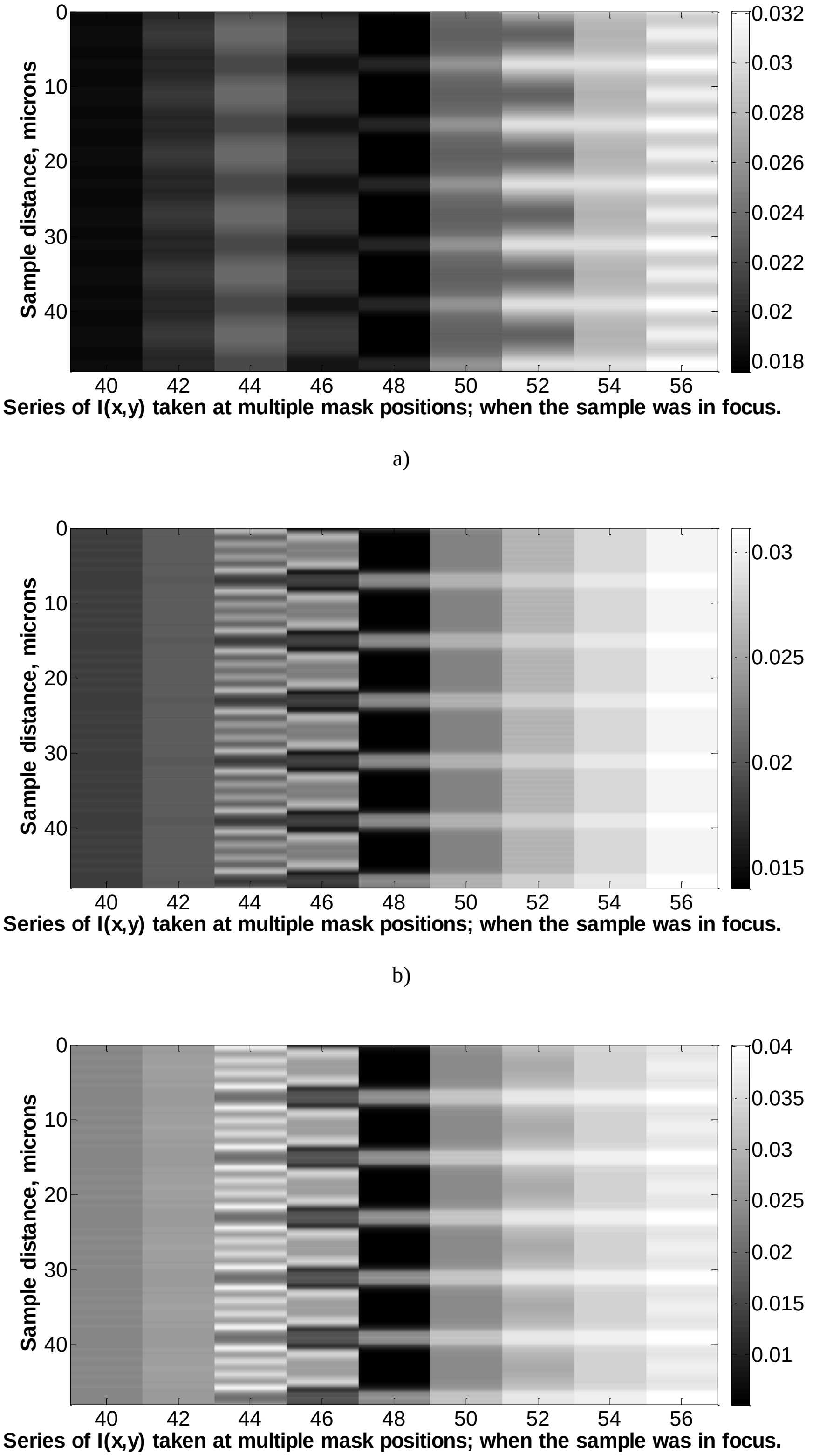


b)



c)

Figure 4.19 series of *I(x,y)* images taken at the centre of mask position as shown in x axis 5 degrees mask width for a) *x polarisation* b) *y polarisation* and c) *r polarisation*



a)

b)

c)

Figure 4.20 series of *I(x,y)* images taken at the centre of mask position as shown in x axis 0.5 degrees mask width for a) *x polarisation* b) *y polarisation* and c) *r polarisation*

4.2.3 Techniques to enhance performance of non-interferometric system

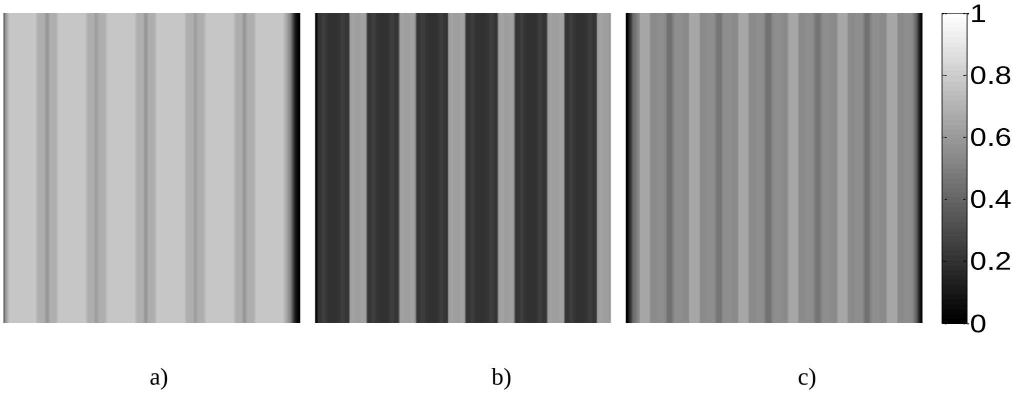
In this section only the radial polarisation results will be shown as it is the best option for an unknown sample. The following techniques have been investigated; firstly the use of shorter propagation length, secondly defocusing the sample and thirdly use of a high NA objective with high refractive index couplant. Using higher refractive index couplant did not improve the performance of non-interferometric microscope, on the other hand, it will be shown later in the interferometric section that the high index couplant can be used to enhance the performance of the interferometric microscope.

Does the sample have to be in focus?

The reason we need to ask this question is because we show in later sections that improved contrast is achieved with the interferometric system when the sample is defocused.

Let us look at 3 cases where we have the sample focused at 3 different z defocus positions, which are $z=-2$ microns (moving the sample 2 microns away from the focal point towards the objective lens), 0 micron (in focus) and 2 microns (moving the sample 2 microns away from the focal point outwards the objective lens). The linescan images for silicon nitride grating sitting on gold 50 nm are shown in figure 4.21. The images were taken with the mask width was 5 degrees cantered at 47.5 degrees mask position with *r polarisation*.

From the figure we can see that for the non-interferometric system the sample is required to be in focus in order to get the best resolution and contrast; as we can see from the figure that if the sample were out of focus the images would be blurry and the contrast ratio was also poor.



a) b) c)

Figure 4.21 shows linescan images of a) *z=-2* microns b) *z=0* micron and c) *z=2* microns when the mask

Shorter propagation length

In order to illustrate the effect of shorter propagation length 50 nm layer of uniform gold was replaced with 30 nm as discussed earlier that the thinner layer of gold will have shorter propagation length. The figure 4.22a shows the linescan of the silicon nitride grating with gold thickness of 30 nm *r polarisation* when the mask width was 5 degrees. We can see from the figure 4.22a that the resolution of 30 nm case was better than the 50 nm case shown in figure 4.22b, however, the contrast ratio was worse. The calculated values of 10%-90% resolution were 0.83 microns and 1.12 microns for 30 nm and 50 nm of gold respectively. In conclusion, one can improve the lateral resolution of the image, but one has to pay the price of reduction in the contrast of the image.

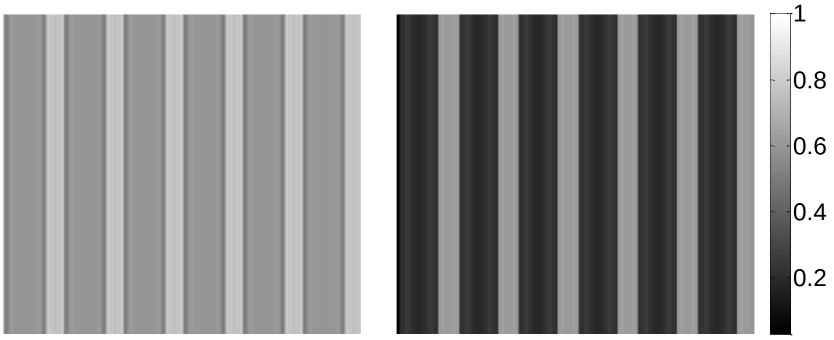


Figure 4.22 shows linescan images of a) 30 nm of gold and b) 50 nm of gold when the mask width was 5 degrees taken at 47.5 degrees centre of mask position *r polarisation*

It is worth pointing out that there is also another way to shorten the propagation length; this can be done by using an aluminium layer of 15 nm instead of the gold layer. It has been experimentally confirmed by (Giebel *et al* 1999) that the resolution can be improved however with an expense of contrast of the image.

To sum up, so far I have shown that the non-interferometric system has a characteristic trade-off between lateral resolution and sensitivity. For the next section, I will show that the *V(z)* interferometric microscope setup can overcome or at least obtain a good compromise the trade-off between resolution and sensitivity.

4.3 Interferometric SPR setup

Let us consider the scanning SPR heterodyne interferometric microscope as in the previous chapter and shown in figure 2.29. We have seen that the interferometric system operates by interfering a reference beam reflected from a mirror with the signal beam incident on the sample. The two beams will then interfere and the interference signal is then detected using lock in amplifier. The interference beam is the so called *V(z)* measurement. The *V(z)* response can be depicted as a phasor diagram shown in figure 4.23.

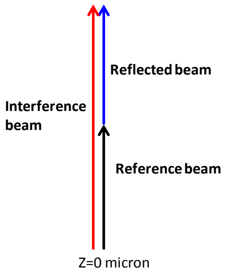


Figure 4.23 shows vector diagram for different z defocuses for the non plasmonic case; interference signal in red, reflected signal in blue, reference signal in black z=0 microns

From figure 4.23, when the sample is in focus, all the spatial frequencies on the back focal plane will be in phase with the reference beam, hence it gives the highest amplitude of the *V(z)* response. On the other hand, if the sample is defocused the DC spatial component on the back focal plane will have the phase difference of $2k_{vec}z$ with respect to the reference signal, whereas the other spatial frequencies will have the phase difference of $2k_{vec}z\cos\theta$.

For the non-plasmonic case, we expect this interference beam to provide a steady reference to interfere with the SPs. The steady reference can be provided by using a smooth pupil function, such as the one shown in figure 4.24. The *V(z)* response calculated using the pupil function shown in figure 4.24 for non-plasmonic case is shown in figure 4.25a. We can see that the pupil function can provide us with a good steady reference. Note that the non-plasmonic case was provided by illuminating the light on a thick gold, such as 2 microns thick, where the SP excitation is suppressed.

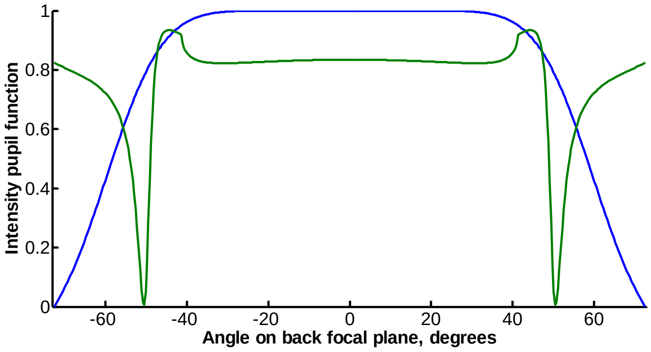
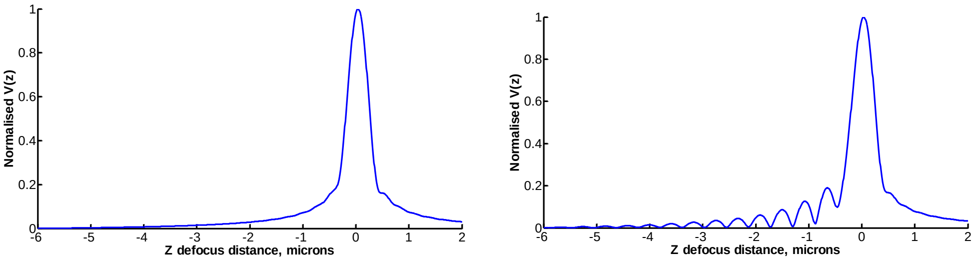


Figure 4.24 shows pupil function used in this study; intensity of pupil function in blue and the back focal plane distribution for $n=2.023$, 20 nm thick in green; $NA=1.45$ and $n_0=1.52$, *633 nm* wavelength *p-polarisation*

If there are SPs propagating along the surface of the metal, these will, of course, alter the phase of the reflected wave and appear as ripples on the *V(z)* curve as shown in figure 4.25b, where the calculation was carried out with 46 nm of gold; 10 nm of layer of $n=1.5$ sitting on gold in water ambient with $NA=1.65$ and $n_0=1.78$ 633 nm wavelength and linear polarised wave.



a) b)

Figure 4.25 shows *V(z)* curve for a) non-plasmonic case 2 microns thick of gold and b) plasmonic case 46 nm of gold; 10 nm of layer of $n=1.5$ sitting on gold in water ambient with $NA=1.65$ and $n_0=1.78$ 633 nm wavelength and linear polarised wave.

More complete details for the *V(z)* calculation and effects of pupil function are provided in the next chapter where the accuracy of the *V(z)* measurement and effects of pupil function have been addressed. However, the aim of this chapter is to provide a basic idea of how the *V(z)* microscope operates when imaging samples with refractive index variation and shows how we use this to enhance the resolution when it is compared with the non-interferometric microscope.

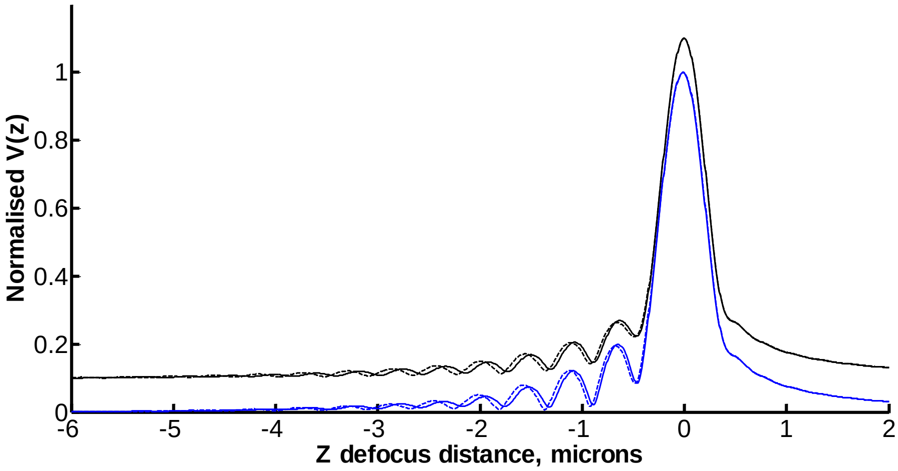


Figure 4.26 shows *V(z)* curves for 10 nm thick sample with refractive index $n=1.33$ (dashed line) and $n=1.5$ (solid line) in water ambient sitting on *36 nm* of gold (in blue) and *46 nm* of gold (in black); the black lines have been displaced by *0.1* for visualisation. $NA=1.65$, $n_0=1.78$, *633 nm* wavelength and linear polarised wave.

From figure 4.26, the results illustrate two key features of the *V(z)* measurement; firstly the period of the ripples is sensitive to the refractive index change and secondly propagation length of the SPs does not have much effect on the period of the ripples as we can see that the blue and the black lines are synchronised with each other. One of the main advantages of using a thinner layer of gold is that higher amplitude of ripples can be obtained, in other words, signal to noise ratio (SNR) of the ripples for the thinner layer case is higher than the thick gold case. The change in the ripple period due to the refractive index change was experimentally confirmed by (Somekh *et al* 2000a; b). The advantage of using thinner layer of metal to enhance the ripples amplitude experimentally was confirmed by (Roland *et al* 2010).

It can be shown that the ripple period Δz is related to the plasmonic angle θp by the equation (Somekh *et al* 2000a; b):

$\Delta z = \frac{\lambda_{free\ space}}{2n_0(1-cos\theta_p)}$ [4.6] (note that this equation will be derived in the next chapter)

Where Δz is ripple period. If Δz is known we can then work out the value of the plasmonic angle. Equation [4.6] was derived with certain assumptions. The accuracy of the *V(z)* measurement is explained in details in the next chapter. The ripple period Δz calculated by using Fourier transform to determine the dominant frequency.

We can see from the figure 4.26 that the ripple period is sensitive to the change in the refractive index, so this gives us a contrast mechanism for imaging. When we scan the focal point along the grating sample the *V(z)* curves will then change its period corresponding to the refractive index under the axis of the beam.

4.3.1 What does the grating do to the V(z) curve?

Let us firstly understand what the effects of the grating on the *V(z)* curves. We have seen that for the non-interferometric system a grating generates two dips on the back focal plane. Figure 4.27 shows that the *V(z)* curves for the silicon nitride grating where the beam is at the centre of silicon nitride (dashed line in blue) and centre of the gap (dashed line in red) compared with the uniform layer of silicon nitride 20 nm thick (solid line blue) and bare gold 50 nm (solid line red).

Figures 4.27 to 4.28 show that the grating sample displaces the *V(z)* ripple position. The amount of the displacement is different for each polarisation.

This leads to another way to investigate the crosstalk effect; crosstalk is an effect that distorts the accuracy of SPs measurement since adjacent regions affect the measurement of the region we are trying to access. A good example of the crosstalk is that when we want to measure the SPs dip position θ_p of a material, which is surrounded by other materials, such as microfluidics or other binding sites. In theory, without the effect of crosstalk we should be able to get exactly the same plasmonic angle as measuring on a uniform material. If that is the case it means the measurement is perfect, however as explained for the grating case the θ_p will change from its uniform value. Even if equation 4.6 is perfectly accurate on a uniform sample the effect of crosstalk means we do not recover the correct value of the θ_p that we would get on a uniform sample. Therefore measurement with the results closer to the uniform material case indicates low level of crosstalk, on the other hand, if the measurement is so different from the uniform value case it means it suffers severely from the crosstalk effect.

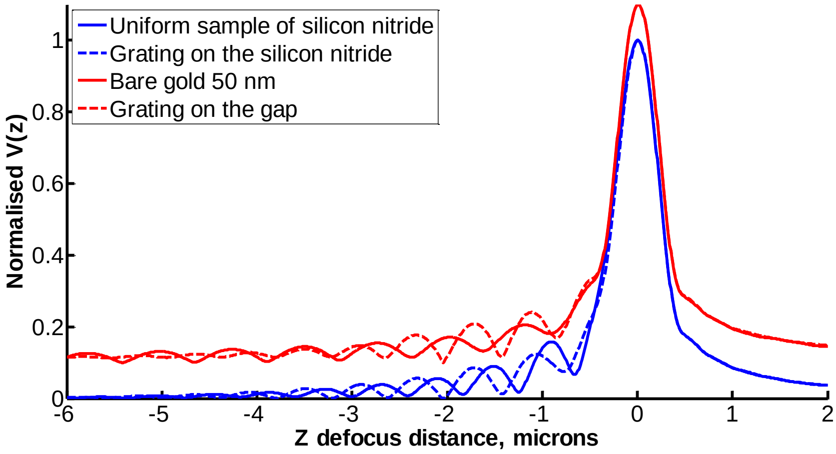


Figure 4.27 shows *V(z)* curves for the silicon nitride grating where the beam is at the centre of silicon nitride (dashed line in blue) and centre of the gap (dashed line in red) compared with the uniform layer of silicon nitride 20 nm thick (solid line blue) and bare gold 50 nm (solid line red) *x polarisation*.

The best resolution can be obtained with the *y polarisation* where the grating *V(z)* ripples were closer to the uniform case curves as shown in figure 4.28. Also it shows one important feature that for the longer region in the grating the *V(z)* result is close to uniform value as can be seen on the figure 4.28 that *V(z)* on the centre of silicon nitride were closer to the uniform silicon nitride case compared to the *V(z)* curves of the gap of grating and the bare cases. The details to quantify this crosstalk effect will be discussed in the next chapter.

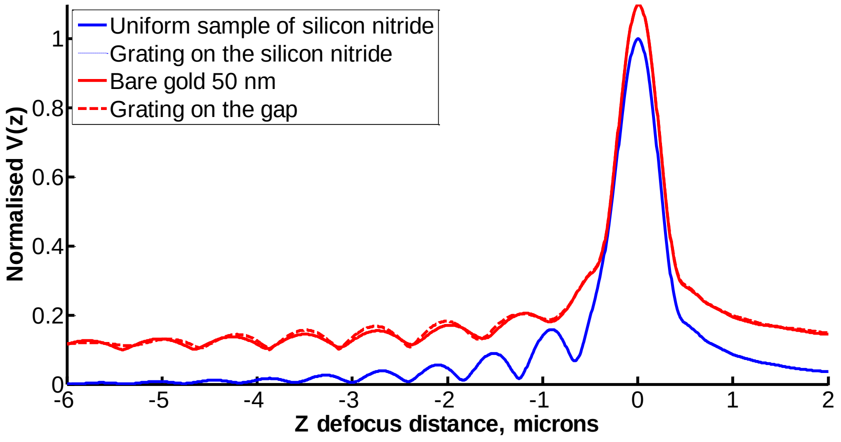


Figure 4.28 shows *V(z)* curves for the silicon nitride grating where the beam is at the centre of silicon nitride (dashed line in blue) and centre of the gap (dashed line in red) compared with the uniform layer of silicon nitride 20 nm thick (solid line blue) and bare gold 50 nm (solid line red) *y polarisation*. Remark that the two blue lines were very similar; the two lines were overlaid.

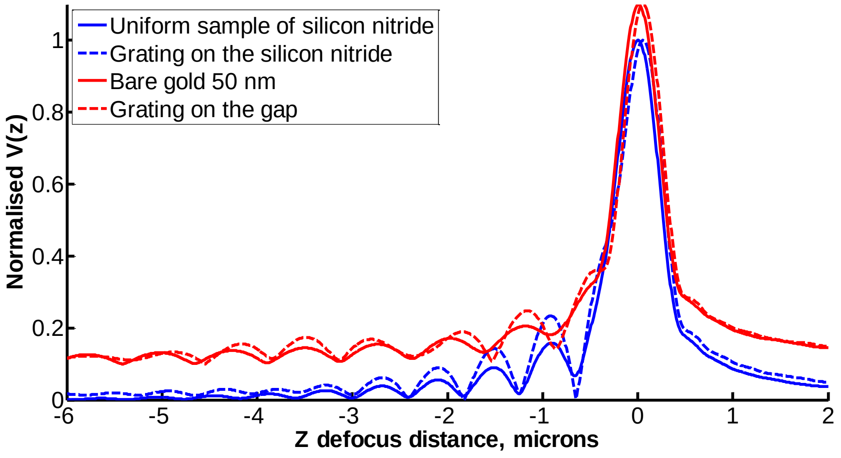


Figure 4.29 shows *V(z)* curves for the silicon nitride grating where the beam is at the centre of silicon nitride (dashed line in blue) and centre of the gap (dashed line in red) compared with the uniform layer of silicon nitride 20 nm thick (solid line blue) and bare gold 50 nm (solid line red) *r polarisation*.

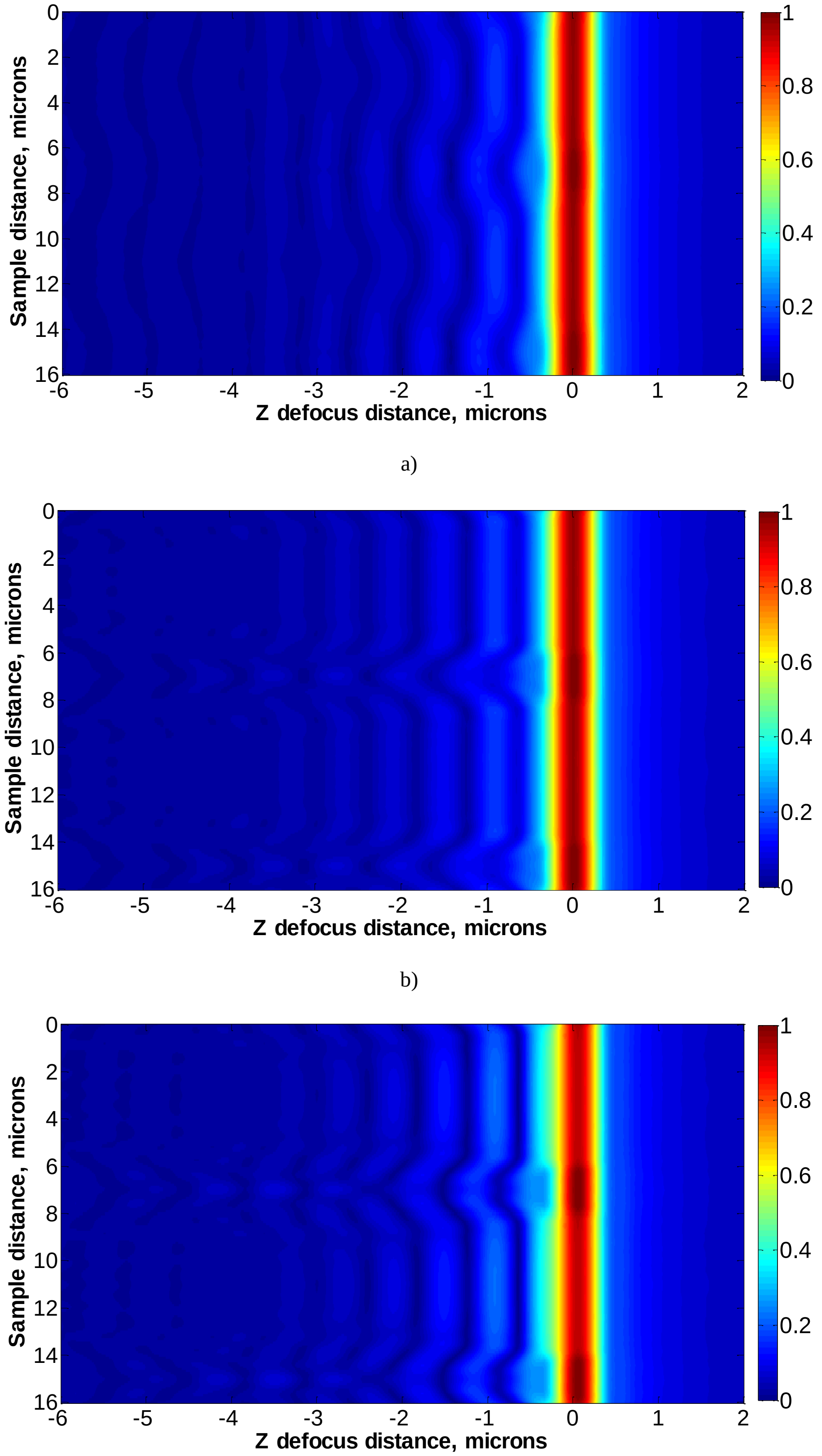
From figure 4.29, we can see that the magnitude of the ripples for radial polarisation are higher than the magnitude of the ripples for the *x* and *y polarisations*. This is due to the fact that the radial polarisation is *p-polarised* wave for all azimuthal angles; this will, of course, lead to the stronger excitation of SPs.

4.3.2 Imaging performance

We can see from the results in figure 4.29 that *V(z)* measurement can be used as an imaging tool to obtain good contrast images of the z defocus range; we can also see that this mechanism not only provides good contrast over the z defocus range, but also good resolution over the range. We can see directly from the *V(x,z)* that firstly *y* *polarisation* gives the best resolution performance and followed by *r* and *x polarisations* respectively. For the contrast of the image *r polarisation* has the highest contrast followed by *y* and *x polarisations*. The *r polarisation* had the highest ripple amplitude compared to the *x* and *y polarisations* as the *r polarisation* has the *p-polarisation* for all the azimuthal angles ϕ on the back focal plane, in other words, it had stronger SP excitation.

Note that over the *z* defocus distance there are many *z* points that give us no contrast at all such as at -1 micron -1.65 microns and also there are many points that give us a very good contrast. As we can see here that the non-interferometric and interferometric microscopes have different imaging mechanisms, therefore we still need a proper method to quantify their performances and compare them. The method to compare them will be discussed in the next chapter.

The purpose of this chapter is to give some general idea about imaging mechanisms. Let us look at some *V(x,y)* images taken around the first *V(z)* ripple in figure 4.30. From the figure 4.30 we can summarise the performance of some z defocus operating points as shown in the table 4.2 below. The definitions of all the terms shown in the table were defined in section 4.1.



a)

b)

c)

Figure 4.30 shows *V(x,z)* linescan images (*x axis*) over the *z* defocus range

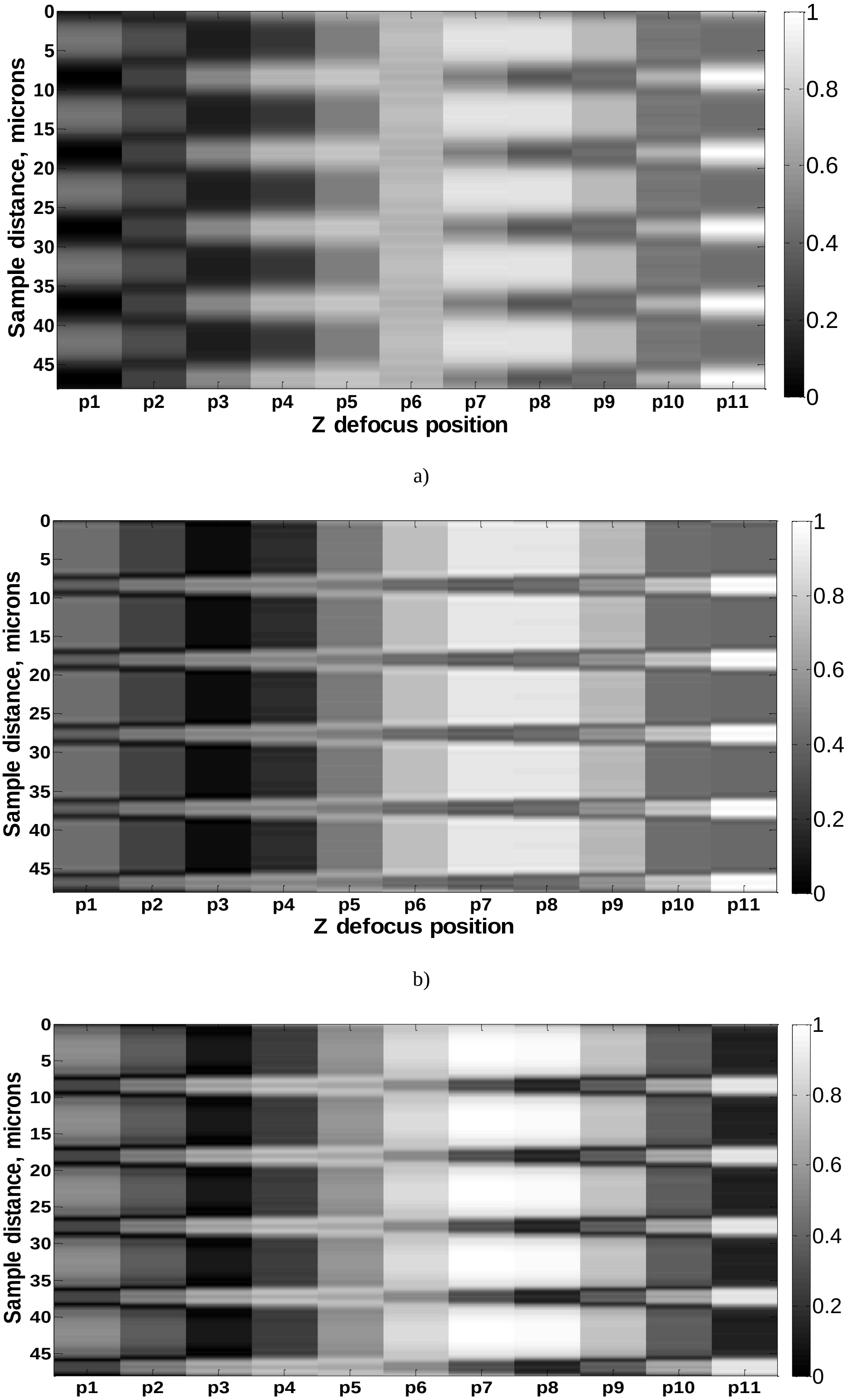
a) *x polarisation* b) *y polarisation* and c) *r polarisation*

Polarisation	Parameters	Z=-0.624 μm (p11)	Z=-1.024 μm (p6)	Z=-1.424 μm (p1)
X polarisation	Resolution	1.5 microns	1.8 microns	2 microns
	Aspect ratio	1.3:1	2.1:1	1.2:1
	Contrast ratio	0.40	0.01	0.98
Y polarisation	Resolution	0.5 microns	0.6 microns	0.8 microns
	Aspect ratio	3.0:1	3.2:1	3.5:1
	Contrast ratio	0.41	0.30	0.05
R polarisation	Resolution	1.0 microns	1.3 microns	1.5 microns
	Aspect ratio	3.0:1	4:1	3.3:1
	Contrast ratio	0.72	0.22	0.36

Table 4.2 shows resolution, aspect ratio and contrast ratio of different mask positions and mask size for *x, y* and *r polarisations*

We can see from the results in table 4.2 that *V(z)* measurement can be used as an imaging tool for a good contrast image of the z defocus range; we can also see that this mechanism does not only provide good contrast over the z defocus range, but also good resolution over the range. We can see directly from the *V(x,z)* that firstly *y polarisation* gives the best resolution performance followed by *r* and *x polarisations* respectively. The table also shows us that the lateral resolution of the *V(x,z)* images is quite stable over the *z* defocus distance. It has also confirmed two important points, firstly, the lateral resolution obtained by the interferometric system is indeed very good and better than the non-interferometric and secondly the contrast ratio is generally better than the non-interferometric microscope. The best lateral resolution for the SPR interferometric image obtained in experiment was around 0.5 microns (Somekh *et al* 2000b).

It is interesting to note that the lateral resolution becomes poorer when z defocus is increased. This is similar to the apparent aspect ratio of the image; the aspect ratio is worse with a bigger negative z defocus as shown in the table 4.2.



a)

b)

c)

Figure 4.31 shows series of *V(x,y)* images at different *z* points a) for *x polarisation* b) for *y polarisation* and c) for *r polarisation*; *p1=-1.424 microns, p2=-1.344 microns, p3=-1.264 microns, p4=-1.184 microns, p5=-1.104 microns, p6=-1.024 microns, p7=-0.944 microns, p8=-0.864 microns, p9=-0.784 microns, p10=-0.704 microns and p11=-0.624 microns*

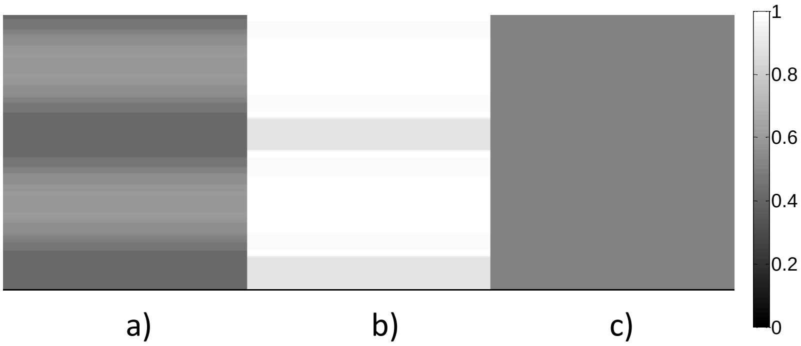


Figure 4.32 shows linescan images for *x polarisation* at a) $z=-1$ micron, b) $z=0$ micron (in focus) and c) $z=1$ micron

From figure 4.32, we can obtain a good contrast image only when the sample is negatively defocused (moving the sample towards the objective), whereas at $z=0$ micron and positively defocused the images have virtually no contrast. This is because at the focal point the amount of the reflected beams from the two regions of the grating had the same amount of interference signal. For the positive defocus, the SPs cannot be detected by the optical system; the details of this will be discussed in the next chapter.

We can see that the aspect ratio for the *x polarisation* was not good over the whole z defocus range. This is due to the SPs reradiated back at a different angle from what we expected passing to the objective lens.

The reason for the wrong aspect ratio image is that the SPs from one region propagate to the other region and reradiate back with different angle as shown in figure 4.33; this will lead to the wrong $V(z)$ ripple period and hence the wrong aspect ratio of the image.

From figure 4.33, we can see that when we scan over the silicon nitride region (blue in the figure), we expect to get the SP reradiation from only k_{p1}, however the k_{p2} mode will also go through the objective leading to the discrepancies in $V(z)$ response.

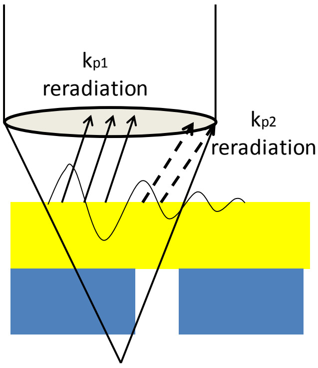


Figure 4.33 shows the SPs propagate to the other region and reradiate at different angle

The aspect ratio for the *y* and *r polarisations* were correct for the *z=0* upto *z=-1* microns. This aspect ratio error was not caused by the multiple reflections but by the footprint on the sample. When the sample is scanned towards the objective the size of the defocused beam on the sample is increased, the size of the focus beam might cover too much of the region than what we want to as shown in figure 4.34. Coverage region can be calculated as:

$$Coverage\ length = 2z\tan\ \theta_p \ [4.7]$$

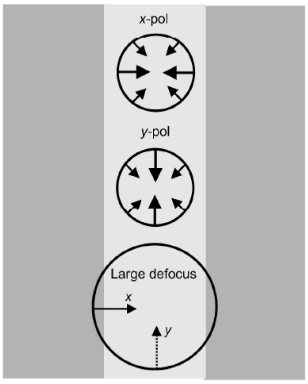


Figure 4.34 shows footprint on the sample at different z defocuses (Pechprasarn & Somekh 2012)

Let us calculate it to get the feeling of how the z defocus focused spot covers the sample, for example, the smaller region for the grating case is two microns and the plasmonic angle for the bare gold case is 44 degrees. Therefore the maximum z defocus that we can scan to is about 1.03 microns. These agreed well with the results presented in the figure 4.31 and the table 4.2; when we scanned over the -1 microns range the aspect ratio seem to be worse.

Therefore we can conclude that the interferometric system has better resolution and comparable sensitivity to the non-interferometric over the z defocus distance. This means the interferometric system can overcome or at least optimize the trade-off between resolution and sensitivity for SPR microscopy.

4.3.3 Techniques to enhance performance of interferometric system

Let us investigate some optical techniques to improve the performance of the interferometric microscope. In this section, we will consider only radial polarisation as it is a good representation between the two extreme cases. The following techniques have been investigated; the use of shorter propagation length and the use of a high NA objective with high refractive index couplant. The technique that can improve the performance is to employ a higher refractive index couplant for the interferometric microscope. This means that the propagation length is not a main consideration, whereas optical configuration is a more important key to improve the performance of the interferometric system.

Higher NA objective with higher refractive index couplant

In this section the $NA=1.49$ with $n_0=1.52$ will be replaced with $NA=1.65$ with $n_0=1.78$. As mentioned earlier, the higher refractive index will move the plasmonic angle inside the back focal plane from the range of 44-52 degrees to 36.1-41.3 degrees. It is interesting to note that when the plasmonic angles were reduced by using higher index couplant, $\Delta\theta_p$ was also reduced from 8 degrees to 5.2 degrees. The lateral resolution was similar to the figure 4.31c; however, the aspect ratio performance is better. This can be explained as the fact the SP angle for the $NA=1.65$ case was lower than the $NA=1.45$ case; this make the footprint on the sample smaller; therefore it will have a better performance and seem to give us a correct aspect ratio over the z points shown. We can calculate the footprint size using equation 4.7 and find that the range that we can defocus to is $z= 1.37$ microns.

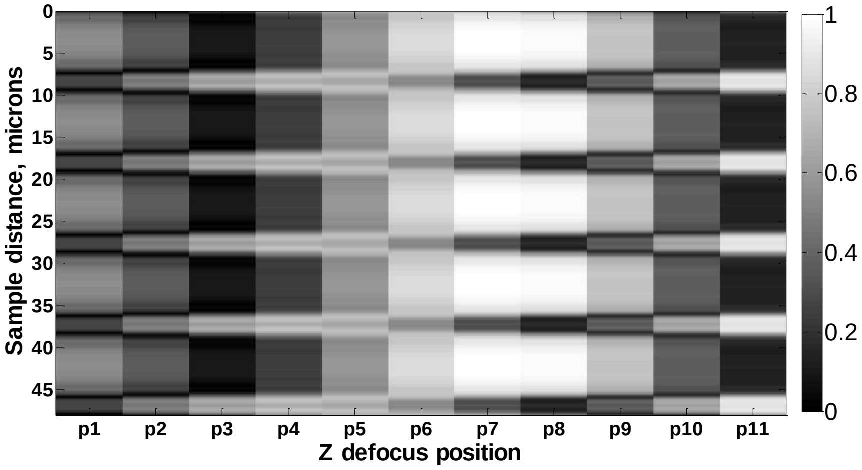


Figure 4.35 shows series of $V(x,y)$ images at different z points for *r polarisation*; $NA=1.65$ with $n_0=1.78$ *p1=-1.424 microns, p2=-1.344 microns, p3=-1.264 microns, p4=-1.184 microns, p5=-1.104 microns, p6=-1.024 microns, p7=-0.944 microns, p8=-0.864 microns, p9=-0.784 microns, p10=-0.704 microns and p11=-0.624 microns*

4.4 Discussion and conclusion

For the non-interferometric microscope, the resolution and sensitivity appear to conflict. For high sensitivity SP non-interferometric sensor, it is highly recommended to use a long propagation length SPs. For imaging, it is recommended to employ a short propagation length of SPs as this will enhance the lateral resolution. In other words, resolution and sensitivity of non-interferometric system depends very much on the propagation length. The effect of using high coupling index does not significantly affect the resolution and sensitivity.

For the interferometric system, we have shown the interferometric setup has comparable sensitivity to the non-interferometric system but with better lateral resolution. The interferometric system can largely overcome the trade-off between the resolution and sensitivity, where we can perform high contrast measurement with a good lateral resolution over the z defocus distance. The interferometric system is independent of the effect of propagation length but, on the other hand, the optical configuration such as NA and coupling oil of the objective seems to have a more pronounced effect on the performance of the interferometric SPR microscope.

We have demonstrated that the interferometric system has a better imaging performance than the non-interferometric system; however, we still have a need to perform a full quantification to ensure that the findings are justified. Therefore in the next chapter I will provide a proper quantification to confirm and compare the resolution and sensitivity of the interferometric and non-interferometric systems.

Chapter 5

Quantitative analysis of SP microscope performance using RCWA

In the previous chapter, we have seen that the interferometric microscope has the capability of imaging with better resolution and contrast than the non-interferometric microscope. However, we still need to quantify this improvement in a measurement environment. Therefore in this chapter, I will present a way to quantify the microscope response and compare sensitivity and localization of the interferometric and non-interferometric setups. We also show that the interferometric system in the defocused condition defines the measured point of excitation and reradiation of the surface plasmons; which greatly improves localization. In other words, the interferometric can be used to measure SP in a confined region.

It will be shown later that the quantitative measurement not only provides us with a way to compare the performance of both types of microscopes, but also gives us an insight to understand behaviour of SPR imaging and crosstalk.

5.1 Definitions of key terms and their definitions

5.1.1 Measurement localization

Measurement localization here means the smallest size of local region that enables one to recover the refractive index within specified tolerances. Clearly, the values obtained depend on both the material system and the tolerance values set.

Although the term ‘measurement localization’ is not exactly the same as the lateral resolution discussed in the previous chapter, it will be shown later that they are related and quite similar in a sense as the measurement localization also depends on the edge response of the sample as well as the contrast of the image. In addition, the measurement localization takes into account the edge response (10%-90% edge response), the aspect ratio and contrast ratio as shown later; this means that the measurement clearly depends on the position that one wishes to recover the refractive index.

5.1.2 Crosstalk

Let us clarify the term crosstalk; crosstalk is an effect that distorts the accuracy of SP measurement by being affected by the presence of adjacent materials. A good example of the crosstalk is when we want to measure the SP dip position θ_p of a material, which is surrounded by other materials, such as micro fluidic channels or grating samples. Without the effect of crosstalk we should ideally be able to obtain exactly the same plasmonic angle as a measurement on the same uniform material. If that is the case it means the measurement is perfect, however, as explained in the previous chapter for the grating case the θ_p will change from its uniform value. This means that there is a presence of crosstalk. Therefore measurement with the results closer to the uniform material case indicates low level of crosstalk, on the other hand, if the measurement is very different from the uniform value case it means it suffers severely from the crosstalk effect.

5.2 Methodology

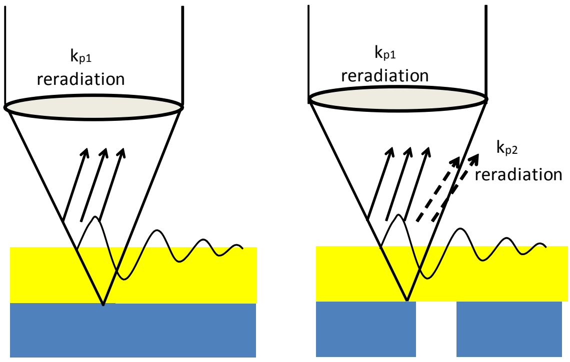
The idea to quantify the microscope response is to measure its plasmonic angle θ_p in a confined region and compare it to the uniform sample case to see how close the measured plasmonic angle is compared to the uniform case. The idea is based on the fact that SP in a material in a confined region will behave differently from the same material in a uniform material. This concept is in fact very important for designing SP sensors. In this chapter three types of samples will be used to demonstrate the effect of crosstalk as shown in figure 5.1. There are two types of the uniform sample cases which are bare gold case shown in figure 5.1a and the uniformly coated layer case shown in figure 5.1b. Grating samples with 50% on-off ratio and grating periods of 1 micron up to 30 microns were used in this simulation study in order to compare the performance of interferometric and non-interferometric microscopes as shown in figure 5.1c.

a) b) c)

Figure 5.1 shows samples used in this study a) bare gold case b) coated gold case and c) grating case

5.2.1Non-interferometric microscope

Let us consider the details of how the θ_p measurement was simulated for the non-interferometric microscope, an amplitude spatial light modulator was employed to vary the diameter of an annulus mask (this can be thought of as ring excitation); light reflected from the sample is then summed using a lens and detected by an intensity detector as a function of mask radius. This idea is similar to the idea illustrated in the previous chapter. The minimum intensity position is then interpolated to accurately determine θ_p.



a) b)

Figure 5.2 shows a) measurement without crosstalk b) measurement with crosstalk due to unwanted modes of SP from the adjacent regions

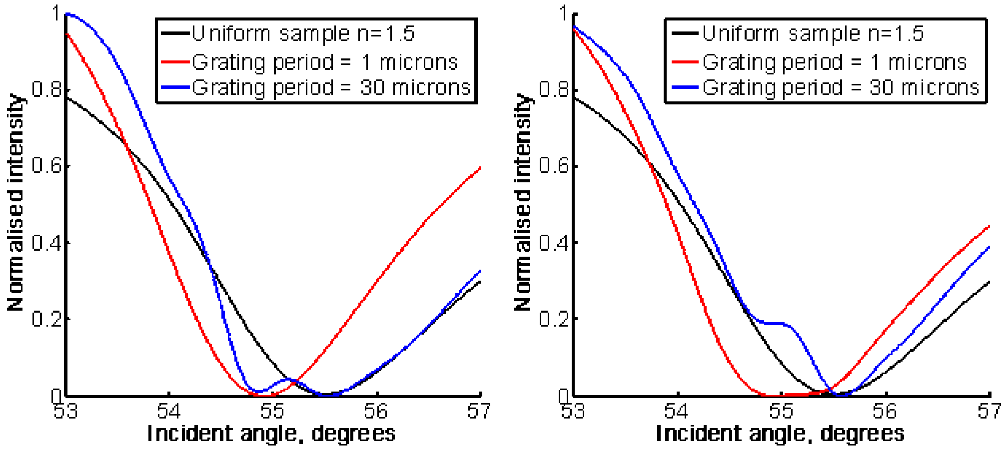
For the non-interferometric system, all the θ_p calculations were, of course, carried out when the sample was placed at the focal point as has been illustrated in the previous chapter that the non-interferometric images obtained at a defocused position had poor contrast and resolution. Figure 5.2a demonstrates the idea of non-interferometric measurement with no crosstalk, where the sample is in focus; the SP angle will not change when we move the focal spot along the *x* axis as the sample is a uniform sample. On the other hand, if we measure the same material on a confined region, such as grating or microfluidic channels; the measured SP such as plasmonic angle θ_p will be affected by reflected SP and the optical system may detect unwanted reradiation from the consecutive region as depicted in figure 5.2b.

Quantification of crosstalk in non-interferometric microscope

We can expect that shorter grating periods will, of course, alter the plasmonic angles θ_p from the value for a uniform layer more than the longer grating periods as the longer periods have their properties more similar to the uniform case. The effect of different grating periods can be illustrated as shown in figure 5.3 for the non-interferometric microscope, where a very small size of 0.05 degrees annulus mask were used in the intensity calculations. The intensity was calculated as a function of back focal plane angle as explained earlier.

From figure 5.3, we can see that the plasmonic angles for the larger grating cases were closer to the uniform sample case although the shape of the plasmonic dip was slightly different, for instance the dip in the case of the grating samples were wider. Figure 5.3 provides us with strong evidence of the existence of crosstalk effect on grating samples for the non-interferometric microscope.

All the results shown so far were measured at the centre of grating, one may question the best position to measure the SP signal on a grating sample. The criterion to choose where to measure the signal is, of course, the position that has the smallest crosstalk, in other words, the position that gives the closest SP parameter (in this case plasmonic angle θ_p) compared to the uniform case.

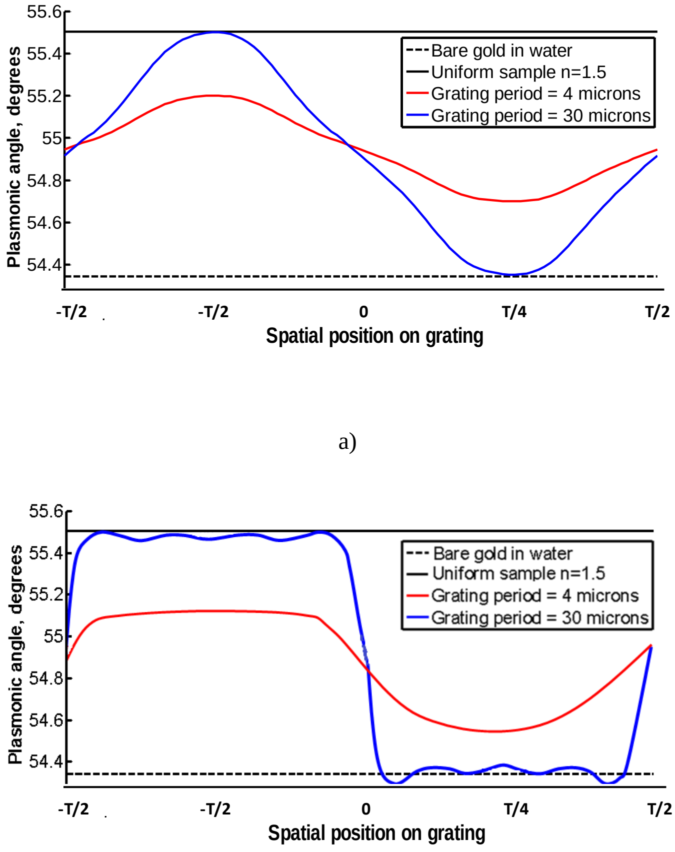


a) b)

Figure 5.3 shows plasmonic dips of non-interferometric microscope for *1 micron*, *30 microns* grating periods and uniform sample measured at the central point material of *n=1.5* with thickness of *10 nm*. n_0*=1.78* gold thickness of *46 nm* and *633 nm*. The sample was at the focal point and in water ambient. a) For *x-polarisation* and b) For *y-polarisation*

The results shown in figure 5.4a tells us, unexpectedly, that the best position to measure θ_p is at the centre of each material as they are closest to the uniform values. Figure 5.4b for 4 micron case also tells us with the same story; however, for the 30 micron case it is hard to determine the optimum position as the result suffered from the Gibbs effect; where we tried to represent a square grating with a Fourier harmonic numbers. In other words, for a long period of grating it is likely to suffer from the undersampling issue in frequency domain.

It is interesting to note that there are two types of Gibb's effect, firstly numerical Gibb's effect and physical Gibb's effect. The numerical Gibb's effect means the calculation is performed with insufficient number of diffracted orders for RCWA. This can be solved obviously by using a higher number of diffracted orders. On the other hand, physical Gibb's effect means that the calculation was carried out with sufficient number of orders for accurate calculation; however the angles of the majority of the diffracted orders were larger than the maximum angle of the objective lens aperture and they cannot be collected by the objective. In other words the physical Gibb's effect represents the situation we would expect to see in an experiment where the numerical Gibb's phenomenon is a limitation of the numerical calculation. The 30 microns case shown in figure 5.4b suffered from the numerical Gibbs effect, although it was computed with the highest number of diffracted orders possible before running out of memory as shown in appendix B. The 30 microns case was calculated using 121 diffracted orders (*N=60*), the following calculation confirms that the 30 microns case suffered from numerical Gibbs effect.



a)

b)

Figure 5.4 shows plasmonic angles θ_p of non-interferometric microscope for 4 micron, 30 microns grating periods and uniform sample measured at the central point material of *n=1.5* with thickness of 10 nm. *NA=1.65* n_0*=1.78* gold thickness of 46 nm and 633 nm. The sample was at the focal point and in water ambient. a) For *x-polarisation* and b) For *y-polarisation*

The maximum angle of the objective lens has a wave vector value of $\frac{2\pi}{\lambda}NA = 1.64 \times 10^7$. The maximum diffracted angle calculated using 61 diffracted orders assuming normal incidence is $60k_g = 60\frac{2\pi}{grating\ period} = 1.26 \times 10^7$. This means the calculation with 121 orders covered about 78% of the full objective lens aperture.

Recently, Kwiecien and Richter demonstrated a way to get around the Gibbs effect, where they extended the concept of rigorous wave coupled analysis by matching the tangential electric and magnetic fields using wave expansion and aperiodic field conditions (Kwiecien & Richter 2011). If there are aperiodic field components, this calculation will, of course, minimise the Gibbs effect.

The results obtained for both the *x* and *y polarisations* are sufficient to conclude that the best position for the non-interferometric microscope is at the centre of each material.

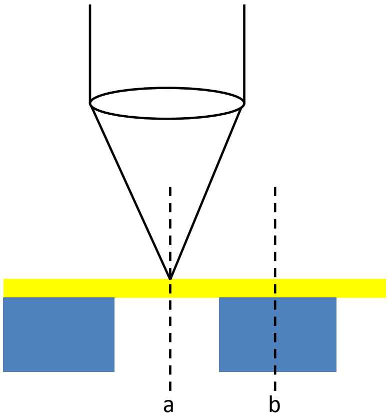


Figure 5.5 shows schematic diagram of measuring points

Therefore in this chapter all the calculations for the non-interferometric system were carried out at the centres of each material as shown in the diagram in figure 5.5.

5.2.2 Interferometric microscope

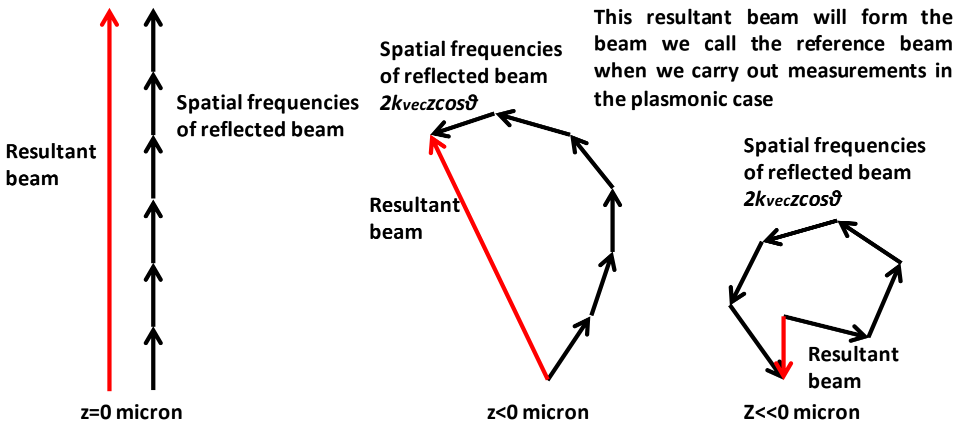
In the previous chapter, we have briefly introduced the scanning heterodyne interferometric microscope. In this chapter we will discuss in detail how the sensitivity of the *V(z)* arises and also the conditions under which it is valid.

Let us begin with the theoretical concept of the *V(z)* response, the *V(z)* response can be written as the summation of amplitude and phase between the reflected signal and the uniform reference signal over the full aperture of the objective, which can be expressed as (derived in chapter 2):

$$V(z) = | \iint P^2(sin\theta)\left[R_p(sin\theta)cos^2\emptyset + R_s(sin\theta)sin^2\emptyset\right]e^{jn_1 kcos\theta z}\, dsin\theta d\emptyset| \; [5.1]$$

Where $P(sin\theta)$ is the pupil function to modulate the intensity profile, θ is incident angle ϕ is azimuthal angle, $k = \frac{2\pi n_1}{\lambda_{free\ space}}$, $R_p(sin\theta)$ and $R_s(sin\theta)$ are complex reflection coefficients for *p-polarised* and *s-polarised* waves respectively.

The *V(z)* response represents the interference signal from the microscope. In the heterodyne interferometer we obtain an interference signal which is an integral of the field over the back focal plane. In the phasor diagram shown in figure 5.6, the interferometer reference beam is not shown; it is suffice to say that the interferometric signal is the complex sum of the back focal plane field. The phase relationship between the different components in the back focal plane plays a crucial role. Let us consider the case of a perfect reflector at different defocuses as shown in figure 5.6b and 5.6c.



a) b) c)

Figure 5.6 shows vector diagram for different z defocuses for the non plasmonic case; interference signal in red, reflected signal in black, reference signal in black a) *z=0* micron b) *z<0* micron for every single spatial frequencies and c) *z<<0* micron for every single spatial frequencies

This resultant beam is the reflected phasor arising from the microscope response as we defocus. The resultant beam will form the beam we call the reference beam when we carry out measurements in the plasmonic case.

Let us now investigate the behaviour of the interferometer in the presence of SP excitation.

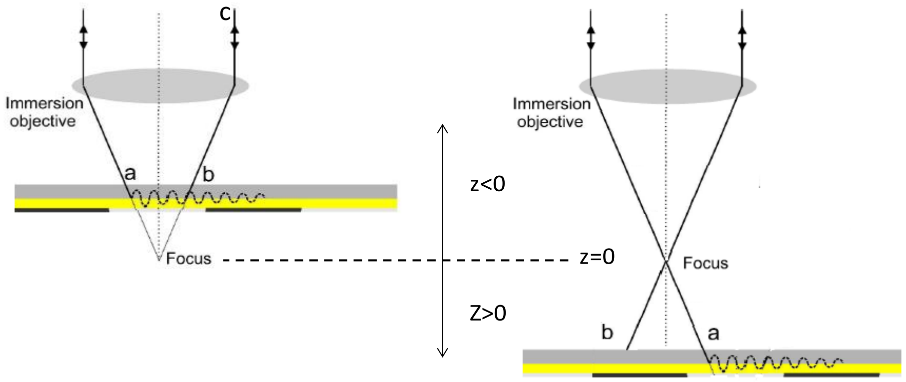


Figure 5.7 shows schematic diagrams for SP excitation a) negative z defocus b) positive z defocus

From figure 5.7a, the SP are excited at point ‘a’, which will propagate to point ‘b’ and then the phase and amplitude of the reflected signal at point ‘b’; optical path ‘bc’ will be modulated by the SP wave. On the other hand, for the positive defocus as shown in figure 5.7b; there will be an excitation of SPs which will propagate along the metal surface, however these will not be detected as its moving outwards the direction of reflected beam path ‘b’. In addition, if there are some ripples detected in the positive z defocus this mean there is a back scattered surface wave coming back at point ‘b’. Therefore for the two arm heterodyne interferometer, the path of the SP is well defined by the sample defocus and is not limited by the propagation length of the SP.

The simplest interpretation of the behaviour at negative defocus in the presence of SP excitation is that there is an effective interference between a reference beam incident close to normal incidence (this reference is not to be confused with the interferometer reference) and a beam that generates the SP. This reference beam is essentially the resultant beam for the non-plasmonic case shown in figure 5.6. We now consider the effect of this interference for an ideal case, we then discuss the conditions necessary to ensure that the behaviour approaches this ideal.

Let us now consider the effect of sample that produces a surface Plasmon. In this case we get a vector sum between the resultant phasor shown in figure 5.6 and another arising from the excitation of detection of a surface plasmon. Let us assume that the change in phasor of the 'reference phasor' goes as $2k_{vec}z$ and the phase of the SP goes as $2k_{vec}z\cos\theta_p$. In this case the relative phase between them will vary as:

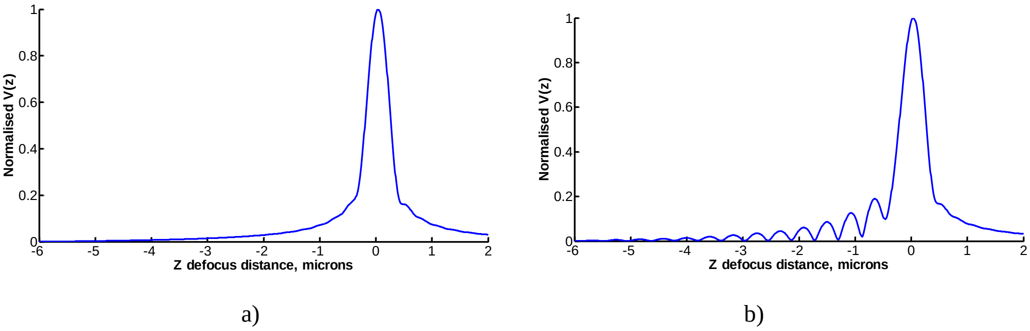
$$2k_{vec}(1 - cos\theta_p)z = \frac{4\pi n_0}{\lambda_{free\ space}}(1 - cos\theta_p)\Delta z.$$

Where θ_p is the plasmonic angle.

This phase will change by 2π when:

$$\Delta z = \frac{\lambda_{free\ space}}{2n_0(1-\cos\ \theta_p)} \qquad [5.2]$$

Where Δz is the ripple period, n_0 is the refractive index of the first material layer and θ_p is the plasmonic angle. The equation [5.2] defines the period of oscillation of the curve.



a) b)

Figure 5.8 shows $V(z)$ curve for a) non-plasmonic case 2 microns thick of gold and b) plasmonic case 46 nm of gold; 10 nm of layer of $n=1.5$ sitting on gold in water ambient with $NA=1.65$ and $n_0=1.78$, 633 nm wavelength and linear polarised wave.

Figure 5.8 shows two $V(z)$ curves one with SP and one without SP. We can see that oscillation which depends on the value of θ_p. Equation [5.2] provides means of determining θ_p. This is only valid, however, if the assumption that the reference varies as $2k_{vec}z$ is valid. We will now examine the condition under what this holds by considering what happens with the uniform sample in more detail.

For the non-plasmonic case, assuming linearly polarised light with its electric field along the x axis; negative z means the sample is moved towards the objective lens, *z=0* means the sample is at the focal point of the objective and positive z means the sample is moved away from the objective as shown in figure 5.9a below.

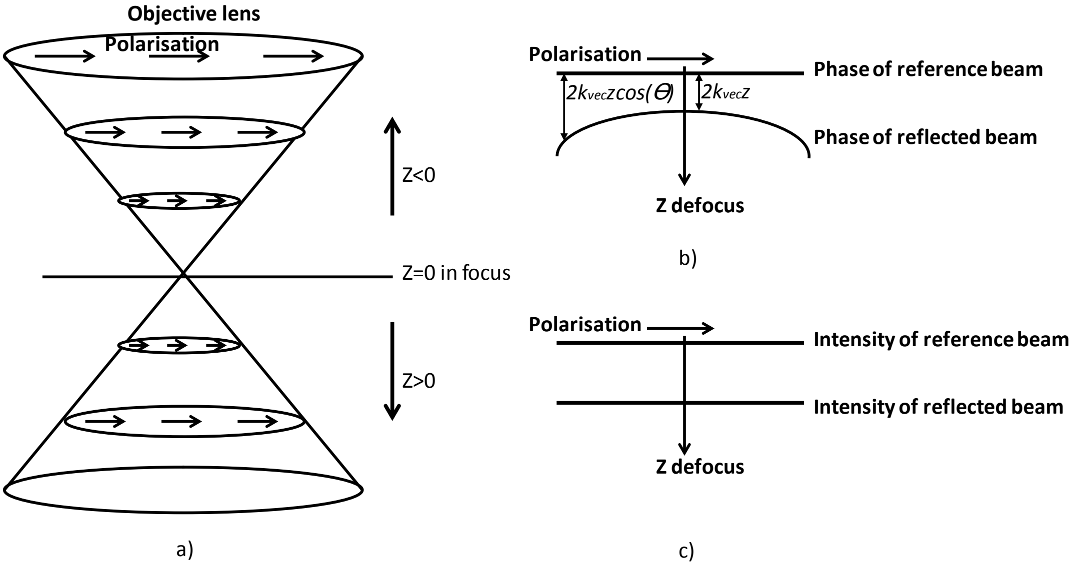
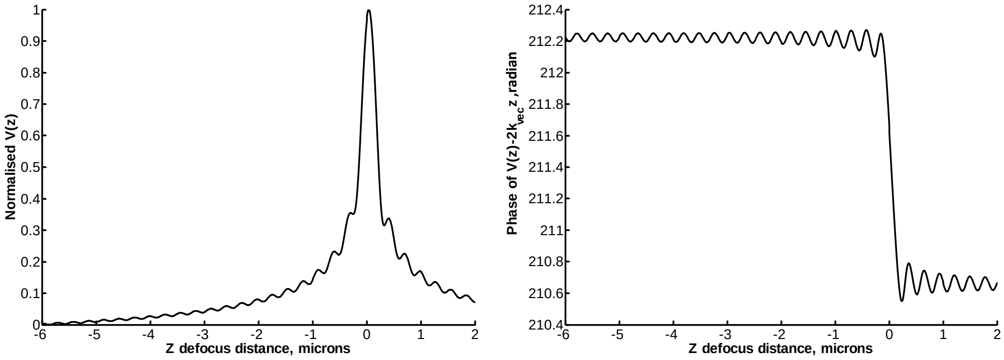


Figure 5.9 shows a) the defocus positions and electric fields b) Phases of reference and reflected beams and c) Intensity of reference and reflected beams

Having discussed in figure 5.6 that the concept of *V(z)* response can be illustrated by the phasor diagram of the individual spatial frequencies on the back focal plane images, therefore in this section we will make use of this concept to illustrate the phase behaviour of the reflected beams. Figure 5.9b shows the phase of the reflected beam at the central of the back focal plane when it is defocused is given by $2k_{vec}z$, on the other hand the phase at the other incident angle is shifted by $2k_{vec}z\ cos\theta$ where θ is the incident angle. However, the intensity on all the points of the back focal plane will be identical to the reference beam assuming the sample is a perfect mirror. The linearly polarised reference and reflected beams will interfere with each other as they both have the same polarisation.

The low spatial frequency component on the back focal plane of the reflected beam will contribute to the main decaying amplitude of the interference signal, which will decay along z, whereas the high frequency components on the back focal plane will cause oscillation around the main amplitude; the interference signal for the uniform mirror case is shown in figure 5.10a. The phase of $V(z)$ subtracted from $2k_{vec}z$ is shown in figure 5.10b.



a) b)

Figure 5.10 shows a) normalized interference signal and b) phase (in radians) of $V(z)$- $2k_{vec}z$. $NA=1.65$ and $n_0=1.78$ with uniform layer of gold *2 microns* thick with incident wavelength of *633 nm in water ambient*. This $V(z)$ curve was calculated with a uniform pupil function

More importantly, if the phase of the reference beam for the perfect mirror does not follow $2k_{vec}z$ as the sample along the z defocus, i.e. the difference is not constant, this will, of course, lead to some systematic error for the equation [5.2]. Therefore a more precise indicator that can be used to design the pupil function is to see how linear the phase of the high frequency components on the bfp are; this can be done by determining the phase of $V(z)$ and subtract it with the $2k_{vec}z$ term as shown in figure 5.10b. The uniformity of the subtracted phase shows how well the reference behaves. We can see that the uniform pupil function cannot provide a steady reference for the SP because the phase difference oscillates so in this case the assumption that the reference varies as $2k_{vec}z$ is not valid, so consequently the recovered θ_p is incorrect. The $2k_{vec}z$ assumption may be achieved by attenuating the high frequency components by a process called ‘apodizaion’. A simple way to get around this problem is, of course, by reducing the intensity of the high frequency components on the back focal plane; this may be achieved using an amplitude spatial light modulator (A-SLM) or by modulating the intensity beam profile by other means.

Therefore our main consideration to design pupil function for a *V(z)* microscope system is to reduce the higher spatial frequencies effect on the back focal plane. Let us consider the case where we have different pupil functions as shown in figure 5.11.

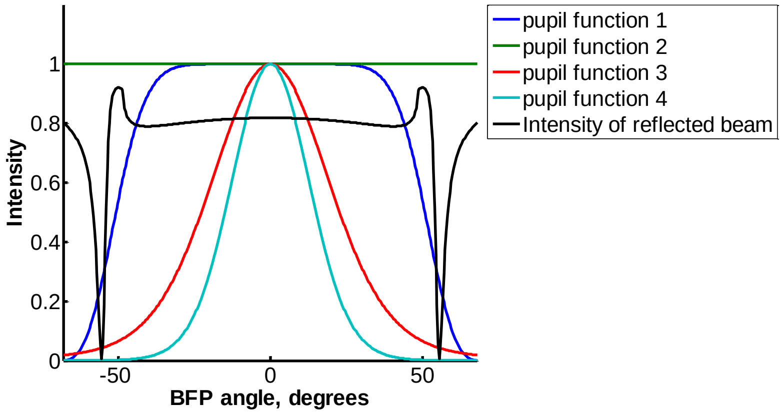
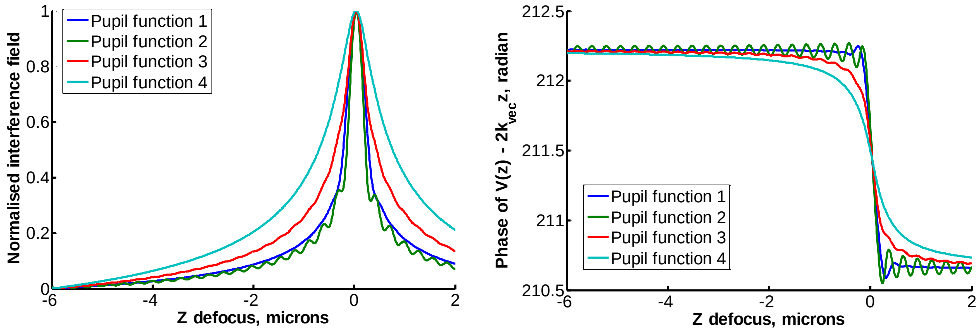


Figure 5.11 shows different pupil functions with reflected beam intensity for plasmonic case. *NA=1.65* and $n_0=1.78$ with uniform layer of gold *46 nm* thick with incident wavelength of *633 nm in water ambient.* The amplitude reflection coefficient in the presence of surface plasmons is overlaid in black.



a) b)

Figure 5.12 shows a) *V(z)* responses and b) phase (in radians) of *V(z)*- $2k_{vec}z$ corresponding to the pupil functions in figure 5.11 for non-plasmonic case. *NA=1.65* and $n_0=1.78$ with uniform layer of gold *2 microns* thick with incident wavelength of *633 nm in water ambient*

From figure 5.12a we can see that for the non-plasmonic case the uniform pupil function suffers severely from the high frequency resonance and we can still see some tiny ripples for the pupil function 3, however, we see no ripple for the pupil functions 1 and 4. This indicates that we might be able to use either pupil function 1 or 4 to provide a steady reference for the non-plasmonic case.

Having discussed that the way I derived the equation [5.2] relies on the phase behaviours of the terms $2k_{vec}z\,cos\theta$ and $2k_{vec}z$, from figure 5.12b we can see that the pupil function 2 (uniform pupil function) oscillated for all the z defocuses. There was still some oscillation at the first ripple for the pupil function 1; this will distort the performance of pupil function 1 for the reference, however, it does show the phase linearity over the z defocus range. The third pupil function also had some tiny amplitude of oscillation over the z defocuses, whereas the fourth pupil function appears to have no ripples at all.

The next question is to investigate which one of the pupil functions is most suitable for SP detection.

Plasmonic case with different pupil functions

For the plasmonic case, as discussed in the second chapter, at the SP angle there will be a 2π radian phase shift as shown in the diagram figure 5.13 below. The 2π phase shift will change the wavefront of the reflected beam around the plasmonic dip; this can be seen by the ripples on the *V(z)* curve as shown in figure 5.8b earlier.

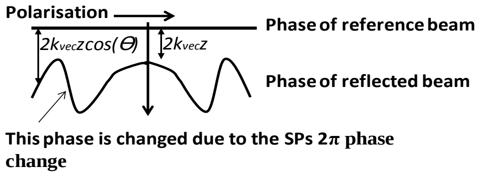
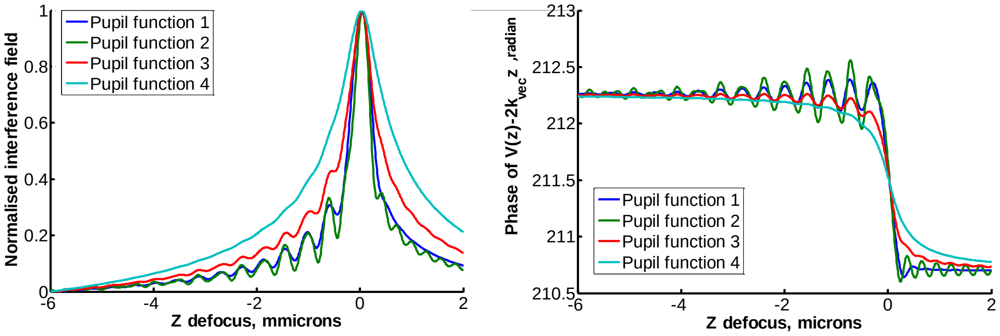


Figure 5.13 shows phase profile of the reference and reflected beams

Let us now consider the effect of the same set of pupil functions as shown in figure 5.14 for the plasmonic case by replacing the thick mirror by 50 nm of uniform gold.



a) b)

Figure 5.14 shows a) *V(z)* responses b) phase (in radians) of *V(z)*- $2k_{vec}z$ corresponding to the pupil functions in figure 5.11 for plasmonic case. *NA=1.65* and n_0*=1.78* with uniform layer of gold *46 nm* thick with incident wavelength of *633 nm in water ambient*

The results in 5.14a show that the uniform case had the highest ripples amplitude followed by pupil function number 1,3 and 4 respectively, where for the pupil function number 4 we can see almost no ripples. Therefore the optimum pupil function we can use is pupil function 1, where the high frequency oscillation is quite low and we can still detect strong SP ripple amplitude. The key trade-off therefore for designing a pupil function is that we need to compromise between the amplitude of the SP ripples and the steadiness of reference. For this simulation study, the pupil function 1 provided a good compromise for this requirement.

Figure 5.14b tells us some interesting stories; the SP ripple period for pupil function 2 is not constant; some of the amplitude where here we expect to see the decaying amplitude of ripple and also the Δz ripple periods changed all over the z defocuses; with z close to zero the ripple periods was shorter than the far negative z defocus; this was because the SP ripples were modulated with high frequency components oscillation. On the other hand, for the pupil function 2 and 3, each ripples of the two pupil functions seemed to be aligned with each other over the z defocus distance.

Let us now determine the plasmonic angle θ_p by determining the ripple period Δz and substituting it into equation [5.2]. It is important to note that there are also quite a number of ways to determine Δz as the $V(z)$ has multiple ripples over the z defocus distance. Also each of these ripples has different ripple period Δz, this is because the rate of change of the phase $V(z)-2k_{vec}z$ is not constant for all the z defocus positions as discussed.

The closer of the z defocus position to the objective lens gives a more accurate θp result; the reason why this is the case is that for the high values of negative z defocus the relative phase $2k_{vec}z\, cos\theta$ compared to $2k_{vec}z$ will move much slower than the low negative defocus case. In other words, the $2k_{vec}z$ approximation improves with defocus. However there is a trade-off between accuracy of the result and signal to noise ratio (SNR) that can be detected.

From equation [5.2], we can see that the parameter that will affect θp is Δz; therefore we have to carefully choose the best way to measure the ripple period Δz since each of the ripples has different value of Δz; this is due to influences of pupil function and high frequency component oscillation as explained.

An obvious choice to calculate the Δz is to determine its dominant ripple frequency by taking Fourier transform or Hilbert transform. This will account for the fact that each ripple has different amplitude and period. In other words, the FFT and Hilbert transform can give us an averaged Δz over the z defocus range. Other ways that can be done are firstly measure the Δz period of each ripple individually and secondly average value of Δz by measuring a few ripple distance; this will, of course, reduce the effect of noise. For the second approach, the ripple period Δz depends on number of ripples to measure.

The table summarizing the performance of pupil functions is shown in table 5.1 below.

Plasmonic angles *θp*, degrees	**1st ripple**	**2nd ripple**	**3rd ripple**	**Accumulative first 2 peaks**	**Accumulative first 3 peaks**	**Fourier transform**	**first ripple/ background ratio**
Pupil function 1	54.90	54.47	54.68	54.69	54.68	54.71	0.07
Pupil function 2	55.98	53.88	55.47	54.90	55.09	55.48	0.12
Pupil function 3	54.98	54.20	54.87	54.59	54.68	54.91	0.05
Pupil function 4	52.37	57.29	52.55	54.66	53.93	53.61	0.01

Table 5.1 shows plasmonic angles *θp* calculated from *V(z)* responses in figure 5.14 with a uniform layer of *n=1.5* 10 nm thick in water

For the non-interferometric case the uniform sample case, the plasmonic angle measured at the centre of *n=1.5* was 55.50 degrees, on the other hand, the interferometric approach seemed to give the plasmonic angles θ_p for all pupil functions below the plasmonic angle θ_p measured using the non-interferometric system. Although the interferometric method gives us very good spatial resolution for imaging, it does not give an accurate value for quantitative measurement; this is due to the conditions as explained earlier.

The longer propagation length will give closer results to the plasmonic angle of non-interferometric case compared to the shorter propagation length as the phase around SP angles have a steeper phase transition as discussed in chapter 2; this makes the angles more accurate as the interference signal is contributed by the angles that are close to the actual θ_p. This is because the magnitude of the reflected beam at the plasmonic angle is low; therefore the actual spatial frequency component that will interfere with the reference beam is the frequency components around the SP angle rather than the angle itself.

For this study I chose the pupil function 1 and used it throughout this chapter. The reasons I chose such a pupil function was that firstly the magnitude of the ripples arising from the SPR was quite strong although it was still less than the uniform pupil function (pupil function2) because we can see that the real advantage of the pupil function1 was that it had smaller $2k_{vec}z\,cos\theta$ oscillation in the absence of SP excitation. Although the uniform pupil function seemed to give better θ_p performance, it will be very hard to distinguish between the SP ripple and the $2k_{vec}z\,cos\theta$ high frequency component oscillation as they might be at a similar frequency so called in band noise. The third pupil function was also a good candidate for this task as it had a closer value to the non-interferometric θ_p, but the ripples size was smaller than pupil function 1 therefore we have to choose the pupil function that allows us to compromise between ripples size and accuracy of measurement. From a simulation point of view, either pupil function 1 or 3 are fine, however for a real experiment we need to, of course, use pupil function 1 in order to make sure that we have sufficient SNR to overcome the noise. The fourth pupil function seems to eliminate all the ripples and this will not be sensitive to SP.

Although FFT is a very good measurement that averages over all the ripples; this is the ideal measurement technique for the uniform sample case, the samples used for this study was grating sample therefore the different z defocus distance will, of course, cover different sizes of footprint on the sample and it might as well cover too much of the area than we intend to measure as depicted in figure 5.15.

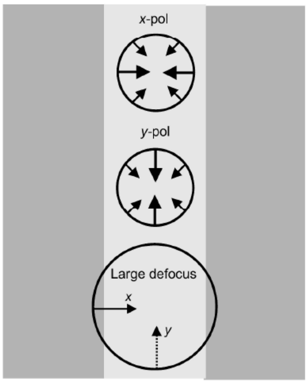


Figure 5.15 shows footprint on the sample at different z defocuses

Assuming that the plasmonic angle is almost at the edge of the objective the coverage area on the sample can be then expressed as:

$$Coverage\ region = \left|2z\tan\theta_p\right| \qquad [5.7]$$

Therefore in this case the FFT and Hilbert methods are not a good choice for the non-uniform sample as they take both of the regions we want to measure into account at the same time. Let us consider the cases that I used in this research which were $NA=1.65$ with $n_0=1.78$ and $NA=1.49$ with $n_0=1.52$. The coverage regions for the two objective lenses are presented in figure 5.16.

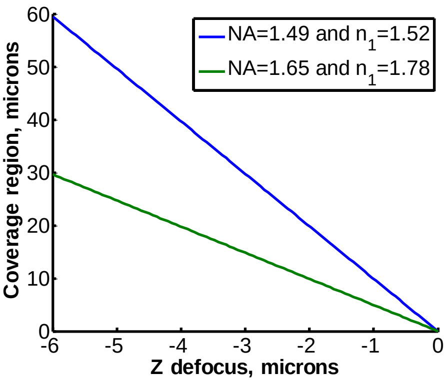


Figure 5.16 shows coverage length on the sample at different *z* defocuses

The coverage length does not only tell us about the *z* defocus length limit, but also enables us to understand the effect of crosstalk when we measure over the z defocus. For example if one tries to measure θ_p at a centre point of *n=1.5 10 nm* thick with a grating period of *8 microns* (*4 microns* is *n=1.5* and another *4 microns* is water *n=1.33*) without considering the effect of crosstalk over the *z* defocus using *NA=1.65* with $n_0=1.78$. They would expect to see the results as shown in the table 5.2.

Centre of the ripple on the z axis, microns	**-0.664**	**-1.088**	**-1.52**	**-1.944**
Region covered, microns	3.262 μm	5.337μm	7.413μm	9.488 μm
n=1.5 region	3.262 μm of n=1.5	4 μm of n=1.5	4 μm of n=1.5	5.488 μm of n=1.5
n=1.33 region	0 μm of n=1.33	1.337 μm of n=1.33	3.413 μm of n=1.33	4 μm of n=1.33
Plasmonic angle of Uniform sample *n=1.5*, degrees	54.90	54.47	54.68	55.14
Plasmonic angle of bare gold in water, degrees	54.08	53.48	53.52	53.72
Plasmonic angle predicted without crosstalk calculated by interpolation using uniform samples θp, degrees	54.90	54.22	54.15	54.54
Plasmonic angle simulation results with crosstalk for x polarisation, degrees	54.61	54.55	54.05	54.68
Plasmonic angle simulation results with crosstalk for y polarisation, degrees	54.67	54.61	54.36	55.03

Table 5.2 shows plasmonic angles measured at the centre of n=1.5 of 8 microns grating period compared with the uniform sample with the same thickness of 10 nm

In addition, table 5.2 also provides us with a strong evidence to say that the *V(z)* measurement is also sensitive to its coverage region as the simulation results agreed very well with the trend of the prediction, although the values of θ_p are different due to the crosstalk.

As explained earlier that the longer negative z defocus will give us a more stable result, however, here we have to also consider the size of footprint of the focused beam on the sample. We cannot go too far with the z defocus therefore the method I used for this study was the first 2 consecutive peaks accumulative measurement. For small gratings 2 ripples were needed here, however for the longer gratings where the footprint remains in the region 3 ripples is superior due to the additional stability.

From an instrumental point of view measurement to the first few ripples encourages the use of thinner layers of gold which give larger amplitude ripples for small defocuses and the rapid decay of these ripples is not an issue because the later ones will not be used.

Quantification of crosstalk in interferometric microscope

The next task is to investigate the interferometric microscope to see whether we can accurately measure the plasmonic angle θ_p if we perform a *V(z)* measurement at the centre of material for different grating period and compare the results with the uniform sample case. The results in figure 5.17 confirm that there were also some crosstalk effects for the interferometric microscope although the crosstalk was very small compared to the non-interferometric case as for the non-interferometric case it was required to go up to 30 microns period to obtain the correct plasmonic angle compared to the uniform case for the both *x* and *y polarisations*, on the other hand we can see clearly that even for the 8 microns grating period the *V(z)* ripples were very close to the *V(z)* for the uniform case.

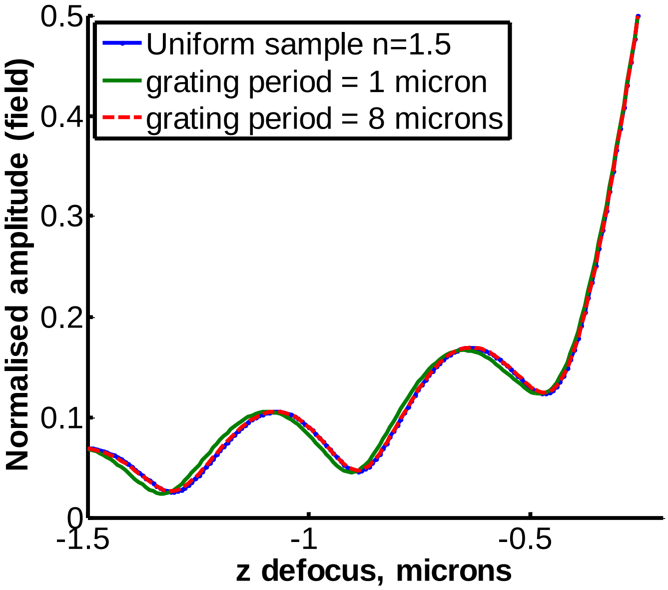
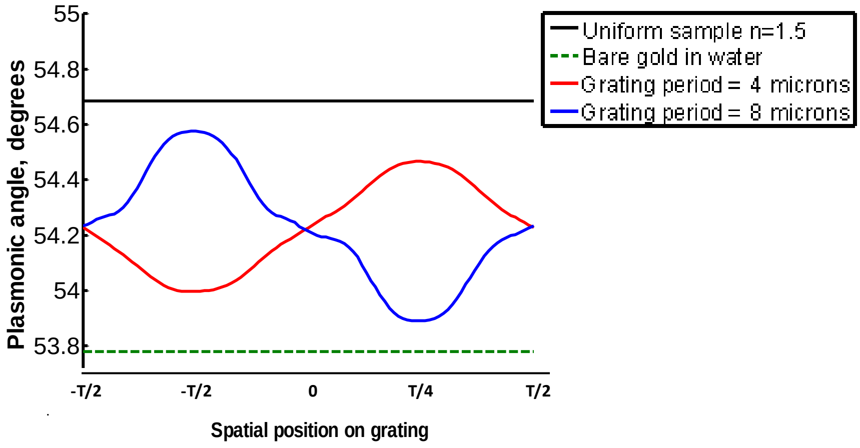
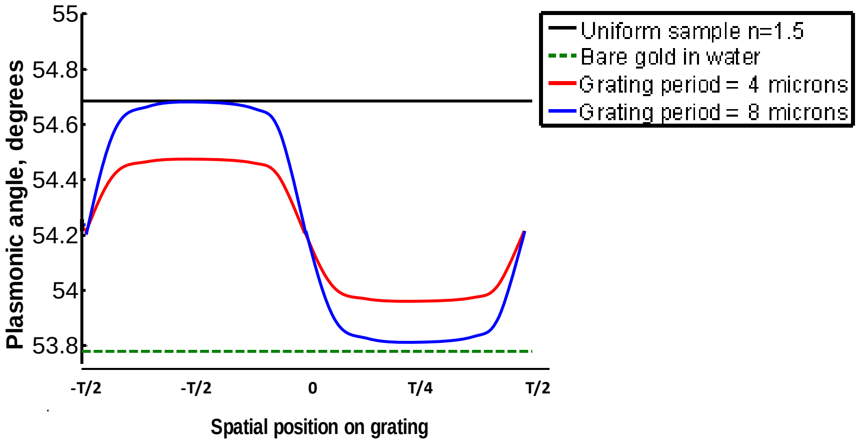


Figure 5.17 shows *V(z)* for *1 micron*, *30 microns* grating periods and uniform sample measured at the central point material of *n=1.5* with thickness of *10 nm*. *NA=1.65* n_0*=1.78* gold thickness of *46 nm* and *633 nm x polarisation* incidence. The sample was in water ambient. Remark that the results for *y* and *r polarisations* are similar to this figure with an exception that the *r polarisation* has bigger ripples.

The equation [5.2] enables us to calculate the plasmonic angle *θp* from the *V(z)* ripples. The *θp* results converted from *V(z)* measurements for different positions on grating samples are shown in figure 5.18, this is similar to the non-interferometric case which confirms that the best position to measure *θp* is at the centre of each materials.



a)



b)

Figure 5.18 shows plasmonic angles *θp* of interferometric microscope for *4 micron*, *8 microns* grating periods and uniform sample measured at the central point material of *n=1.5* with thickness of *10 nm*. *NA=1.65* n_0*=1.78* gold thickness of *46 nm* and *633 nm*. The sample was in water ambient. The plasmonic angle was measured using accumulative measurement of the first two ripples a) For *x-polarisation* and b) For *y-polarisation*

There is an interesting contrast reversal figure 5.18a, which shows that measurements on the 4 micron periodicity gratings are not valid for *x polarisation*. This will be explained in the results section later.

5.2.3 Comparison of the non-interferometric and interferometric systems

It is necessary to compare performances of the non-interferometric and interferometric microscopes, which have different plasmonic angles. Therefore the best way to compare them is to use relative value η, which is given by:

$$\eta = \frac{\theta_{p,grating\ at\ the\ centre\ of\ n=1.5} - \theta_{p,grating\ at\ the\ centre\ of\ n=1.33}}{\theta_{p,uniform\ sample\ of\ n=1.5} - \theta_{p,bare\ gold\ with\ water\ of\ n=1.33}} \quad [5.9]$$

The η term tells us about how closely we can resolve the SP plasmonic angle in a confined region compared with the uniform sample case.

5.2.4 Polarisation consideration

It is worth pointing out at this stage that many authors have advocated the use of radial polarisation in SP microscopy (Moh *et al* 2008). As explained in the previous chapter, this allows all the incident light to be *p-polarised* so that the magnitude of the *V(z)* oscillations is increased. Moreover for the contrast mechanism that requires a tightly focused spot, for radial polarisation the diffraction limited spot is tighter and more symmetrical; it is therefore gives clear advantages.

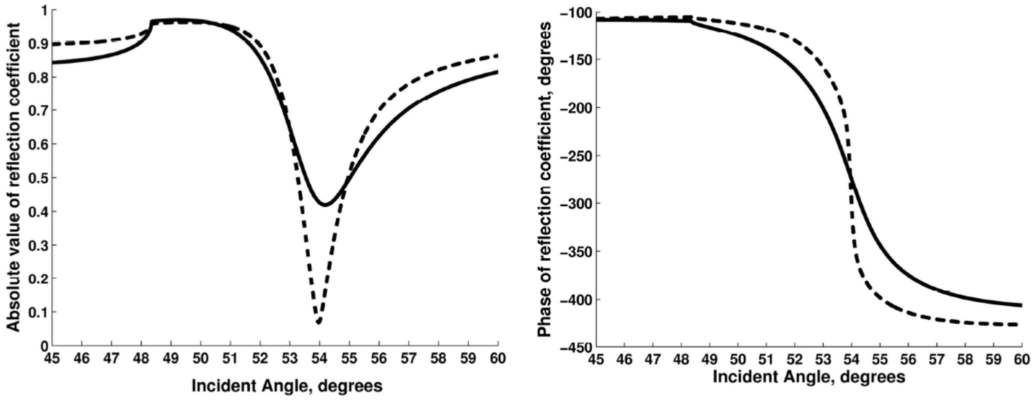
For the grating sample case here, however, the merits of radial polarisation are less clear cut and in this paper we will consider linear polarisation oriented both parallel and perpendicular to the grating vector as this will allow us to obtain a better physical picture of the mechanisms involved. Simulations were also produced for radial polarisation which shows curves intermediate between the *x*- and *y-polarisation* cases and these are omitted for brevity.

5.3 Simulation results

Simulation results obtained from full back focal plane microscope simulations using RCWA calculation will be presented; firstly the non-interferometric microscope results, followed by the interferometric results in order to confirm that the interferometric microscope performance is better than the other and investigate the effects of crosstalk on the interferometric system as from this chapter onward results presented will be about the interferometric and confocal microscopes, where some mathematical techniques and optical devices are employed to extract some SP information from the interference signal described in this chapter. Therefore it is very important to understand the effect of crosstalk on the interferometric microscope. Although the crosstalk results for the non-interferometric microscope are not considered here, the concept of crosstalk in the interferometric microscope is still applicable.

5.3.1 Simulation results for the non-interferometric microscope

Firstly let us look at the plasmonic dips of two cases 36 nm and 46 nm of uniform gold in figure 5.19 to remind ourselves about the propagation length effect.



a) b)

Figure 5.19 shows a) Modulus of reflection coefficient from gold sandwiched between two half spaces. *Wavelength=633nm, NA=1.65,* $n_1=1.78$, $n_{gold}=0.17+3.52i$, $n_{lower}=1.33$ (water). The sample is at the focal point of the objective. Solid curve 36nm gold thickness, dashed curve 46nm gold thickness.

b) Phase of reflection coefficient corresponding to case of figure 5.19a.

I explained in chapter 2 that steeper gradient of the phase of the reflection coefficients indicates the longer propagation length of the SP. Therefore in this case the 46 nm has a longer SP propagation length than the 36 nm case. We can see from figures 5.20 and 5.21 that for the non-interferometric case the propagation length of the SP affects the rate of convergence to the value associated with the uniform sample, the 36nm coating giving better performance than the thicker coating with the longer propagation length. The effect of the NA is not so pronounced as the higher NA does not improve the measurement resolution. The only factor that improves the performance of the non-interferometric system is the propagation length; the shorter propagation length gives a better resolved plasmonic angle resolution.

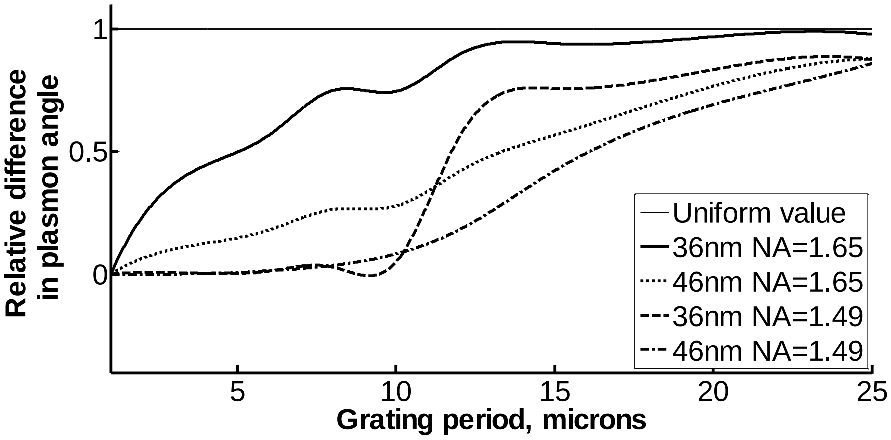


Figure 5.20 shows difference in recovered value *n* of SP excitation angle, *θp*, on coated (10nm dielectric n=1.5) layer and bare layers versus grating period for incident *x-polarisation* for non interferometric microscope, determined by position of dip in back focal plane. Solid curve: 36nm gold layer *NA=1.65*, dotted curve: 46nm gold *NA=1.65*, dashed curve: 36nm gold *NA=1.49*, dot-dashes: 46nm gold *NA=1.65*.

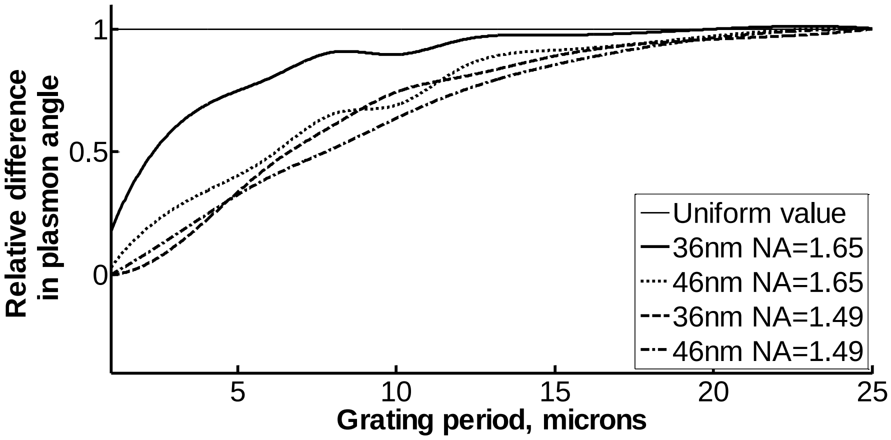


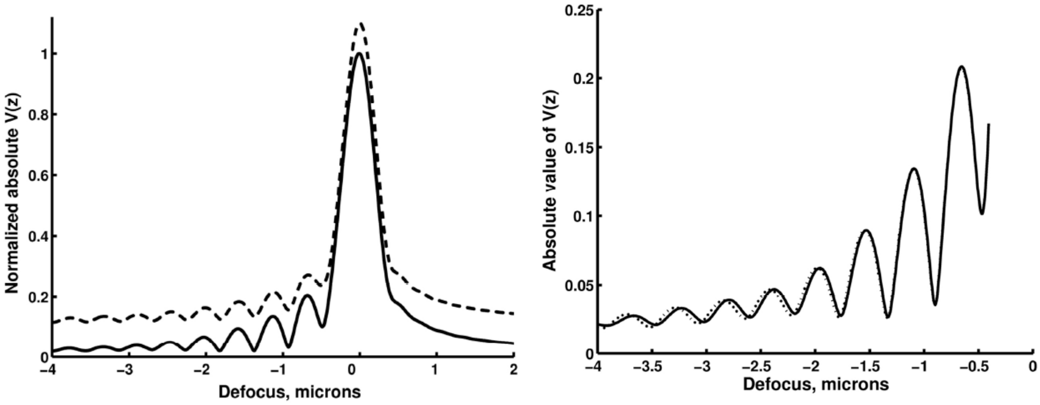
Figure 5.21 shows difference in recovered value *n* of SP excitation angle, θp, on coated (10nm dielectric *n*=1.5) layer and bare layers versus grating period for incident *y-polarisation* for non interferometric microscope, determined by position of dip in back focal plane. Solid curve: 36nm gold layer *NA=1.65*, dotted curve: 46nm gold *NA=1.65*, dashed curve: 36nm gold *NA=1.49*, dot-dashes: 46nm gold *NA=1.65*.

By comparing the performance of *x* and *y polarisations* in figures 5.20 and 5.21, we can clearly conclude that *y polarisation* can be better used to measure the SP parameter over a confined region, however, even for quite big grating periods such as

25 microns the value of η does not reach a hundred percent. In other words, even though *y polarisation* is better than *x polarisation*, it is still recommended that the SP measurement is carried out using either uniform sample or microfluidic channels with a long separation such as 15 microns and 15 microns pad of detection area.

5.3.2 Results for the interferometric microscope

From figure 5.22a, we can see that the 36 nm case ripples have higher amplitude than the 46 nm case. In other words, thinner layer of metal gives stronger ripples amplitude. This finding has been experimental confirmed by (Roland *et al* 2010). They have recently discussed the use of thinner layers in microscopy where the SP is more strongly excited. It is interesting to note that for a conventional SP sensor better sensitivity would be obtained with the thicker layer which supports SP with a longer propagation length, while the stronger excitation with thin layers encourages the use of thinner layers for microscopy applications. Figure 5.22b shows *V(z)* curves obtained on a uniform sample compared with 6 microns grating, this is strong evidence of the presence of crosstalk. The *V(z)* on the grating was obtained at the midpoint of the region of the material directly under the axis of the beam is the same, in the grating case, however, the properties of the adjacent material vary. We see that initially the curves obtained on the grating are close to those on the uniform sample but gradually start to deviate substantially as the defocus increases, since the points of excitation and detection 'a' and 'b' (Figure 5.7a) become separated they are no longer enclosed in the region we are trying to measure. The period of the dashed curve associated with the grating becomes longer since the SP now also propagate on the bare region where the *V(z)* periodicity as predicted by equation [5.7] is larger as discussed in section 5.2.2.



a) b)

Figure 5.22 shows *V(z)* curves calculated for the same case as shown in figure 5.19 Solid curve 36nm gold thickness, dashed curve 46nm gold thickness. a) full range of z defocus scan from -6 microns to 2 microns and b) *V(z)* curves on uniform sample (36 nm gold and 10 nm dielectric *n=1.5*), solid curve. The dashed curve is obtained in grating sample as shown in figure 5.1c, with same parameters as uniform sample on measured region *point 'a'* with n_b=1.33. Grating period 6μm, mark space ratio 50-50 and input polarisation along *x* direction.

In order to get a better idea of the effect of structure on the *V(z)* we present results of simulations which show the effect of the periodicity of the grating structure. These are shown in figure 5.23 incident *x-polarisation* and in figure 5.24 for *y-polarisation*. We can see as expected that as the periodicity of the grating increases towards 10 microns (corresponding to a local region of 5 microns) the values for both x- and y-incident polarisation tends to the value obtained with a uniform sample. It is also clear that the *y-input polarisation* tends to the value obtained on a uniform sample for much smaller grating periods compared to the case where the input polarisation is *x-polarised*. It is well known and established experimentally that SP resolution in non-interferometric imaging is superior when the SP propagates perpendicular to the grating (along the grooves) compared to the case where the SP propagates across the grating (Stabler *et al* 2004). The physical reason for these seemingly analogous effects is, however, different as considered in the discussion section. For a layer of thickness of 46nm the results converge as almost as quickly as those associated with the 36nm thick layer where the propagation length of the SP is shorter. This confirms our assertion that the confinement is not propagation length limited in the interferometer system but arises from the optical configuration.

We do notice a considerable difference between the performance of the *NA=1.65* and the *NA=1.49* cases with the rate of convergence between the value associated with the uniform sample being approximately between a factor of two or three greater for the higher numerical aperture. This is not due in any sense to the higher resolution associated with the greater numerical aperture but arises from the different angles at which the SPs are excited. In the case of the *NA=1.49* lens the index of the immersion fluid is 1.52 which means the SP are excited at approximately 72 degrees whereas in the case of the *NA=1.65* and immersion fluid with index 1.78 is approximately 54 degrees. If we now refer to the upper diagram in Figure 5.7a we see that points ‘a’ and ‘b’ are separated in the simple geometrical picture by$2(\tan\theta_p)z$, where z is the defocus, this means that the footprint is larger by the ratio of the tangents of the SP excitation angle, which taking the values above is approximately 2.2; we believe it is this geometrical effect that degrades the performance of the lower NA.

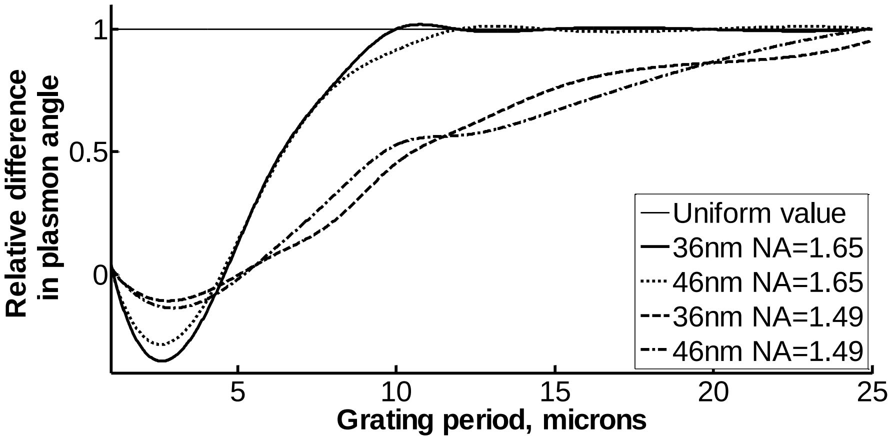


Figure 5.23 shows difference in recovered value *n* of SP excitation angle, *θp*, on coated (*10 nm* dielectric *n=1.5*) layer and bare layers versus grating period for incident *x-polarisation*. Solid curve: *36 nm* gold layer *NA=1.65*, dotted curve: *46 nm* gold *NA=1.65*, dashed curve: *36 nm* gold *NA=1.49*, dot-dashes: *46 nm* gold *NA=1.65*. Values of *θp* were obtained from the periods of the first two ripples and converted using [5.2].

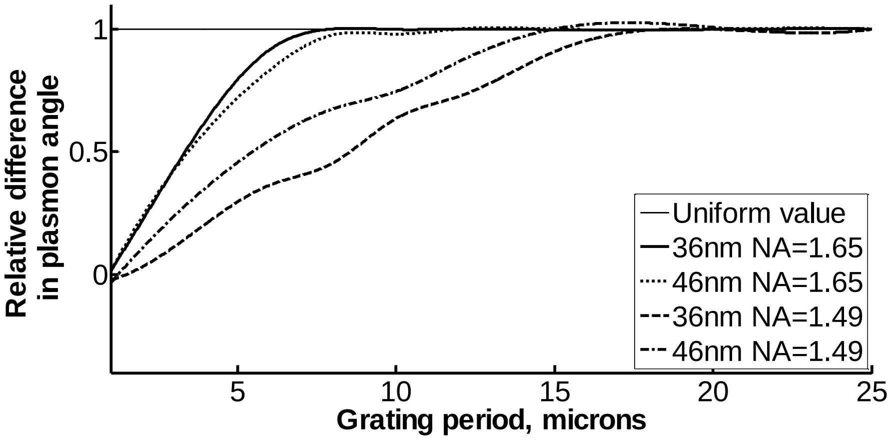


Figure 5.24 shows difference in recovered value η of SP excitation angle, *θp*, on coated (*10 nm* dielectric *n*=1.5) layer and bare layers versus grating period for incident *y-polarisation*. Solid curve: *36 nm* gold layer *NA=1.65*, dotted curve: *46 nm* gold *NA=1.65*, dashed curve: *36 nm* gold *NA=1.49*, dot-dashes: *46 nm* gold *NA=1.65*. Values of *θp* were obtained from the periods of the first two ripples and converted using [5.2].

It is interesting to point out that for the grating period below 5 microns, η is negative for *x polarisation* as the *θp* of the two regions seem to switch over each other in figures 5.18a, 5.23 and 5.25. We are definitely of a firm view that the switching over effect does not actually mean either we can measure the property of region 'b' by measuring it at the centre of region 'a' or the other way around. The effect can be seen for the both NAs, only for *x polarisation* and for short period of grating; therefore it must be that the effect occurs because of the propagating length of SP with strong reflection when the SPs excited are propagated to the end on the material region and then reflected back. The reflected SPs will then modulate the phase and amplitude of the propagating SPs. The reflected SPs (shown in red line in figure 5.26) are also scattered back by the other end of the material and this forms multiple reflections within the grating period.

The evidence of this effect is that when mismatched intrinsic impedance of the SP is reduced (reflection coefficient is reduced), the plasmonic angle should get closer to its uniform value. Therefore let us assume that we can change the refractive index n_{2b} shown in figure 5.1c from 1.33 to 1.2 and 1.4 respectively by keeping $n_{1b}=1.5$. This means for the case $n_{2b}=1.2$ the mismatched impedance is increased, on the other hand for the case $n_{2b}=1.4$ the mismatched impedance is decreased. Therefore we expect to see that for the case $n_{2b}=1.4$ the plasmonic angle θ_p measured at the centre of $n_{1b}=1.5$ will get closer to its uniform value and in contrast for the case $n_{2b}=1.2$ the plasmonic angle θ_p measured at the centre of $n_{1b}=1.5$ will be even worse. The result has come up as predicted and been shown in figure 5.25, where the θ_p calculated from the *V(z)* curves in figure 5.27 using 3 consecutive peak measurements are 53.86, 54.03 and 54.48 degrees for $n_{2b}=1.2$, $n_{2b}=1.33$ *and* $n_{2b}=1.4$ respectively, where the uniform value for $n=1.5$ is 54.68 degrees.

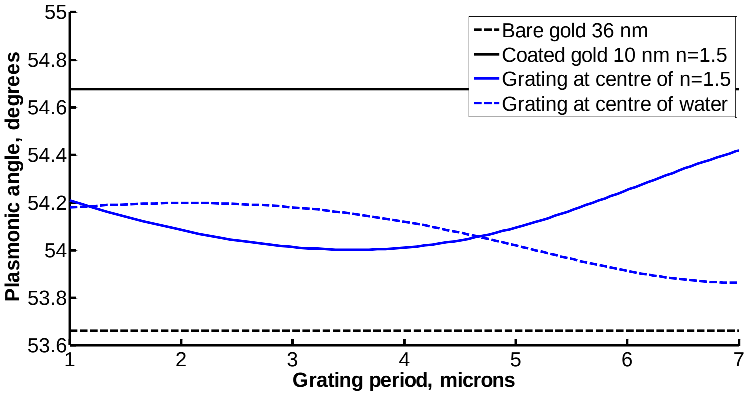


Figure 5.25 shows switching over effect of plasmonic angles; $NA=1.65$ $n_0=1.78$ $wavelength=633$ nm uniform gold with thickness of 36 nm. The plasmonic angle was calculated using 3 consecutive peaks measurement method. Black solid curve is uniform sample case, black dashed line is bare gold case, blue solid is the plasmonic angle measured at the centre of $n=1.5$ of the grating sample and blue dashed line is the plasmonic angle measured at the centre of $n=1.33$ of the grating sample.

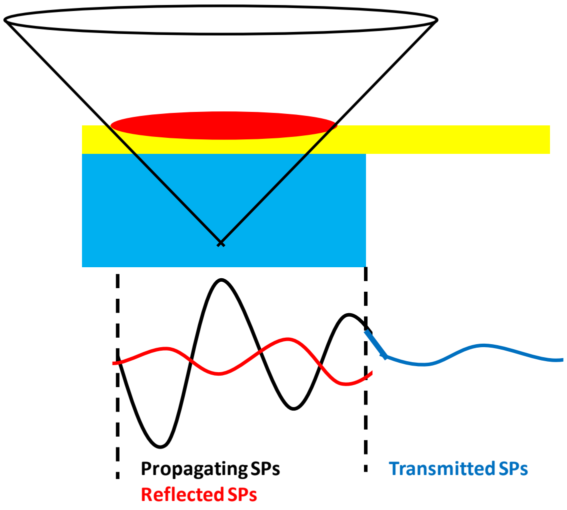


Figure 5.26 shows multiple reflections in a small grating period *x polarisation* on one side of material;

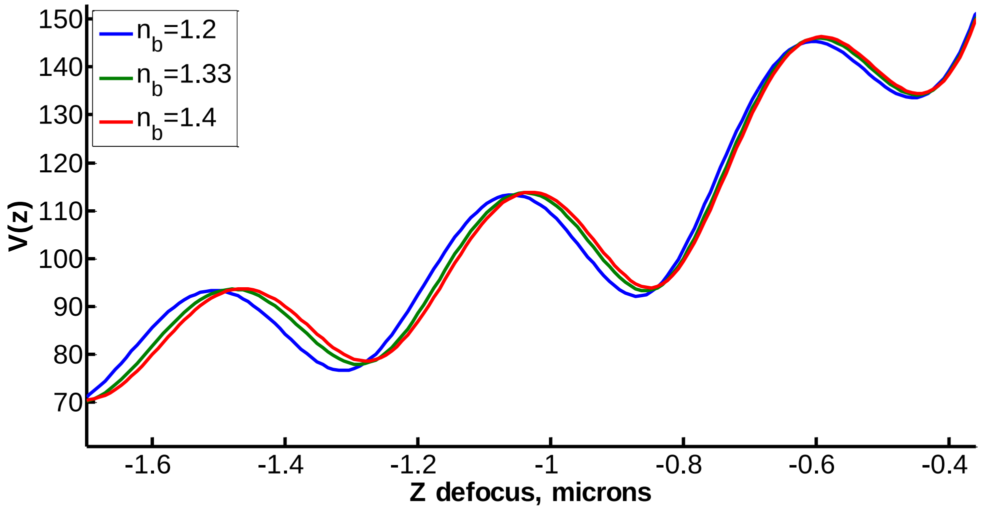


Figure 5.27 shows *V(z)* responses measured at the centre of $n_{2a} = 1.5$. NA=1.65 n0=1.78 wavelength=633 nm uniform gold with thickness of 36 nm. $n_{2a} = 1.5$(fixed) and $n_{2b}=1.2$, $n_{2b}=1.33$ *and* $n_{2b}=1.4$ (varied) grating period = 3 microns

Crosstalk in the interferometric microscope

It is also important to examine the effect of crosstalk between adjacent regions. We now focus on a region of index 1.5 and thus keep n_{2b} unchanged but consider the effect of varying the index of the adjacent region, n_{2a}, between a low value of 1.3 and high value of 1.7. Figure 5.28 shows the effect of changing the index of the adjacent region for a 10 micron grating using the first 2 *V(z)* ripples to perform the measurement. The *y*-axis shows the variation of the measured value of surface plasmon angle compared to a uniform sample, so a value of zero indicates no crosstalk. We see as expected that the crosstalk increases going from *y*- to *r*- to *x*- *polarisations*. We also see that the high NA objective with high refractive index couplant *(NA=1.65)* gives smaller crosstalk compared to the lower index (*NA=1.49* blue) as expected from figures 5.23 and 5.24.

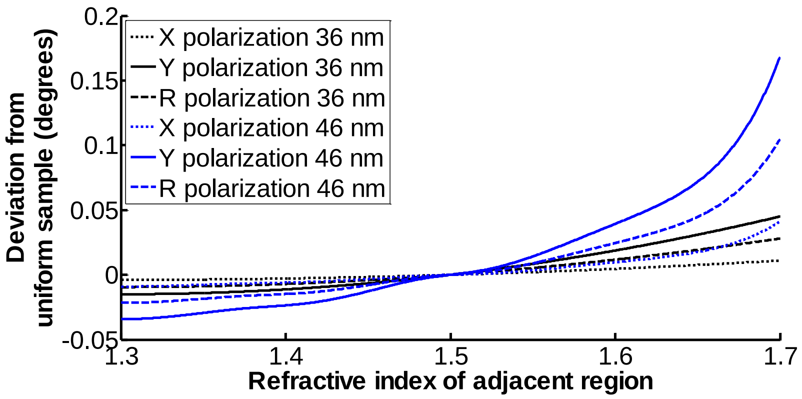


Figure 5.28 Deviation from recovered surface plasmon angle, θp, compared to a uniform sample as a function of refractive index of adjacent region for different layer thicknesses and polarisation angles. Black solid *y-polarisation*, dashed *r-polarisation*, dotted *x-polarisation* layer thickness 36 nm *NA =1.65*. Blue solid *y-polarisation*, dashed *r-polarisation*, dotted *x-polarisation* layer thickness 46 nm *NA=1.49*.

5.4 Discussion and conclusion

In this chapter, we have analysed both the interferometric and non-interferometric SP microscopes by numerical simulation and also given physical arguments to show that the interferometric performance in terms of measuring local refractive index is far superior to that of the non-interferometer configuration on account of the fact that the points of excitation and, more particularly, detection are strongly localized by the interference process. When operating the interferometric microscope in defocused *V(z)* mode far better confinement is obtained when the angle for SP excitation is kept as small as possible as demonstrated by the results that show superior performance when the SP are excited from a high index as is the case with the 1.65NA objective, with high index couplant.

The results presented here have been analysed for scanning heterodyne interferometer but may be applied to a widefield system as well (Somekh *et al* 2009) provided the pupil function can be controlled. The results discussed here should also be applicable to a conventional confocal system since the confocal system has identical coherence properties and transfer function to the interferometer. Finally, we reiterate that the concept of resolution in SP imaging needs to be related to the precise type of measurement we are trying to perform, with this in mind we believe that the present study complements (Elezgaray *et al* 2010) to give a more complete picture of resolution issues in SP microscopy.

The table 5.3 provides a summarised story for the interferometric microscope performance.

Grating period (microns)	η_2 **for *x-polarisation***	η_3 **for *x-polarisation***	η_2 **for *y-polarisation***	η_3 **for *y-polarisation***	η_2 **for *r-polarisation***	η_3 **for *r-polarisation***	η_3 **for *x-polarisation* at 25% position**	η_3 **for *y-polarisation* at 25% position**
5	9.952	-4.674	78.297	63.040	44.125	29.183	-24.396	47.280
7	56.835	35.761	92.806	78.260	74.821	57.011	6.959	57.130
9	87.410	62.500	98.561	92.390	92.986	77.445	28.849	70.216
11	96.043	89.565	99.880	97.830	97.962	93.698	42.813	78.264
13	97.722	92.717	100.360	99.674	99.041	96.196	51.862	79.640
15	100.120	98.370	100.959	100.109	100.540	100.217	53.730	80.087
17	99.880	100.000	100.600	100.000	100.240	100.000	56.789	81.531
19	99.400	100.000	99.880	100.000	100.000	100.000	58.102	82.563

Table 5.3 depicts ratio of recovered difference in surface plasmon angle on grating samples compared to difference in recovered surface plasmon angle on uniform sample for a 10nm layer with refractive index values of 1.5 and 1.33 respectively. Gold thickness 36nm, NA=1.65.

Where η_2 and η_3 in the table refer to the first 2 and the first 3 consecutive peaks accumulative measurement methods respectively. As explained earlier that the longer negative z defocus will give us a more stable result. From the table 5.3, it is clear to conclude that firstly the accumulative 2 peaks measurement method reaches the uniform sample condition quicker than the accumulative 3 peaks measurement method; however once the accumulative 3 peaks measurement method reaches the uniform case the results are a lot more stable compared to the other case. The table also confirms to us that the best position to measure the SP is at the centre of the grating, where the last two columns of the table show η when θ_p is measured at 25% length of grating. The table also provides us with a complete story of the performances of different polarisations; *y polarisation* is the best option for SPR measurement in a confined region follows by *r* and *x polarisations* respectively. Very accurate results compared with the uniform case can be obtained at around 9-11 microns periods of grating, which means the smallest microfluidic channel we can use is about 4.5-5.5 microns for the interferometric microscope which also minimizes the effect of crosstalk. This is, of course, is not an absolute guideline for designing a SPR sensor device as the sensitivity of the *V(z)* is dependent on refractive index contrast between the two channels and also the thickness of the sample and in addition what parameter that *V(z)* is sensitive to the change in the local SP coupling condition Δk-*vector* which in this case directly proportional to Δnd, where Δn is refractive index contrast of the two channel and d is the thickness of sample. The reason I modelled the material with $n=1.5$ in water ambient and 10 nm thick was because it is of the same order as biological membranes such as cell membranes and protein.

Chapter 6

SPR confocal microscopy

In this chapter, I will show that the complexity of the two arm heterodyne SPR microscope can be reduced using confocal microscope configuration and moreover that a similar *V(z)* measurement performance can be obtained from the confocal system. However, there are some issues which need to be addressed. Firstly, effects of the confocal pinhole size on the *V(z)* measurement and the ripple amplitude as the confocal will detect the interference signal in intensity rather than field; therefore the amplitude of ripples in the confocal will be smaller than the interferometric one. All the issues and some ways to get around them will be discussed in this chapter.

6.1 SPR Confocal microscopy theory

It has been shown in the previous chapters that a scanning heterodyne interferometric microscope with an oil immersion objective can be used for high resolution surface plasmon (SP) imaging. The essential idea is that when the sample is moved above the focal plane of the objective there are two major contributions to the output signal; one arising from the SP and the other arising from light directly reflected from the sample. As the sample is defocused the relative phase between these contributions changes leading to an oscillating signal whose period depends on the angle of incidence at which SPs are excited. Recently, this idea has been extended to a wide-field configuration (Somekh *et al* 2009).

The problem associated with the interferometric configuration is that it places severe demands on system stability and in the case of the heterodyne system acousto-optic modulators and associated electronics are required. In a scanning interferometer the interference signal is recovered from the product of the returning field in the back focal plane (BFP) and the reference beam so that this signal is proportional to the integral of the reflected field (assuming a uniform reference beam).

The two arm heterodyne microscopy setup is shown in figure 6.1. For the heterodyne system the reference beam is provided separately with the reflected beam. The idea to reduce the complexity of the heterodyne system is to provide the *V(z)* reference beam at the centre of the reflected so called ‘common path’ interference as shown in figure 6.2a.

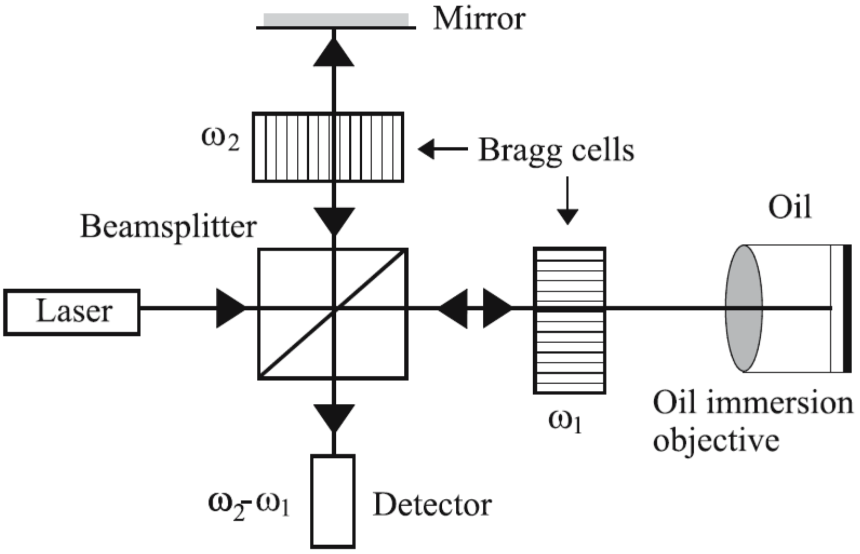
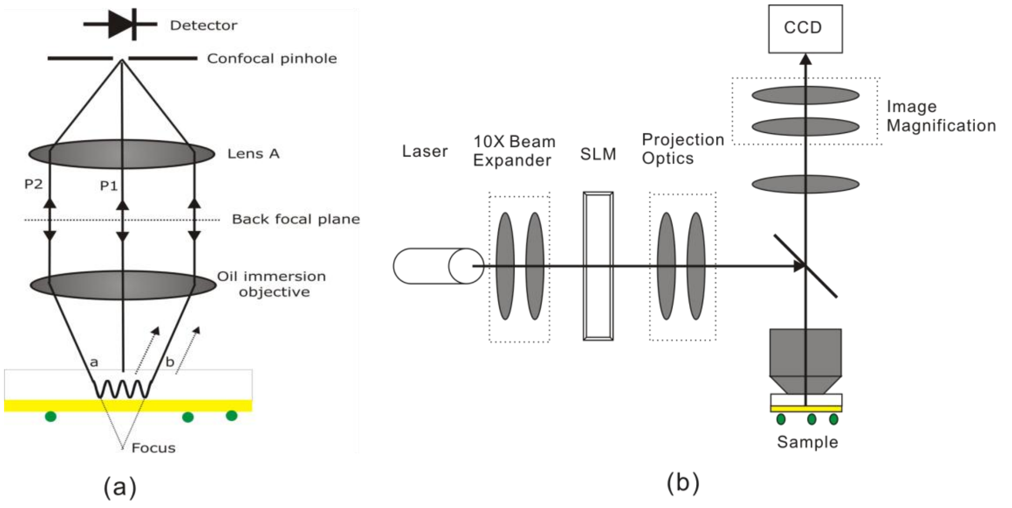


Figure 6.1 shows a scanning interferometric SPR microscope setup (Somekh 2007)

The common path interference configuration can be achieved by confocal setup as shown in figure 6.2b. When the sample is defocused the major contributions to the signal arise from the paths P1 and P2. Path P1 involves light incident and reflected close to normal incidence and path P2 comprises light incident at the angle for excitation of SPs, this generates SPs at position ‘a’; this couples back to light at all positions, however, in the confocal system only the light appearing to come from the focus (reradiating at ‘b’) passes through the pinhole. Of course, SPs excited at ‘b’ and reradiated at ‘a’ make a similar contribution.



(a) (b)

Figure 6.2 shows a) Conceptual diagram of SP confocal microscope; b) simplified schematic diagram of the experiment setup (Zhang *et al* 2012)

The other difference between the confocal and heterodyne systems is the magnitude of ripples detected at the output; the confocal system will have lower ripple amplitude as there is no interference beam for the reflected beam to interfere with. In other words, the confocal detects $I(z) = |V(z)|^2$ rather $|V(z)|$; this, of course, means that the background becomes large and, on the other hand, ripple become smaller relative to its background. It will be shown later in section 6.2.2 that this issue can be largely addressed by pupil function engineering. The purpose of pupil function engineering is to suppress light that does not contribute to the signal.

The other difference between the heterodyne system and confocal system is that for the heterodyne system we do not have to take effect of pinhole size in to account, whereas here for the confocal system the pinhole needs to be small for ideal performance as when the pinhole becomes wider SPs emitted from different positions are detected and we would expect the ripples to become less well defined. The effect of pinhole size will be discussed in section 6.2.1.

For the experimental setup and simulation parameters, the confocal system was setup as shown in figure 6.2b. A 632.8nm He-Ne laser (10mW) was used as the illumination source. Beam expanders were used to increase the beam diameter of the beam incident onto the spatial light modulator (SLM) which was conjugate with the BFP of the immersion objective. The SLM was used to control the effective illumination pattern in the BFP. A pellicle beam splitter was used to separate illumination and imaging paths. The microscope objective used in this study was a 1.25 NA oil immersion objective, which had sufficient aperture to excite SPs in air. The light distribution reflected from the sample was imaged onto the CCD camera which served as a variable pinhole. The light from the sample was magnified by approximately 1000 times from the sample to the CCD plane so that a point spread function occupied >100 pixels, this allowed the pinhole radius to be readily controlled by selecting different regions of camera, moreover, spreading the returning beam reduced the problem of saturating individual pixels. Samples were mounted in the 3-dimensional (3D) scanning system, which consisted of a 3axis mechanical stage drive, piezoelectric actuators (P621.1CD, Physik Instrumente) stage and (P-541.2CD, Physik Instrumente) giving 1 and 2 axes of movement respectively. The SLM, PZT stages and camera were controlled by software based on LABVIEW and all data were processed with MATLAB. Samples were prepared by coating gold with or without BSA on cover glasses; 2nm chromium was coated between gold and cover glasses to improve the adhesion.

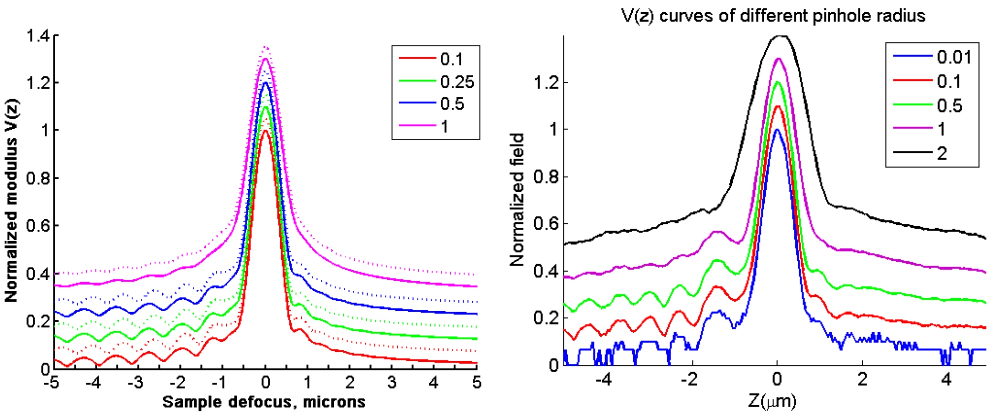
Acknowledgement: the experimental results presented in this chapter were carried out by my colleague Bei Zhang and published in Optics express 2012 (Zhang *et al* 2012).

6.2 Simulation and experimental results

6.2.1 Effect of pinhole size

In order to investigate the effects of pinhole size, we define the size of pinhole by radius of Airy disc radius (0.61λ/NA). In the experiment, the image plane images as a function of z defocus were captured and post-processed to calculate the intensity as a function of z defocus *I(z)* for different sizes of pinhole. This was slightly different from the simulations, where, for the simulation, bfp images for every z defocus position were calculated and Fourier transformed to calculate the Airy disc pattern. We can then calculate the total intensity for different pinhole sizes as a function of z defocus distance.

Experimental and simulation case: $NA=1.25$, $n_0=1.52$, n_1 was gold layer with 50 nm. Wavelength 633 nm *p-polarised* wave; there were cases investigated in this section firstly bare gold in air and secondly coated gold, where the gold is coated by a 10 nm of a refractive index of 1.5.



a) b)

Figure 6.3 shows *V(z)* curves for different pinhole diameter. Solid curve 50nm bare gold, dashed curves gold with 10nm overlayer with refractive index 1.5. Each pinhole diameter is displaced by 0.1 in the *y*-axis and overlayers are displaced by 0.05 in the *y*-axis. Pinhole radii are defined in terms of radius of Airy disc radius ($0.61\lambda/NA$) are shown in the legend (Zhang *et al* 2012)
a) simulation results and b) experimental results

Figure 6.3b shows experimental *V(z)* curves obtained from a coverslip that was coated with 50nm gold and 2nm chromium. For the large pinhole diameters the system has the same transfer function as a wide-field imaging system and the ripples are not observed; as can be seen in the simulations of figure 6.3a. For very small pinhole diameters the SNR is poor since little light is detected, so a suitable compromise between satisfactory confocal response and good SNR is, for pinhole size between 0.1 and 0.5 of the Airy disc; similar to the values used in most conventional confocal imaging experiments.

The figure also demonstrates that the *V(z)* effect is periodic with a period of 741nm close to the expected value of 752nm, however errors in film thickness and properties and also changes in the pupil function (Elezgaray *et al* 2010) account for this difference. A systematic error in absolute period does not prevent the measurement of very small changes, which is the primary purpose of most SP sensors. The figure also shows that the contrast is low and to this end we discuss the effect of pupil function engineering to mitigate this problem.

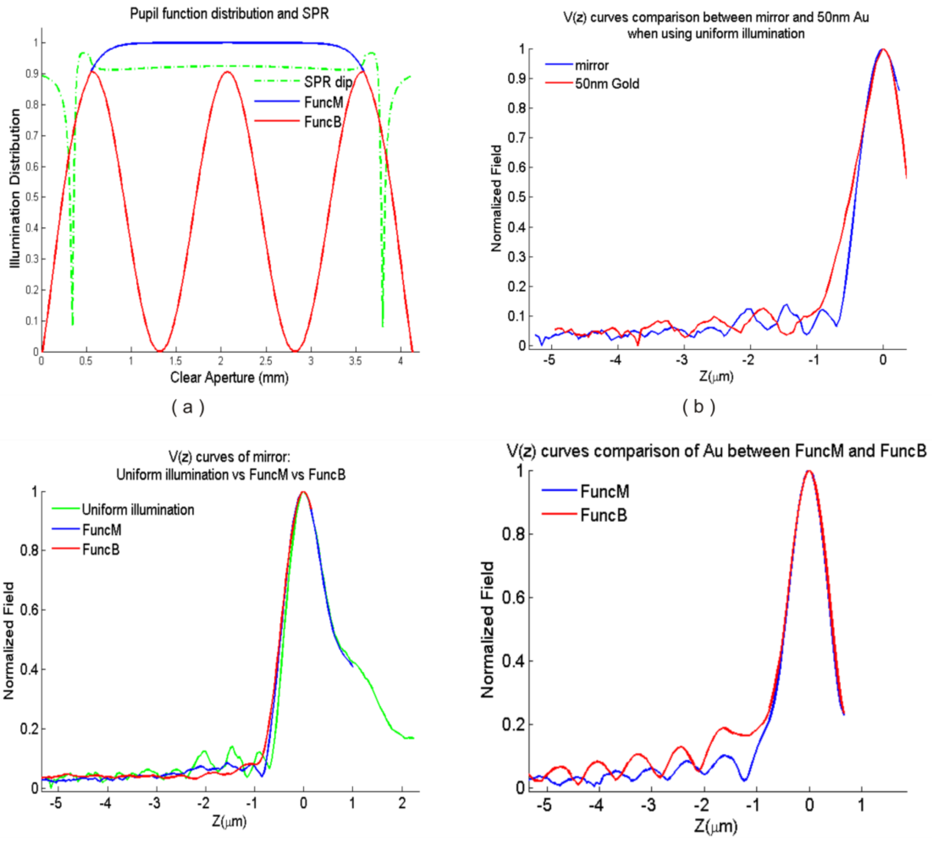
6.2.2 Increase the ripples magnitude by pupil function engineering

In this section, I will demonstrate that the ripples magnitude can be enhanced by pupil function engineering; where we try to reduce the effect of the spatial frequencies that do not contribute the interfere signal. This can be done by employing either an amplitude SLM or phase SLM to suppress the intensity of the unwanted frequencies. The pattern from the SLM was imaged onto the BFP of the objective to modulate the pupil function of the scanning system. In other words, the ripple contrast can be increased by attenuating some of the spatial frequencies as shown in figure 6.4.

For a confocal system, it is expected that even if the sample is just a pure mirror, *V(z)* curve still presents some oscillatory behaviour which arises from the edges of the pupil function (Gong & Hsu 1994). The experimental results in figure 6.4b show that the period is approximately 559 nm, and the strength of the signal is also about 15%, which means that when using uniform illumination, it is difficult to obtain a good quantification.

Both the theory and simulation results show that the oscillatory behaviour arises from the sharp edge of the pupil function. Clearly, the effect of this edge obscures the ripples due to SPs, however, it also affects the period of the ripple as determined by the equation 5.2 as explained in detail in the previous chapter.

A reference beam whose phase does not vary in a perfectly linear fashion will also lead to variations in the measured period when SPs are present; it will not affect, however, the detection of SPs just the absolute quantification. In order to overcome this problem and also to increase the relative size of the SP contribution, pupil function modulation by a SLM was used and two functions to smooth the patterns on the BFP are discussed, we name these FuncM and FuncB as shown in figure 6.4. The experimental results show that the two functions work well in the system and can be used to reduce the ripples due to the pupil function as shown in figure 6.4c. We can also see that reducing the background signal in the angular range between normal incidence and the region where SPs are excited greatly improves the SP contrast; this can be clearly seen from the comparison between FuncM and FuncB in figure 6.4d.



(a) (b)

(c) (d)

Figure 6.4 shows effects of different pupil functions on recovered *V(z)* for mirror and gold samples. (a) Pupil function distributions. Green dashed curve shows the *p-polarisation* reflection coefficient with respect to the aperture of the microscope object. This shows the SP dip relative to the objective aperture. The blue (FuncM) and red (FuncB) curves show the modified pupils to optimize contrast and reduce oscillations due to hard cut off in the lens aperture. (b) *V(z)* comparison between mirror and 50nm Au when using uniform illumination. (c) *V(z)* curves from a mirror using uniform, FuncM, and FuncB pupils. (d) *V(z)* curves of 50nm gold by FuncM and FuncB pupil functions. (Zhang *et al* 2012)

6.2.3 Imaging capabilities

In order to demonstrate that the technique can be used for spatially resolved measurements, a protein, bovine serum albumin (BSA) grating as shown in figure 6.5 was imaged. BSA is a typical biological protein and the effective refractive index of BSA on gold surface is approximately 1.4 (Beketov *et al* 1998). In this experiment, the BSA was fabricated on coverslips (refractive index 1.52) coated with 50nm gold with a period of 25 µm, with the BSA occupying 10 µm. The AFM measurement results show that the thickness of the grating was approximately 30nm.

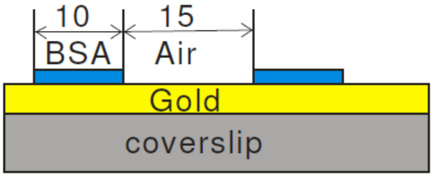


Figure 6.5 shows structure of BSA grating used in this study (Zhang *et al* 2012)

Figure 6.6 shows the 1D line traces of the BSA grating when the defocus distances are −1.25 µm (top left), −1.8 µm (top right) and at the focal plane (bottom left) respectively. It can be seen that the images at defocus −1.25 µm and at −1.8 µm are inverted, which can be easily explained by the bottom right figure, with different defocus (say position A and B), *V(z)* values of coated and uncoated grating are inverted. At the focal plane the grating contrast is extremely poor. It should be mentioned that at the focal plane, the intensity is much larger than the values when the sample is defocused. In order to get better contrast for the defocused values the light level was increased which saturated the signal at focus, which accounts for the saturation seen in Figure 6.6d, of course, Figure 6.6c was taken at the lower power level where the focal distribution was unsaturated.

Figure 6.7 shows the 1D images of the BSA grating using different pinhole radii. It can be seen that contrast varies with the pinhole radius. When the normalized pinhole radius is 2, that is essentially, the conventional scanning microscope, the shape of the grating almost vanishes. With smaller pinhole radius, like 0.8, the grating shape can be recognized but with lower contrast; while the radius shrinks to 0.1 or 0.01, the image can be obtained with better contrast. For the very small pinhole radius of 0.01 the SNR is poorer than 0.1 as expected although the values are still acceptable. We can see than confocal imaging gives greatly improved contrast and quantification. The polarisation direction of the illumination was along the grating fringes.

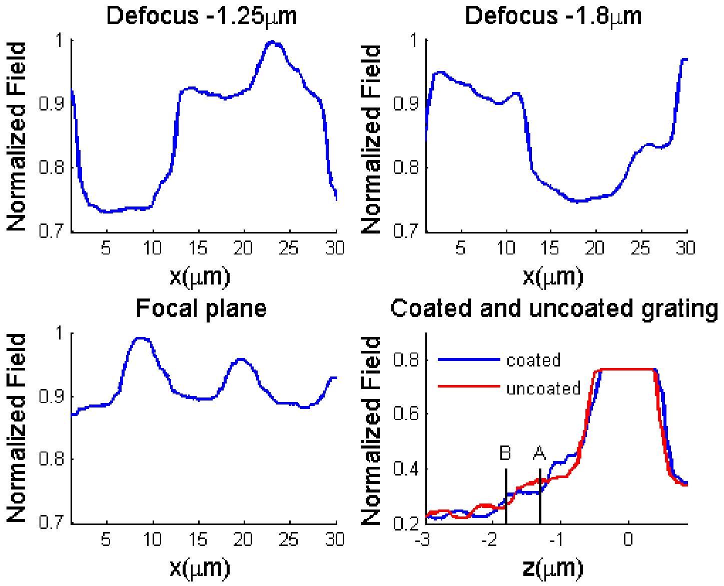


Figure 6.6 shows 1D grating images with different defocuses. Bottom right shows *V*(z) curves on coated and uncoated regions explaining contrast reversal. (Zhang *et al* 2012)

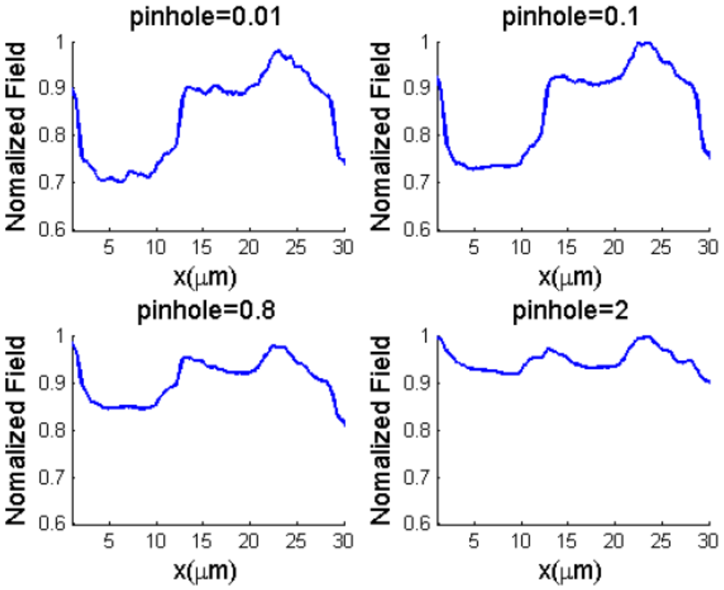


Figure 6.7 shows one dimensional grating images at −1.25μm defocus using different pinhole radii. (Zhang *et al* 2012)

6.2.4 Crosstalk measurement

In this section, I will show that the confocal system can give us a similar performance to the heterodyne system. This is done by simulating the relative difference in plasmonic angle η as a function of grating period as discussed in the previous chapter. The relative difference in plasmonic angle η was calculated for heterodyne system and confocal system with different pinhole sizes ranging from 0.1, 0.25, 0.5 and 1 of the airy disc radii for *x* and *y polarisations* as shown in figures 6.8 and 6.9. Note that the term *V(z)* measurement in figure caption of figure 6.8 and 6.9 means *V(z)* response from the scanning heterodyne system discussed in the previous chapter.

From the figures 6.8 and 6.9, we can see that the heterodyne and confocal systems gave us similar results; however the performance of the confocal depends on the pinhole size. For the pinhole size of 1, the ripple amplitude was very small and the ripple period Δz cannot be determined. In other words, for the confocal system as long as we can determine the ripple period Δz, the measurement performance would be similar to the heterodyne system, although the bigger pinhole size would degrade the accuracy of the SP measurement.

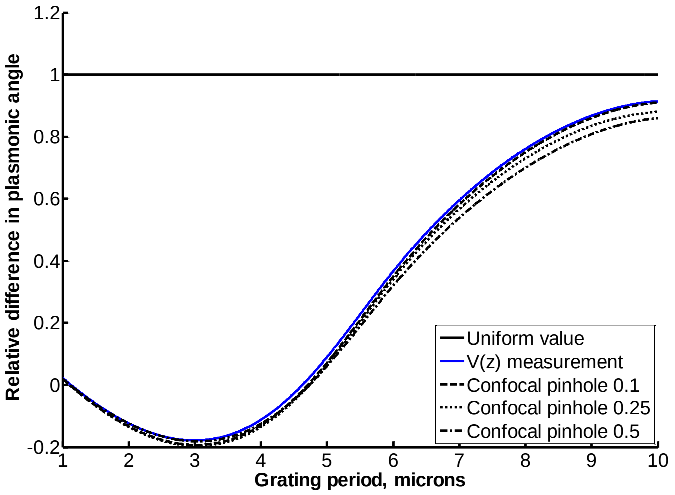


Figure 6.8 shows the relative difference in plasmonic angle η as a function of grating period for heterodyne and confocal system with different pinhole sizes; *x polarisation*

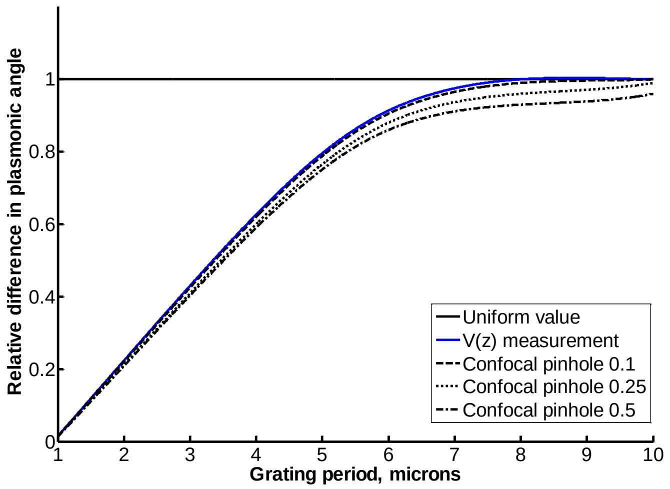


Figure 6.9 shows the relative difference in plasmonic angle n as a function of grating period for heterodyne and confocal system with different pinhole sizes; *y polarisation*

6.3 Discussion and conclusion

In this chapter, I have demonstrated the feasibility of using a confocal arrangement to perform localized measurement of SP propagation. We have examined the effect of pinhole diameter and also shown the effects of different illumination pupil functions. The method offers an alternative to interferometric SP imaging which is simpler and more stable, and perhaps the most significant advantage is that is can be incorporated into a conventional confocal microscope with minimal adaptation. The use of alternative input polarisation states will provide a useful extension of the technique.

It is interesting to note that not only the pinhole size affects the *V(z)* response, but also the pinhole position. In the future work section, I will show that the pinhole position can shift the position of the ripples. Also it is interesting to investigate the use of multiple pinholes; this might give us some rich information about the SP effect.

Chapter 7

SP imaging using a spatial light modulator

In this chapter we consider ways to extract the surface plasmon (SP) signal from the interference signal, so that we might gain some insight of physical SP behaviour in the interferometric system as well as looking at new ways to perform the SP measurements. More importantly, one of the main aims of extracting the SP signal is to be able to measure the SP without mechanical scanning.

In this chapter therefore, we will present a way to extract SP wave profile from the interference response obtained from the SP confocal microscope using a spatial light modulator (SLM). It will be shown that this calculation not only gives insight into the physical nature of the SP, but also provides a mean to improve spatial resolution.

We present two types of SP extraction algorithms, using either amplitude SLM (A-SLM) or phase SLM (P-SLM). Although we are of the firm view that the P-SLM gives us the better performance in terms of noise, the A-SLM was also studied as it can carry out the similar task at a lower cost, albeit with certain limitations as discussed later.

More importantly, we also discuss the imaging performance when using the amplitude and phase stepping algorithms to quantify the SP measurement. This shows an additional advantage of the method that they give less crosstalk from adjacent regions compared to z scanning of *V(z)* measurement.

The chapter begins with an explanation of the assumptions required and validation of the assumptions. Secondly, I will discuss amplitude stepping algorithms and we will show that they require a 100% fill factor. We then present a phase stepping algorithm, which will be shown that it does not suffer from fill factor problem and it is more robust to noise. We then demonstrate the main aim of this chapter by demonstrating that the algorithms enable us to quantify the crosstalk effect without having to measure the ripple period as discussed in chapter 5.

7.1 Assumption and validation

It is important to note that in this chapter, all the simulations were calculated with a pinhole size of 0.1 times the Airy disc size. It was shown in the previous chapter that as long as the pinhole is no bigger than 0.5 times radius of Airy disc; we should be able to get an accurate *V(z)* measurement performance.

As discussed in the chapter 5, the interference effect *V(z)* of the SPR can be depicted as a vectorial summation of two vector fields, namely the SPR field $|S|$ and its reference field $|R|$ as shown in figure 7.1. This is our main assumption, namely that the *V(z)* can be reconstructed by interfering the two vector beams.

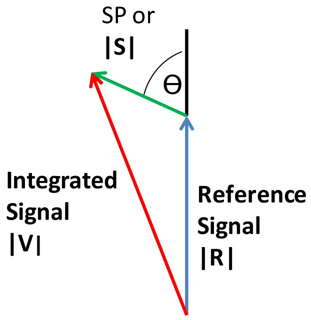


Figure 7.1 shows the vectorial sum of SP and reference beams; reference signal in blue, SP signal in green and integrated signal in red.

From figure 7.1, we can see that the interference term can be described as a vectorial sum using the cosine rule, which is given by:

$$|V|^2 = |R|^2 + |S|^2 + 2|R||S|\cos\theta \qquad [7.1]$$

All the SP interferometric microscopes have been so far studied through the term $|V|$ rather than directly extracting the term $|S|$ and its relative phase θ. Therefore the main aim of this chapter is to determine the SP amplitude $|S|$ and the relative angle θ.

It is important to note that in this chapter the symbol θ represents the relative angle between $|S|$ and $|R|$ vectors; and the symbol γ represents the angle of incidence. This is not to be confused with the symbol of the angle of incidence in the previous chapters.

Validation to show that V(z) can be approximated by $|S|$ and $|R|$ vectors

In this section, I will show that the *V(z)* curve can be reconstructed by equation [7.1]. This was done to test how well our assumption holds as we now assume that the complex field integration of the *V(z)* can be approximated by summation two vectors, which are $|S|$ and $|R|$.

Note the expression for the *V(z)* calculation is given by the complex field integration over the back focal plane, which can be expressed as:

$$V(z) = | \iint P_1(sin\gamma)P_2(sin\gamma)^2\left[R_p(sin\gamma)cos^2\emptyset + R_s(sin\gamma)sin^2\emptyset\right]e^{jn_0kcos\gamma z}\, dsin\gamma d\emptyset| \quad [7.2]$$

Where $P_1(sin\gamma)$ is the amplitude pupil function of the SLM and $P_2(sin\gamma)$ is the pupil function of the objective lens. R_p and R_s are complex electric fields of *p-polarised* and *s-polarised* waves respectively, ϕ is the azimuthal angle and γ is the angle of incidence. It is important to note that the pupil function $P_1(sin\gamma)$ was not squared as it is now controlled by the SLM and the light illuminated on the SLM only once.

Therefore instead of using equation [7.2] to calculate the *V(z)* curve, we want to see the effect of approximating the complex integral above by interference of two vectors, which is given by:

$$V(z) = \sqrt{|R(z)|^2 + |S(z)|^2 + 2|R(z)||S(z)|\cos[\theta(z)]} \qquad [7.3]$$

Where *R(z)* is the reference beam, *S(z)* is the amplitude of SP and *θ(z)* is the relative angle between $|R|$ and $|S|$ vectors as depicted in figure 7.1. Therefore in this section, before using the equation [7.3] to extract the SP profile we should determine how well we can approximate equation 7.2 with the equation 7.3. This can be done by comparing the *V(z)* curves calculated from the two equations to see how well agreed they are.

Pupil functions used in this calculation are shown in figure 7.2 with the following simulation case; *NA=1.65*, $n_0=1.78$ with a bare gold layer 50 nm in water ambient was illuminated with 633 nm incidence linear polarised wave.

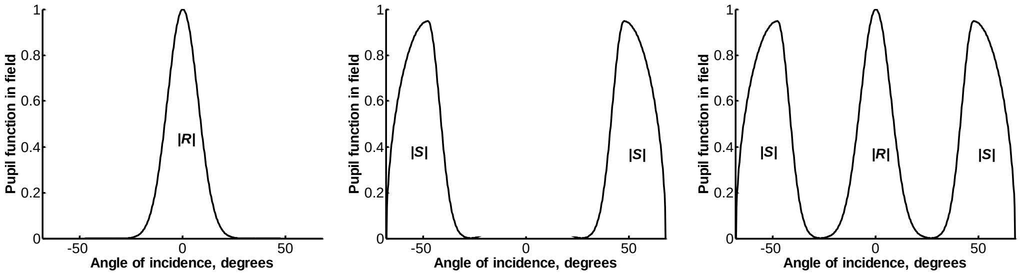


Figure 7.2 shows the pupil functions used in this section a) pupil function for determining |*R(z)*|, b) pupil function for determining |*S(z)*| and c) pupil function for determining *V(z)*

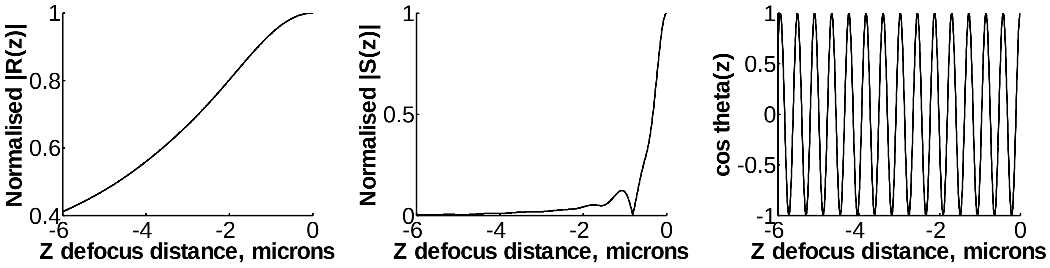
For the two vector approximation, we firstly determine the |*R(z)*| which can be calculated by the same way as determining the *V(z)* with the pupil function 7.2a using equation [7.2]. This |*R(z)*| as shown in figure 7.3a was carried out in the similar way as the full back focal plane integration in order to ensure that the two methods produce the similar DC background level over the z defocus distance. Similarly the |*S(z)*| term can be also calculated using equation [7.2] with the pupil function shown in figure 7.2b.

The last term that we have to work out is $\cos\theta(z)$, where it can be determined by the relative phase between two vectors, i.e. of $2kz$ and $2kzcos\theta_p$. In this case, the plasmonic angle θ_p was 53.78 degrees, where we calculated the plasmonic angle by Fourier transform calculation and determining the dominant frequency of ripples for the *V(z)* calculated from equation [7.2], which will be shown in figure 7.5 later. The $\cos\theta(z)$ calculated is shown in figure 7.3c. The equation [7.4] was derived based on the assumption that the *V(z)* ripples were formed by phase contributions only from $|R|$

and $|S|$. The central point of the bfp has the phase shift varying as *2kz* and the SP contribution has the phase shift varying as $2kz\cos\theta_p$.

$$\cos\theta(z) = Real(\exp(i\delta - i2kz(1 - \cos\theta_p))) \quad [7.4]$$

Where δ is a phase factor to match the phasor model with the calculated *V(z)* from equation [7.2]. In other words, the δ term accounts for the phase difference in the two *V(z)* calculations.



a) b) c)

Figure 7.3 shows a) $|R(z)|$ b) $|S(z)|$ and c) $\cos\theta(z)$

From the figure 7.3b, we can see that the $|S(z)|$ has two lobes; these are a combined effect of plasmonic and non-plasmonic reflections from the sample as described in Velinov *et al* and chapter 2. Velinov *et al* (Velinov et al 1999) captured an SP propagation image shown in figure 2.5, where they reported a similar shape to figure 7.3b.

It is also interesting to see whether the discrete pupil function as shown in figure 7.2c can give us a similar measurement performance compared to the continuous pupil function, such as the one used in chapter 5. The continuous pupil function used to compare the *V(z)* measurement performance in this section is shown in figure 7.4. The *V(z)* curve calculated using the pupil function in figure 7.4 by calculating the full back focal plane integration equation [7.2] is shown figure 7.5 (in red).

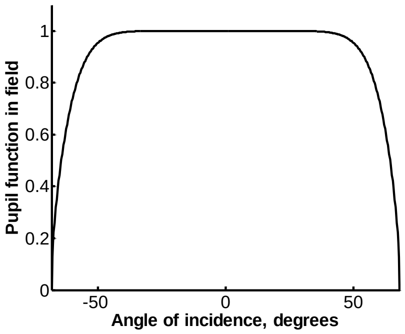


Figure 7.4 shows continuous pupil function

From the results in figure 7.3, we can then work out the *V(z)* curve by substituting the results in the figure into the equation [7.3], we can then get the *V(z)* curve as shown in figure 7.5 (solid green curve). The figure also shows the *V(z)* calculated using full back focal plane integration in equation [7.2] as shown in figure 7.5 (solid blue curve).

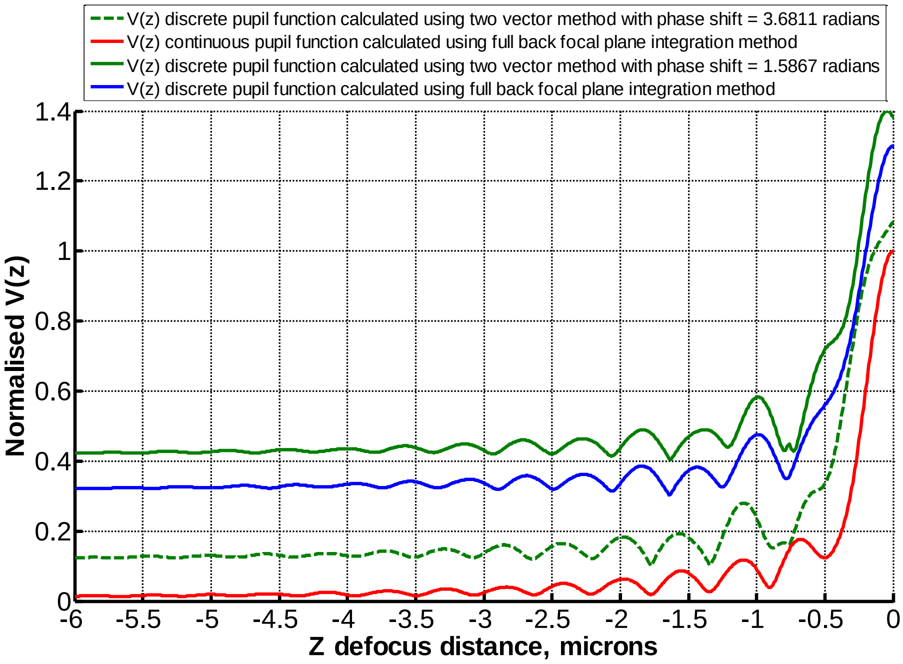


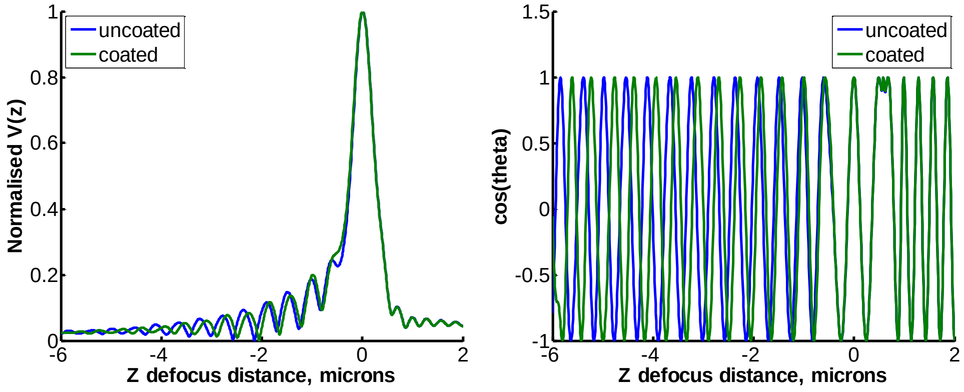
Figure 7.5 shows *V(z)* curves calculated using
1) (Solid green) two vector approximation in equation [7.3] using discrete pupil function shown in figure7.2c with phase factor of 1.586 radians to match with the ripple position on the blue curve
2) (Blue) full bfp integration method in equation [7.2] using discrete pupil function shown in figure7.2c
3) (Dashed green) two vector approximation in equation [7.3] using discrete pupil function shown in figure7.2c with phase factor of 3.6811 radians to match with the ripple position on the red curve
4) (Red) full bfp integration method in equation [7.2] using continuous pupil function shown in figure7.4

The first comparison made here is a comparison between the two vector approximation and full back focal plane integration in order to determine whether the two vector method can be used to model the full back focal plane integration. From the figure 7.5, solid green *V(z)* and solid blue *V(z)* curves (the first two upper lines) tell us that the two vector method can give us similar *V(z)* response to the *V(z)* response calculated using the full back focal plane integration, although there are still some discrepancies around 0 micron to -0.75 microns z defocus range. This is as expected as the two vector approximation will not apply close to focus. The analysis above quantifies when this approximation is valid and also allows one to assess the errors associate with using two vectors to represent the curves.

The second comparison made here is a comparison between the *V(z)* responses obtained from the discrete pupil function and continuous pupil function. The figure (see the lower two *V(z)* curves in figure 7.5) also tells us that the continuous pupil function (solid red curve) and the discrete pupil function (dashed green) can give us a similar measurement performance, although there were some discrepancies; firstly in the DC level and secondly the shape of the first ripple were different; this was due to the use of different pupil function. The two vector approximation can give us even better approximation when we reduce the middle like the discrete pupil function.

The next question is to investigate whether the two vector approximation responds to changes in refractive index in the same way as the full back focal plane integration as having discussed that the *V(z)* curves are sensitive to the refractive under the axis of the beam. This is investigated by using uncoated and coated cases as explained in chapter 5. The uncoated case is provided by simulating the following parameters: 46 nm of gold in water ambient; $NA=1.65$ with $n_0=1.78$ and $\lambda=633\ nm$ linear polarised wave. For the coated case, a uniform layer of dielectric with 10 nm thick uniform sample $n=1.5$ is deposited on the gold layer. In order to calculate the two vector approximation.

The *V(z)* calculations for the coated and uncoated for the full back focal plane integration method were performed using pupil function shown in figure 7.2c. On the other hand, the θ_p for the two vector approximation were calculated from firstly determining $|R|$ and $|S|$ using pupil functions in figure 7.2a and 7.2b. After determining $|R|$ and $|S|$ we can then substitute these $|S|$ and |R| into equation [7.3] and work out Δz by measuring the period of the $cos\theta$ function; we can then determine θ_p using equation [5.2].



a) b)

Figure 7.6 shows a) *V(z)* curves calculated for (Blue) bare gold case and (Green) coated gold using the discrete pupil function with full back focal plane integration and b) *cosθ(z)* curves calculated for (Blue) bare gold case and (Green) coated gold using the discrete pupil function and two vector approximation.

From figure 7.6, we can then measure the period of the ripples from the responses in figures 7.6a and 7.6b, we can then determine the difference in plasmonic angles $\Delta\theta_p$ of the two uniform cases. For the full back focal plane integration shown in figure 7.6a, the plasmonic angles θ_p for coated and uncoated cases were 52.82 and 54.32 degrees respectively, which correspond to the difference of 1.50 degrees. Similarly, for the two vector approximation shown in figure 7.6b, the plasmonic angles θ_p for coated and uncoated cases were 52.82 and 54.33 degrees respectively, which correspond to the difference of 1.51 degrees. Note that the period of the ripple was determined by determining the mean value of the first six ripple periods. These results can provide us with strong evidence that the two vector approximation has a similar sensitivity to refractive index change.

To sum up, this enables us to conclude that the two vector approximation can give us a good approximation to the full back focal plane integration method, although there are still some discrepancies over the z defocus distance. The key conditions that we have to satisfy are to ensure that the discrete pupil function is designed to cover the plasmonic angle position and it is smooth enough to provide a steady reference signal as discussed in chapter 5.

7.2 Amplitude SLM algorithms

Having validated our assumption in the previous section that the *V(z)* curve can be accurately approximated by two vectors under certain circumstances, let us now begin with some ideas to extract some SP propagation parameters using firstly amplitude SLM (A-SLM). The A-SLM algorithms suggested in this section can be categorised into two types, which are direct and indirect methods.

7.2.1 A-SLM direct method

As shown in the validation section that by using such pupil functions shown in figure 7.2. We can directly work out $|S(z)|$, $|R(z)|$ and $\cos\theta(z)$ by firstly using the pupil function shown in figure 7.2a to determine $|R(z)|$; similarly by using the pupil function shown in figure 7.2b, we can then work out $|S(z)|$. Once the $|S(z)|$ and $|R(z)|$ are determined, we can then substitute them back into the equation [7.3], where the *V(z)* term in the equation can be obtained from performing a *V(z)* calculation using the pupil function shown in figure 7.2c. Therefore we can then determine the $\cos\theta(z)$ term.

We can see that the A-SLM direct method is quite straightforward; however the method does depend on 100% fill factor of the A-SLM as the method requires the SLM to be able to either suppress the central part or the side band part of the pupil function completely. If the A-SLM does not have a 100% fill factor, this makes the method fail.

7.2.2 A-SLM indirect method

By employing the amplitude SLM to the system, we can now change the amplitude of the $|S|$ without changing θ and $|V|$ as shown in the figure 7.7 below.

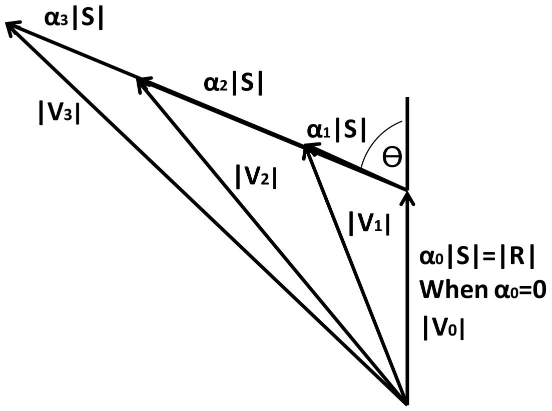


Figure 7.7 shows amplitude stepping can be done by varying α_n

The mathematical representation of the above diagram is given by:

$$|V_n|^2 = |R|^2 + \alpha_n^2|S|^2 + 2\alpha_n|R||S|\cos\theta \qquad [7.5]$$

Where α_n is the SLM amplitude at the SPR frequency range.

We now have three unknowns that we need to solve for which are $|R|$, $|S|$ and θ respectively.

Let $\alpha_0 = 0 \rightarrow |V_0| = |R|$ [7.6] here the first unknown |R| is now solved.

We need at least another 2 equations to solve the remaining unknowns.

Hence we can have the following equations by doing two more amplitude steps on the SP angle:

$$|V_1|^2 = |R|^2 + \alpha_1^2|S|^2 + 2\alpha_1|R||S|\cos\theta \qquad [7.7]$$

$$|V_2|^2 = |R|^2 + \alpha_2^2|S|^2 + 2\alpha_2|R||S|\cos\theta \qquad [7.8]$$

Unfortunately, the unknowns are not linearly independent, so we cannot set a matrix system to solve them.

As $\cos\theta \in [-1,1]$ so we can rewrite the equations [7.7] and [7.8] in terms of two polynomial equations. Therefore we can plot two curves of the polynomial equations $|S|$ as a function of $\cos\theta$; the intersection point of the two curves is the point that can satisfy the equations [7.7] and [7.8] as shown in figure 7.8. In other words, it is the solution to the equations [7.7] and [7.8].

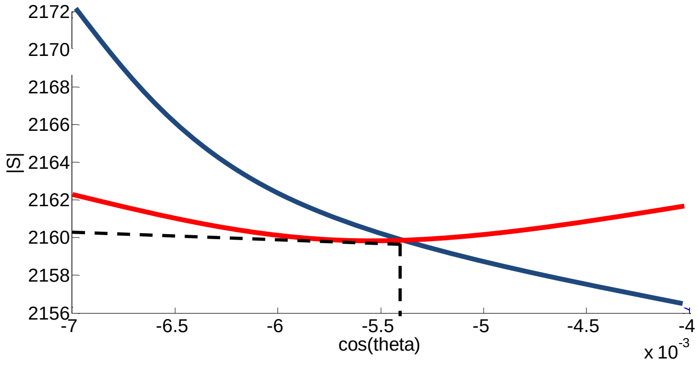


Figure 7.8 shows intersection point of two root locus lines

All the amplitude SLM methods presented cannot be used to solve for θ since we have the ambiguity due to the cosine function. The technique can partially solve the system of equations above by telling us $|R|$, $|S|$ and $cos\theta$ not the relative angle θ. Therefore the main aim of this amplitude SLM is to solve |S| and $cos\theta$ for any particular pupil function.

7.2.3 Amplitude SLM algorithm: criteria to design an amplitude pupil function

For the A-SLM methods, it is required by the methods to set the amplitude pupil functions by the following criteria.

1. Design the pupil function so that it can give us distinguishable responses between reference and SP signals which enable us to vary the amplitude of the SP waves without changing the reference signal.

2. The pupil function has to be smooth enough otherwise it could induce spurious oscillation in the integrated signal |*V*| (described in chapter 5).
3. The sidelobes of the discrete pupil function will be modulated by using constant multiplying factor α_n amplitude steps and we also have to ensure that summation of the |*S*| for all amplitude steps has the same phase of resultant |*S*|. In other words, we have to ensure that only the amplitude of |*S*| is modulated without altering the phase of |*S*| as can be described as equations [7.9] and [7.10] below.

   From figure 7.7, the |*S*| can be written as

   $$S = |S|e^{i\theta} = \alpha_{01}|S_1|e^{i\theta_1} + \alpha_{02}|S_2|e^{i\theta_2} + \alpha_{03}|S_3|e^{i\theta_3} + \cdots \qquad [7.9]$$

   Where α_{0i} is the amplitude of the A-SLM.

   $$\alpha_n S = \alpha_n|S|e^{i\theta} = \alpha_n\left(\alpha_{01}|S_1|e^{i\theta_1} + \alpha_{02}|S_2|e^{i\theta_2} + \alpha_{03}|S_3|e^{i\theta_3} + \cdots\right) \; [7.10]$$

   Where α_n is the scalar value used to modulate the amplitude of the A-SLM.

   We can see clearly that [7.9] and [7.10] have the same relative phase.

7.2.3 Amplitude SLM algorithm simulation results

Let us assume that we have 100% fill factor A-SLM and a noiseless system, we can now look at some results obtained from the algorithm in order to understand what we can obtain by employing such a technique. Note that for the noiseless case, both of the direct and indirect A-SLM algorithms give us the same results.

It is important to subtract the non-plasmonic effect from the plasmonic effect for the amplitude stepping methods presented in this chapter, otherwise we will get the |*S*| response that is contributed from both non-plasmonic and plasmonic effects.

In order to illustrate this point let us consider the following simulation cases. The simulation results in this section were calculated using the set of pupil functions shown in figure 7.9. The pupil functions were designed so that they satisfied the conditions mentioned in the previous section.

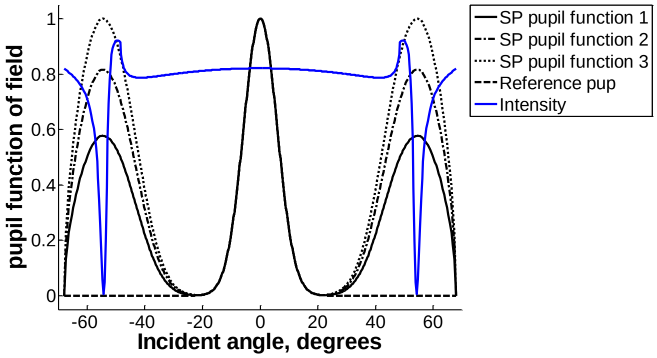


Figure 7.9 shows intensity pupil functions used in the simulation and the back focal plane distribution for *p-polarisation* (in blue); pupil functions 1 to 3 are for modulating the amplitude the SP with different coefficients; reference pup is for solving the reference term |*R*|.

Simulation case for non-plasmonic case: 4 μm of gold (a thick layer gold) in water ambient; $NA=1.65$ with $n_0=1.78$ and $\lambda=633$ *nm* linear polarised wave with 10 nm thick uniform sample $n=1.5$. The *V(z)* curves for the non-plasmonic case are shown in figure 7.10 and the |*S(z)*| calculated using indirect A-SLM method is shown in figure 7.12a.

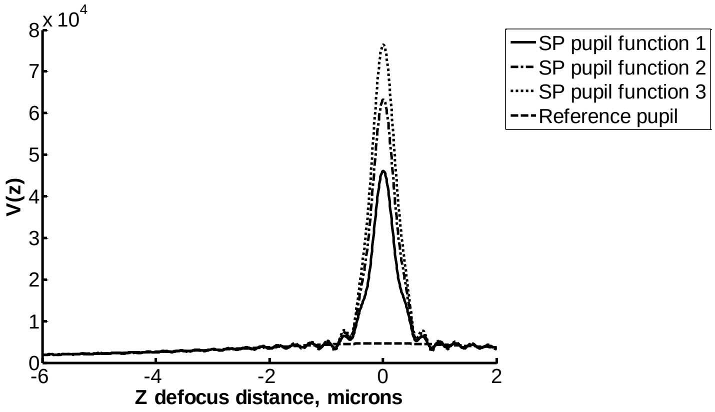


Figure 7.10 shows *V(z)* responses for non-plasmonic case calculated using the pupil functions in figure 7.9

Simulation case for plasmonic case: all the simulation parameters were the same as above case except that the gold layer was replaced by 46 nm of gold. The *V(z)* curves for the plasmonic case are shown in figure 7.11 and the |*S(z)*| calculated using indirect A-SLM method is shown in figure 7.12a.

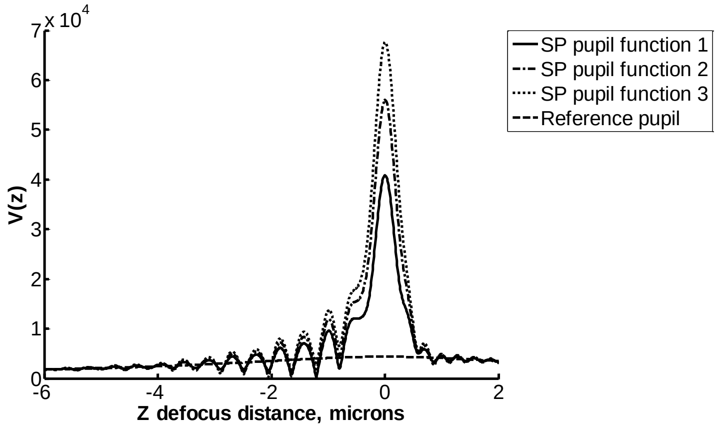


Figure 7.11 shows *V(z)* responses for plasmonic case calculated using the pupil functions in figure 7.9

From the figure 7.12a, we can see that |*S(z)*| curves for the plasmonic and non-plasmonic have a common peak amplitude centred at *z=0* micron. As explained in chapter 5 that, because of the optical scanning configuration we will not expect the system to capture any SP at *z=0* micron and positive defocus. Therefore, it is necessary to subtract the non-plasmonic |*S(z)*| from the plasmonic |*S(z)*| in order to get the effect only coming from the SPs as shown in figure 7.12b.

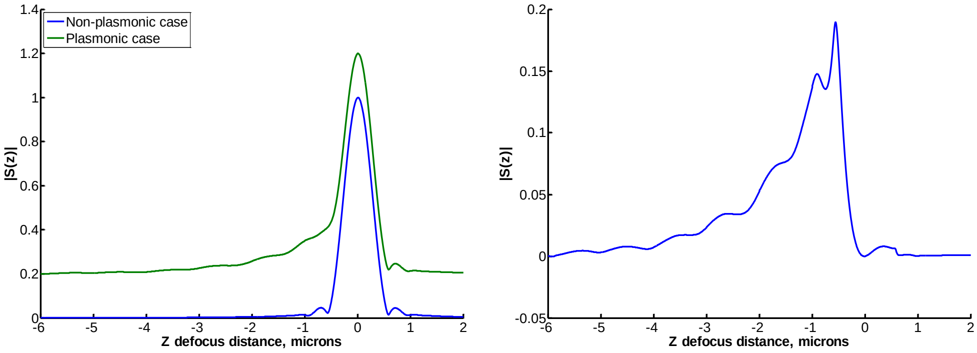


Figure 7.12 shows a) |*S(z)*| calculated from the *V(z)* curves in figures 7.10 (non-plasmonic case) and 7.11 (plasmonic case) and b) |*S(z)*| after subtracting the non-plasmonic effect from the plasmonic effect

From the |*S(z)*| curve shown in figure 7.12b, we can then work out *cosθ(z)* by substituting the result |*S(z)*| back into either equation [7.7] or [7.8]. The *cosθ(z)* curve obtained by using the |*S(z)*| results is shown in figure 7.13.

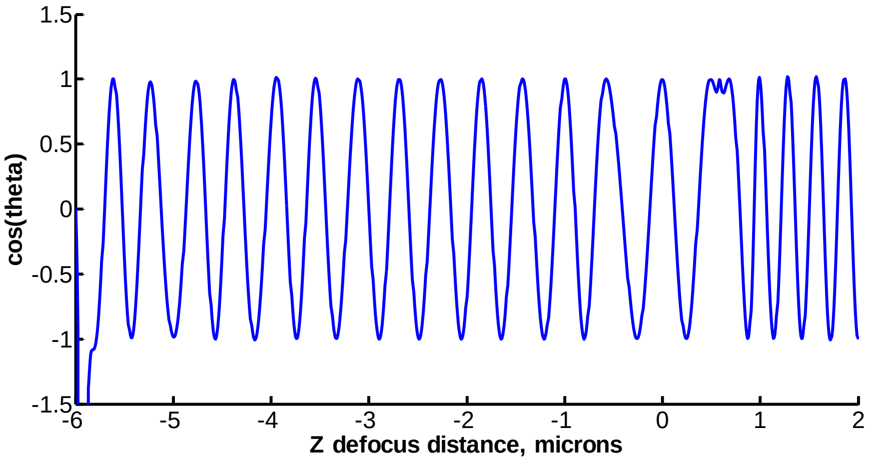


Figure 7.13 shows *cosθ(z)* calculated from the |*S(z)*| shown in figure 7.12b

The next interesting task is to measure the period of the cosine function in figure 7.13, this was done by taking Fourier transform to measure Δz and then calculate θ_p using the equation [5.2] as discussed in chapter 5. The value of the θ_p obtained was 54.75 degrees, which agreed well with the plasmonic angle calculated using 2 consecutive peaks measurement (54.90 degrees) discussed in chapter 5.

As explained earlier in chapter 2 that the propagation length of the SPs is proportional to $\frac{d\emptyset}{dk_x}$ at the plasmonic angle, where ϕ is the phase of the complex reflection coefficient as a function of k_x. Therefore it is interesting to see whether the propagation length of the SPs measured by this method is reasonable. The $\frac{d\emptyset}{dk_x}$ term for the 46 nm of gold is 7.32 microns. We can then calculate the SP exponential equation by applying a low pass filter to eliminate the oscillation on the |*S(z)*| curve in figure 7.12b; we then take natural logarithm to the filtered results and work out the slope of the log scale, which is attenuation factor of the SP, σ. We can then rewrite the exponential decay term in form of $y = e^{\sigma x}$, where x is the propagation distance. It is important to note here that the |*S(z)*| is presented in term of magnitude of electric field as a function of z defocus position. Therefore we will need to convert the z defocus axis to propagation x axis by using equation [5.7] described in chapter 5.

The exponential equations obtained for 46 nm of gold is $y = e^{-0.1421x}$. We can then calculate the actual decay length by working out the position that gives *y=1/e*. The propagation length is 7.04 microns, which agreed well with the $\frac{d\emptyset}{dk_x}$ term (7.32 microns).

7.2.4 Problem in A-SLM algorithm: Is a 100% fill factor required?

Amplitude SLMs normally come with fill factor less than 100%, this means when we set the A-SLM to zero it does not give us a perfectly switched off pixel, but instead it will act as a weak mirror. This issue will, of course, become problematic as we cannot set the SPs fields to completely zero level and solve for $|R|$ as described in the equation [7.6]. Therefore in this section, I will present a new algorithm that will improve the stability of $|S(z)|$ measurement even at the presence of background of the SLM. Let us assume that all intensity pupil functions have the background intensity of β (treated as an unknown in this calculation) as shown in figure 7.14. There are four unknowns in this problem, which are β, $|S|$, $|R|$ and $\cos\theta$.

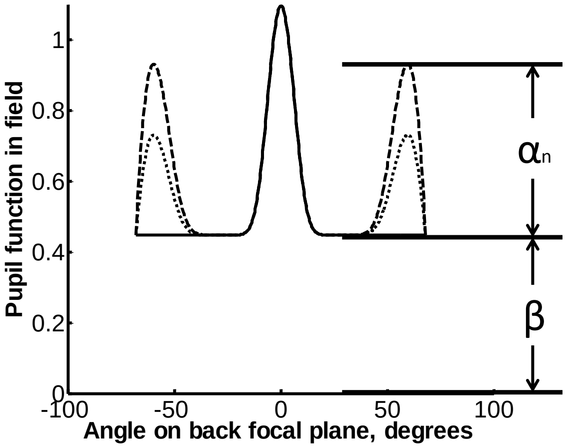


Figure 7.14 shows pupil functions with background

Therefore we can then rewrite the vector summation equations [7.6] to [7.8] based on the figure 7.14 by:

$$|V_0|^2 = (\beta + 1)^2|R|^2 + (\beta + \alpha_0)^2|S|^2 + 2(\beta + 1)(\beta + \alpha_0)|R||S|\cos\,\theta \text{ [7.11]}$$

$$|V_1|^2 = (\beta + 1)^2|R|^2 + (\beta + \alpha_1)^2|S|^2 + 2(\beta + 1)(\beta + \alpha_1)|R||S|\cos\,\theta \ [7.12]$$

$$|V_2|^2 = (\beta + 1)^2|R|^2 + (\beta + \alpha_2)^2|S|^2 + 2(\beta + 1)(\beta + \alpha_2)|R||S|\cos\,\theta \ [7.13]$$

The case that we let $\alpha_0 = 0$, we now cannot solve the term $|R|$ independently, so we have to solve this by using difference square rules. This can be expressed by:

[7.12]-[7.11]: $|V_1|^2 - |V_0|^2 = \alpha_1|S|(\alpha_1|S| + 2\beta|R|cos\theta + 2cos\theta|R|)$ [7.14]

[7.13]-[7.11]: $|V_2|^2 - |V_0|^2 = \alpha_2|S|(\alpha_2|S| + 2\beta|R|cos\theta + 2cos\theta|R|)$ [7.15]

We can then solve the term |S| by eliminating the other variables other than $|S|$ by:

α_1[7.39]- α_2[7.38]: $\alpha_1(|V_2|^2 - |V_0|^2) - \alpha_2(|V_1|^2 - |V_0|^2) = \alpha_1\alpha_2(\alpha_1 - \alpha_2)|S|^2$ [7.16]

From [7.16], we can then solve the term $|S|$ by:

$$|S| = \sqrt{\frac{\alpha_1(|C_2|^2 - |C_0|^2) - \alpha_2(|C_1|^2 - |C_0|^2)}{\alpha_1\alpha_2(\alpha_1 - \alpha_2)}} \qquad [7.17]$$

In order to calculate the *cosθ(z)* term, we require more equations which, of course, require more sophisticated algorithm and amplitude steps. However in this section we are not intending to solve the full set of solution, but instead to show the feasibility of solving the SP profile with less than a 100% fill factor.

The *V(z)* curves simulated for 46 nm of gold in water ambient with *NA=1.65* and n_0*=1.78* linear polarised wave is shown in figure 7.15, the SLM had 20% background intensity level. The figure shows that even for the reference pupil function ($\alpha_0 = 0$) the *V(z)* curve had a strong ripple pattern and a high DC background level. This was due to the fact that the background level was quite high at every point on the back focal plane including the SP angle.

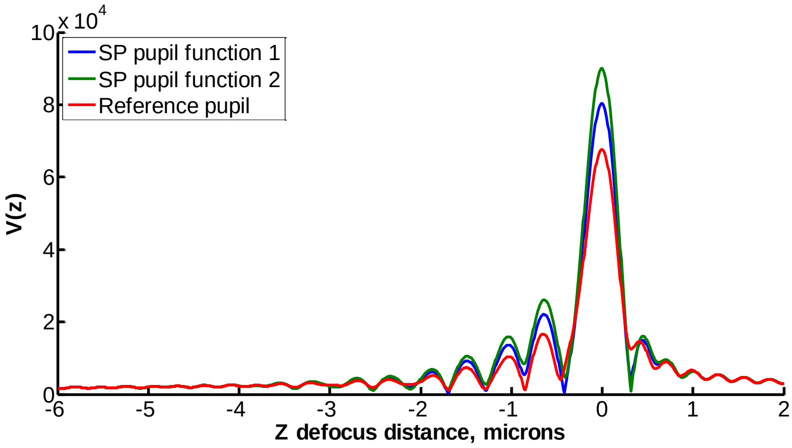
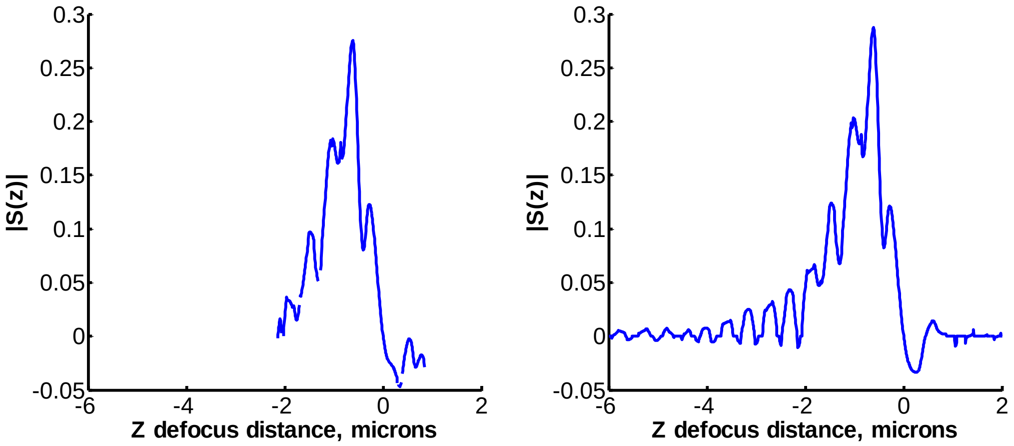


Figure 7.15 shows *V(z)* curves calculated with 20% background intensity level

From the results in figure 7.16b, we can see that the majority of $|S(z)|$ was successfully reconstructed using the new algorithm presented in this section, however, there were still some regions that cannot be reconstructed and some of the reconstructed points seemed to be unstable. This was due to the fact that the new algorithm was derived by assuming that vectorial summation of the interference signal only comes from the central path of the back focal plane and the maximum SP pupil function on the back focal plane by introducing the new amplitude terms as $(\beta+1)$ and $(\beta+\alpha_n)$ respectively, whereas the actual scenario is that all the points on the back focal plane contribute to the *V(z)* interference signal as seen by comparing the discrete pupil function and continuous pupil function as explained earlier. In other words, the two vector approximation was not hold. Therefore it is either required to have a more sophisticated algorithm or a 100% fill factor SLM. Many companies, such as BNS, Hamamatsu have been developing very high fill factor amplitude SLMs by embedding the electronics at the back of each SLM pixels; this possibly makes the fill factor goes up to 95%-100%.



a) b)

Figure 7.16 shows |*S(z)*| reconstructed using a) indirect method and b) the method described in this section.

Let us look at some experimental results; we will show the feasibility of achieving the SP profile extraction in experiment, when there is a presence of A-SLM background.

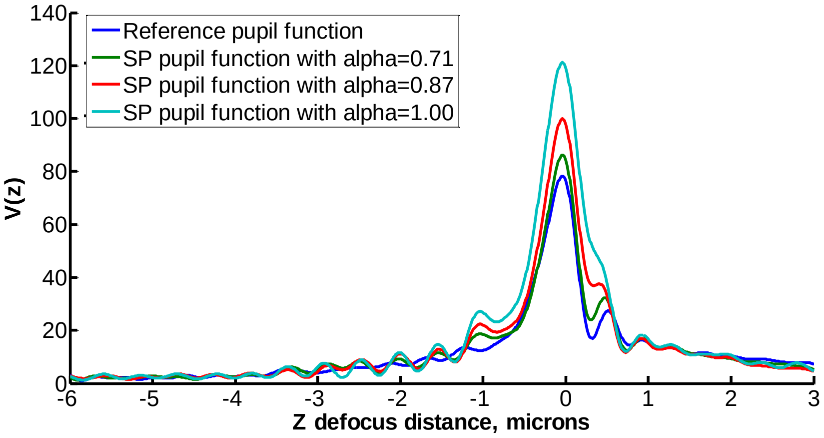


Figure 7.17 shows experimental *V(z)* for 50 nm of gold $NA=1.25$ with $n_0=1.52$ p-polarised wave in air using the set of pupil functions described in the figure 7.14

Figure 7.17 shows us that the experimental results obtained suffered from the background intensity level of the SLM, we checked that by capturing the back focal plane image when the SLM was off, however, we can still see some background light intensity with the plasmonic dip on the back focal plane. Therefore, we will use the algorithm described in the section to calculate the |*S(z)*| curve using equation [7.17]. The |*S(z)*| calculated from the experimental results shown in figure 7.17 is presented in figure 7.18.

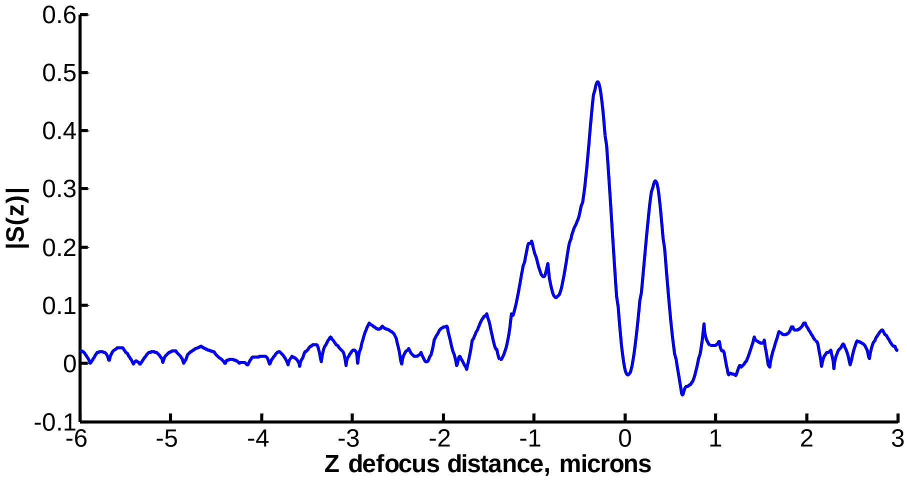


Figure 7.18 shows experimental |*S(z)*| for 50 nm of gold *NA=1.25* with $n_0=1.52$ p-polarised wave in air calculated from the experimental results in figure 7.17

We can see that the majority of the |*S(z)*| points can be reconstructed, which agreed with simulation results. It is also interesting to calculate the SP propagation to see whether it is agreed with theoretical calculation. The theoretical propagation length for 50 nm of gold in air is 17 microns calculated using equation [2.10] described in chapter 2. By filtering the result shown in figure 7.17 and taking a natural log to the filtered |*S(z)*|, we can then calculate the exponential decay function as $y = e^{-0.0595x}$ and the propagation length corresponding to the equation is 16.83 microns, which agreed with the value calculated using equation [2.10].

In conclusion, the amplitude algorithms do require a 100% fill factor of A-SLM and it cannot solve relative phase θ as it suffers from the ambiguity of cosine function. Therefore in the next section I will show that phase modulation technique can be used to get around these problems.

7.3 Phase SLM algorithm

Similar to the amplitude SLM algorithms, the pupil function is now modulated using the phase pupil function this means the amplitude pupil function is now fixed, but the relative phase θ on the back focal plane is modulated. Therefore the amplitudes of $|S|$ and $|R|$ are the same with changing relative phase θ between the $|S|$ and $|R|$. In order to develop a simple algorithm to reconstruct the $|S|$ and relative phase θ, we use the π/2 radian phase step. The phase stepping technique can be illustrated by the phasor diagram shown in figure 7.19.

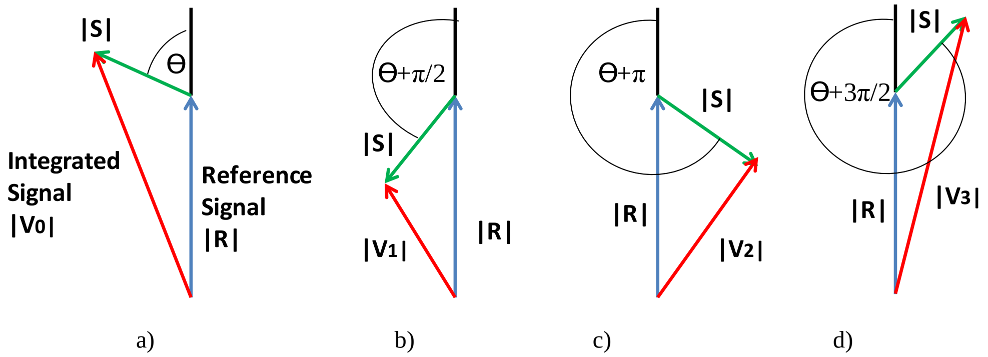


Figure 7.19 shows phasor diagram of the two vector approximation for phase stepping algorithm a) 0 degree phase shift b) 90 degree phase shift c) 180 degree phase shift and d) 270 degree phase shift

The diagram can be expressed by:

$$|V_n|^2 = |R|^2 + |S|^2 + 2|R||S|\cos\left(\theta + (n-1)\frac{\pi}{2}\right) \qquad [7.18]$$

Where $|R|$ is the reference beam, $|S|$ is the SP beam, V_n is the *n^th^* *V(z)* curve corresponding to the amount of phase shift described in the equation [7.18].

In this system we have three unknowns, which are $|S|$, $|R|$ and θ, we then need at least four equations to solve the system as θ needs one more equation to avoid its ambiguity.

$$|V_0|^2 = |R|^2 + |S|^2 + 2|R||S|\cos(\theta) \qquad [7.19]$$

$$|V_1|^2 = |R|^2 + |S|^2 + 2|R||S|\cos\left(\theta + \frac{\pi}{2}\right) \qquad [7.20]$$

$$|V_2|^2 = |R|^2 + |S|^2 + 2|R||S|\cos(\theta + \pi) \qquad [7.21]$$

$$|V_3|^2 = |R|^2 + |S|^2 + 2|R||S|\cos\left(\theta + \frac{3\pi}{2}\right) \qquad [7.22]$$

From the above standard equations of phase stepping algorithm, we can then derive number of unknowns in the system by:

$$[7.19]\text{-}[7.21]: |V_0|^2 - |V_2|^2 = 4|R||S|cos\theta \qquad [7.23]$$

[7.22]-[7.20]: $|V_3|^2 - |V_1|^2 = 4|R||S|sin\theta$ [7.24]

Therefore we can solve for the value of θ by [7.23]/[7.24]:

$\theta = tan^{-1}\left(\frac{|V_3|^2-|V_1|^2}{|V_0|^2-|V_2|^2}\right)$ [7.25]

Once the θ term is obtained, we can then solve for $|R||S|$ term by substituting the θ term back into either the equation [7.23] and [7.24].

$|R||S| = \frac{|V_0|^2-|V_2|^2}{4cos\theta} = \frac{|V_3|^2-|V_1|^2}{4sin\theta}$ [7.26]

We can then form another two sets of $|R|$ and $|S|$ equation by [7.19]+[7.21] and [7.20]+[7.22], we can then have:

[7.19]+[7.21]: $\frac{|V_0|^2+|V_2|^2}{2} = |R|^2 + |S|^2$ [7.27]

[7.20]+[7.22]: $\frac{|V_3|^2+|V_1|^2}{2} = |R|^2 + |S|^2$ [7.28]

We can then add $2|R||S|$ and $-2|R||S|$ to the [7.27] and [7.28] to form two sets of perfect square rule:

[7.27]+$2|R||S|$: $|S| + |R| = \sqrt{\frac{|V_0|^2+|V_2|^2}{2} + 2|R||S|}$ [7.29]

[7.28]-$2|R||S|$: $|S| - |R| = \pm\sqrt{\frac{|V_3|^2+|V_1|^2}{2} - 2|R||S|}$ [7.30]

Therefore we can then solve for $|R|$ and $|S|$ as:

[7.29]+[7.30]:$|S| = \left(\sqrt{\frac{|V_0|^2+|V_2|^2}{2} + 2|R||S|} \pm \sqrt{\frac{|V_3|^2+|V_1|^2}{2} - 2|R||S|}\right)/2$ [7.31]

[7.29]-[7.30]: $|R| = \left(\sqrt{\frac{|V_0|^2+|V_2|^2}{2} + 2|R||S|} \mp \sqrt{\frac{|V_3|^2+|V_1|^2}{2} - 2|R||S|}\right)/2$ [7.32]

This phase stepping algorithm is similar to the general phase stepping interferometer (Wizinowich 1990), however the derived equations here did not assume that $|R|^2 + |S|^2$ equal to a constant. This solves for the relative θ, and also the terms $|S|$ and $|R|$. Although we also used the 90 degrees phase shifts, arbitrary phase steps can be also implemented however with a more sophisticated algorithm as discussed in (Stoilov & Dragostinov 1997).

7.3.1 Phase SLM algorithm: criteria to design a phase pupil function

In this section, we again make an assumption that the *V(z)* interference can be described as a vectorial summation of two vector beams. Therefore we need to design a phase pupil function so that we can distinguish between the phase of the reference and SP beams. As we know sharp edges in the amplitude give spurious oscillations, in fact any discontinuity in the complex amplitude of the pupil function is undesirable; however, a sharp discontinuity in phase can be tolerated provided the amplitude at that point is low. This is explained in figures 7.20 and 7.21.

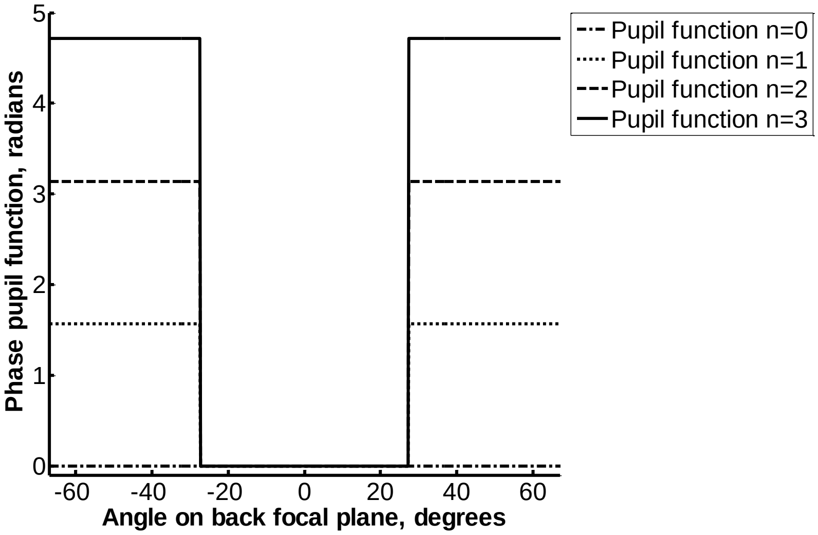


Figure 7.20 shows phase pupil functions used in this study

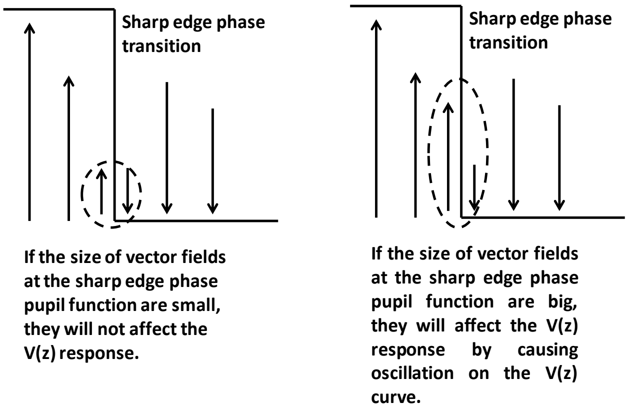


Figure 7.21 shows vector fields of the spatial frequencies around the sharp edge pupil function

It is important to note that the additional phase shift added by the P-SLM has to be the positive sign direction. The reason can be explained since the *V(z)* interference is formed by the centre point of the bfp and the phase of the spatial frequencies around the SP dip; for the positive phase shift, we can see from figure 7.22a that all the positive steps can provide steep gradients, which will generate ripples on the *V(z)* curves. On the other hand, for the negative phase shift as shown in figure 7.22b the phase transition at the plasmonic angle is now distorted and, of course, the ripples on the *V(z)* curves will be virtually eliminated although they will become visible on the as the lens is moved away from the sample.

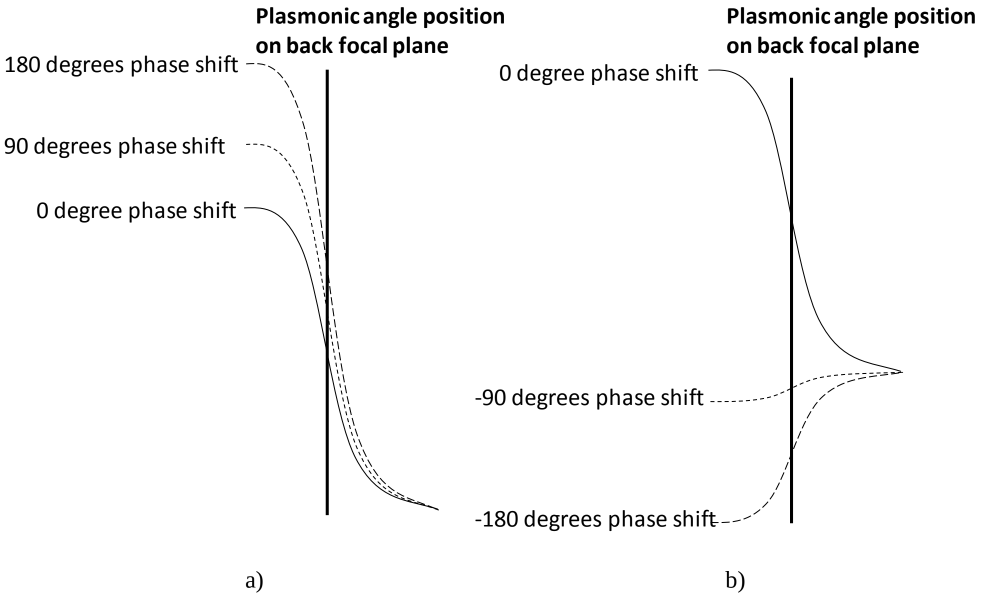


Figure 7.22shows a) positive phase shift effect for plasmonic case and b) negative phase shift effect for plasmonic case

7.3.2 Phase SLM algorithm simulation results

Let us now look at some simulated results obtained using P-SLM technique. The first task is to investigate whether this algorithm requires a smooth amplitude pupil function.

Simulation case: *46 nm* of gold in water ambient; *NA=1.65* with $n_0=1.78$ and $\lambda=633$ *nm* linear polarised wave with *10 nm* thick uniform sample *n=1.5*. We firstly look at the uniform pupil function case. It represents the situation where there is no amplitude pupil function, where all the points on the back focal plane are not amplitude modulated.

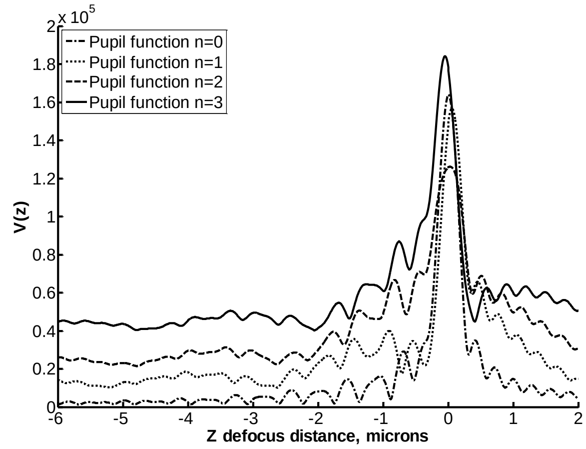


Figure 7.23 shows *V(z)* responses calculated using the phase pupil functions in figure 7.20 and a uniform amplitude pupil function

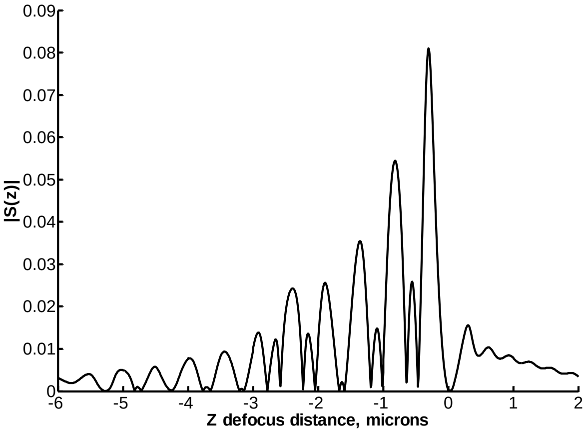


Figure 7.24 shows |*S(z)*| calculated from the *V(z)* curves shown in figure 7.23

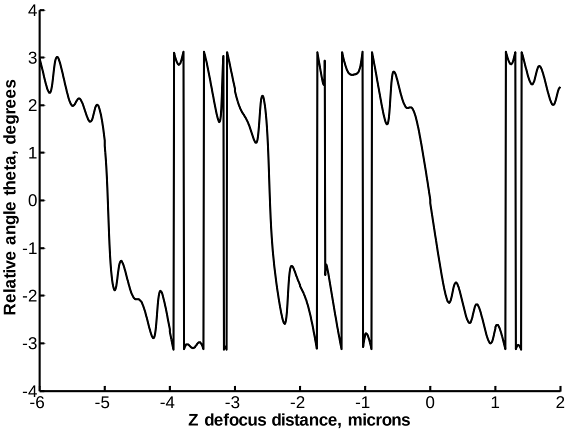


Figure 7.25 shows θ(z) calculated from the *V(z)* curves shown in figure 7.23

The results shown in figures 7.23 to 7.25 confirm us that the *V(z)*, *|S(z)|* and *θ(z)* curves were not stable. The next task is to employ the amplitude pupil function to see whether it can improve the stability of the results.

The next task is to determine whether the stability of the results can be improved if we introduce an amplitude pupil function to the system.

The amplitude pupil function used in this chapter is shown in figure 7.26 below and the phase pupil functions used were the same set of phase pupil functions shown in figure 7.20.

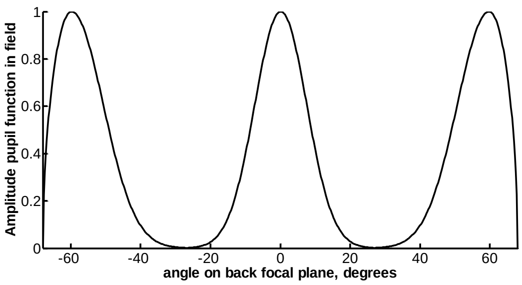
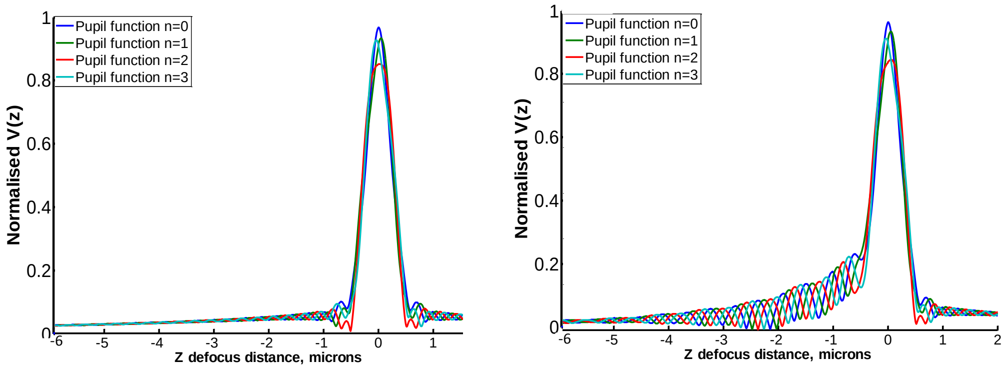


Figure 7.26 shows the amplitude pupil function used in the simulation

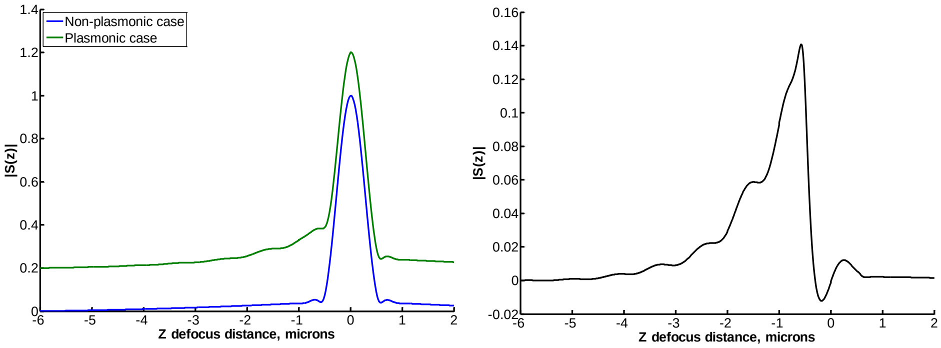
The *V(z)* curves calculated by using the amplitude pupil function in figure 7.26 and phase pupil functions in figure 7.20 is shown in figure 7.27b. Note that the non-plasmonic case described in figure 7.27a was provided by replacing the 46 nm layer of gold with 4 microns thick of gold. The *V(z)* curves for the non-plasmonic case calculated using the same pupil functions are shown in figure 7.27a.



a) b)

Figure 7.27 shows *V(z)* responses for a) non-plasmonic and b) plasmonic case

From the *V(z)* curves above, we can then substitute these into equation [7.31] to calculate |*S(z)*| curves as shown in figure 7.28.



a) b)

Figure 7.28 shows |*S(z)*| responses calculated from the results shown in figure 7.27 for a) non-plasmonic and plasmonic cases and b) |*S(z)*| after subtracting the plasmonic case with the non-plasmonic case

Similar to the amplitude stepping technique in order to calculate the |*S(z)*| term, it is required to subtract the |*S(z)*| of non-plasmonic case with the plasmonic case to avoid the non-plasmonic effect. The |*S(z)*| curve after suppressing the non-plasmonic effect is shown in figure 7.28b. As explained it is not only the |*S(z)*| curve that can be extracted using this phase modulation technique, but also the relative phase *θ(z)*. The *θ(z)* can be determined by substituting the *V(z)* results in figure 7.28 into equation [7.25] as shown in figure 7.29 below.

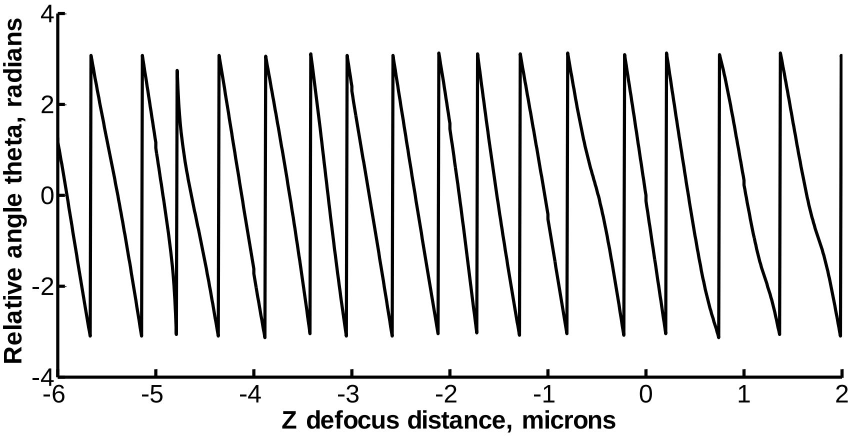


Figure 7.29 shows *θ(z)* calculated from the *V(z)* curves in figure 7.28b

The plasmonic angle θ_p calculated by Fourier transforming the *θ(z)* curve shown in figure 7.29 is 54.73 degrees. The exponential decay equation calculated is $y = e^{-0.1427x}$. We can then calculate the actual decay length by working out the position that gives *y=1/e*. The propagation length is 7.01 microns. The results calculated here agreed well with the amplitude stepping algorithm and the non-interferometric system discussed earlier.

Figures 7.28 and 7.29 show that the results obtained with a smooth amplitude pupil function were more stable compared to the uniform amplitude pupil function. This means that it is necessary to provide a smooth amplitude pupil function as well as the phase pupil functions.

Thanks to holographic concept, this enables us to use only a phase SLM device to modulate both amplitude and phase simultaneously. This can be done by using out of phase cancelation from the adjacent pixels on the phase SLM to modulate the amplitude. For example if we want an SLM pixel that give us 90 degrees phase modulation with 0.25 amplitude modulation in field. This technique requires quite a high number of pixels to do this; assuming that an SLM response of a single pixel is actually a contribution from 4 pixels. We know that waves with 180 degrees out off phase, such as -90 and 90 degrees will be cancelled out. Therefore we can now set the amplitude pupil function by: 3 pixels of 90 degrees and 1 pixel of -90 degrees. This process will, of course, change the overall phase profile of the pupil function; this problem can be sorted by carefully distributing the 2 opposite phases on the back focal plane pupil function to ensure that we get only the amplitude modulation with a uniform phase response. Once the amplitude response has been achieved, we can then add up the amount of phase shift we want on to the pupil function. For example, the amplitude pupil function required is shown in figure 7.26 and also at the same time we require the phase steps as shown in figure 7.20. The amplitude pupil function can be achieved by setting the P-SLM as shown in figure 7.30. We can then add three additional phase steps as shown in figure 7.31.

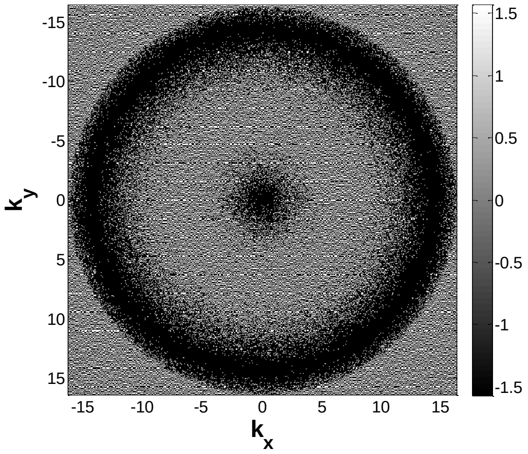
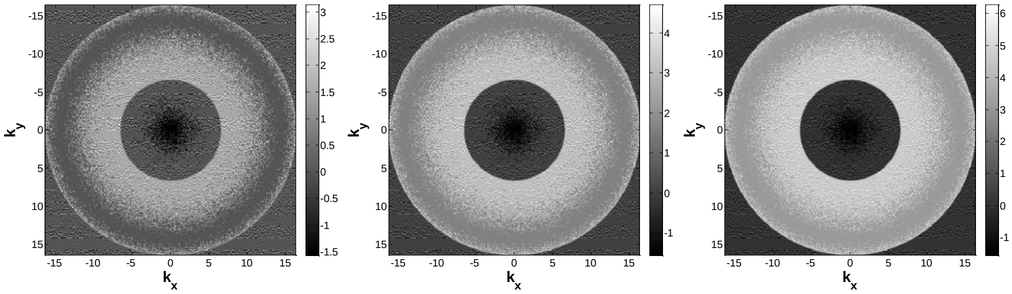
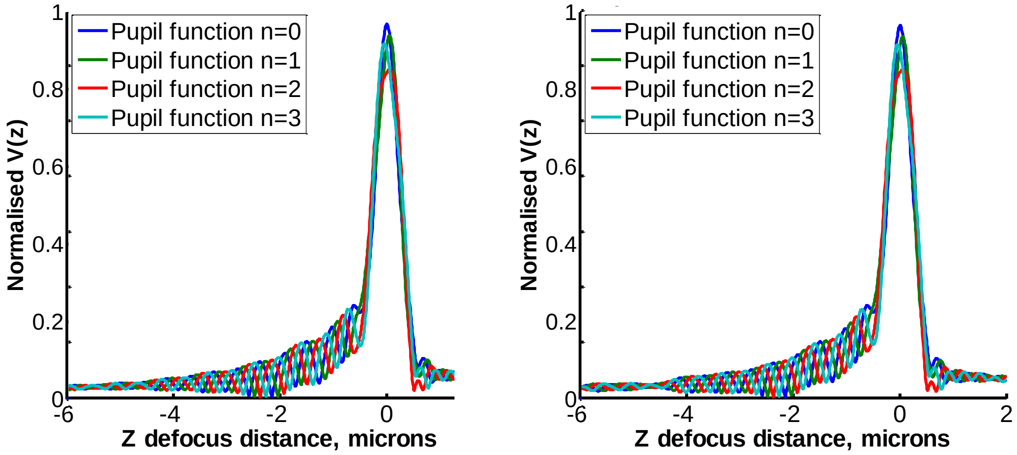


Figure 7.30 shows amplitude pupil function prepared by a phase spatial light modulator (P-SLM)



a) b) c)

Figure 7.31 shows amplitude and phase modulations using a phase spatial light modulator (P-SLM) a) for 90 degrees phase step, b) for 180 degrees phase step and c) for 270 degrees phase step



a) b)

Figure 7.32a shows *V(z)* curves calculated using two different spatial light modulators one for phase and one for amplitude modulation and b) calculated using the pupil functions in figure 7.31, using a single phase modulator.

From figure 7.32, we can see the *V(z)* curves of the two cases. The first case was calculated by employing two SLM devices one operating in amplitude mode and the other operating in phase mode (the first case shown in figure 7.32a) and the second case was calculated using only one P-SLM to control both phase and amplitude pupil functions (the second case shown in figure 7.32b). The results for the both cases were almost the same. There were some discrepancies around *z=-4.5 microns*. We can now conclude that by using only one P-SLM we can control both amplitude and phase pupil functions simultaneously.

7.3.3 Problem in phase SLM algorithm: Is a 100% fill factor required?

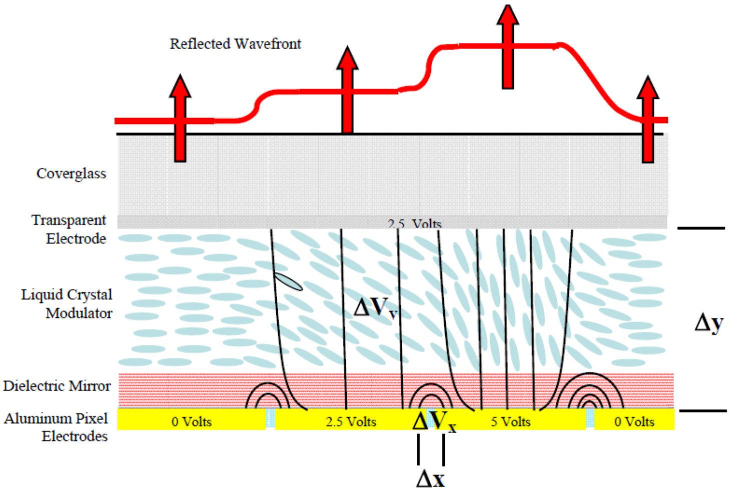


Figure 7.33 shows diagram of phase SLM device of BNS (BNS company website)

Phase SLM with 100% fill factor has become commercially available, such as XY liquid crystal array 512x512 of BNS. The fill factor is achieved by placing the electronics and employing a dielectric mirror at the back of each pixel as shown in figure 7.33.

7.4 Sensitivity to change in deposited material without defocusing

Part of the attraction of these phase stepping methods is to extract values without changing the defocus. Therefore if we measure the *V(z)* at a fixed z defocus, we can then extract $|S|$ and the relative angle θ. Although we cannot probably measure θp, we can measure the change in θp using these methods.

Let us consider the cases where the 46 nm case was deposited with a different thickness of $n=1.5$, the values the difference in relative angles $\Delta\theta$ between bare gold and coated gold cases for different thicknesses were calculated and shown in figure 7.34.

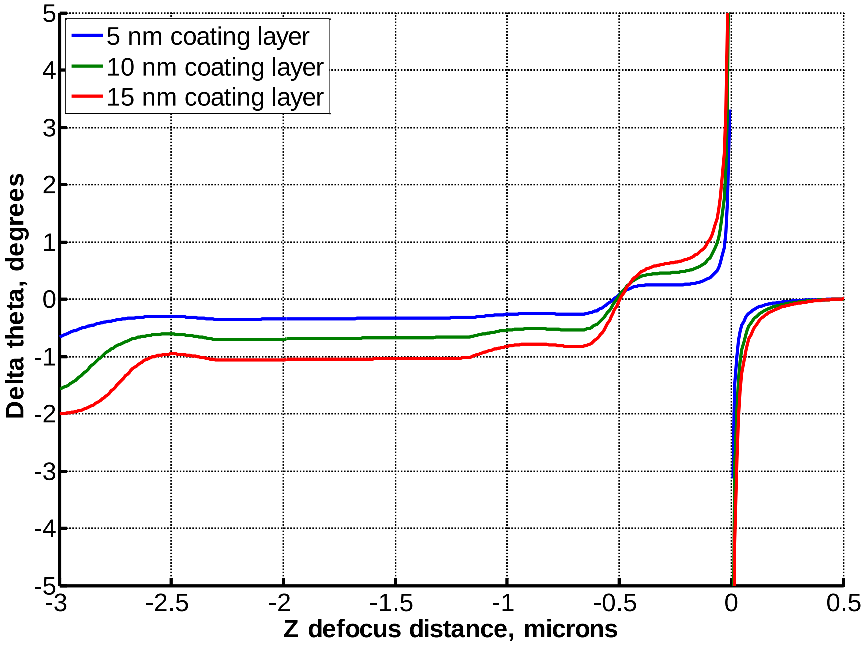


Figure 7.34 shows $\Delta\theta$ between bare gold and coated gold cases for different thicknesses of uniform coating layer

From figure 7.34, we can see that at any z defocus points between approximately -0.5 microns to -2.5 microns, the stepping algorithm can provide us with good sensitivity to detect the thickness change of the sample. The values of $\Delta\theta$ are 0.34, 0.68 and 1.04 degrees for 5 nm, 10 nm and 15 nm coating respectively when the defocus is –2 microns; this demonstrates that we have excellent linearity in the thickness measurement without defocusing provided the correct defocus is chosen. However, when the z defocus is beyond -3 microns the stability of the method seems to degrade dramatically; this may be due to the fact that the signal to noise (SNR) for the ripples at the far negative defocus is low.

7.5 Comparison of robustness to noise performance between the amplitude stepping and phase stepping techniques

Now let us consider the case where there is a presence of noise. In this section, we will consider an overall noise effect of the system by simulating *V(z)* responses with a random root mean square noise (RMS noise) of the electrical voltage *V(z)*. The RMS noise level was modelled with a Gaussian distribution with mean value of μ and variance of σ^2 of electrical voltage level, *Gaussian*(μ,σ^2).

The signal to noise ratio for the electrical voltage is defined as:

$$SNR_{electrical\ signal} = \frac{\mu^2}{\sigma^2} \qquad [7.33]$$

Note that the $\sqrt{SNR_{electrical\ signal}}$ term is directly proportional to signal to noise ratio (SNR) of the optical signal (Somekh et al 2011). It can be shown that when the variance σ^2 approaches the mean value μ; this condition is the condition for shot noise limit, which is the best case scenario for the noise simulation.

In this section, we consider effect of the noise on the both methods, where the mean is simply its voltage *V(z)* and variance $\sigma^2=0.03\mu^2$; this equivalents to the SNR of 33 (30.37 dBV).

Random Gaussian noise corresponding to the SNR level was added to the *V(z)* curves calculated from the confocal simulation. After adding the noise, the noisy *V(z)* curves were then processed using the amplitude stepping and the phase stepping techniques as explained in detail earler. For the amplitude stepping techniques, the pupil functions in figure 7.9 were used in this calculation. For the phase stepping technique, the phase pupil functions were modulated using the phase profiles shown in figure 7.20 and the amplitude pupil fucniton was modulated using the pupil function shown in figure 7.26.

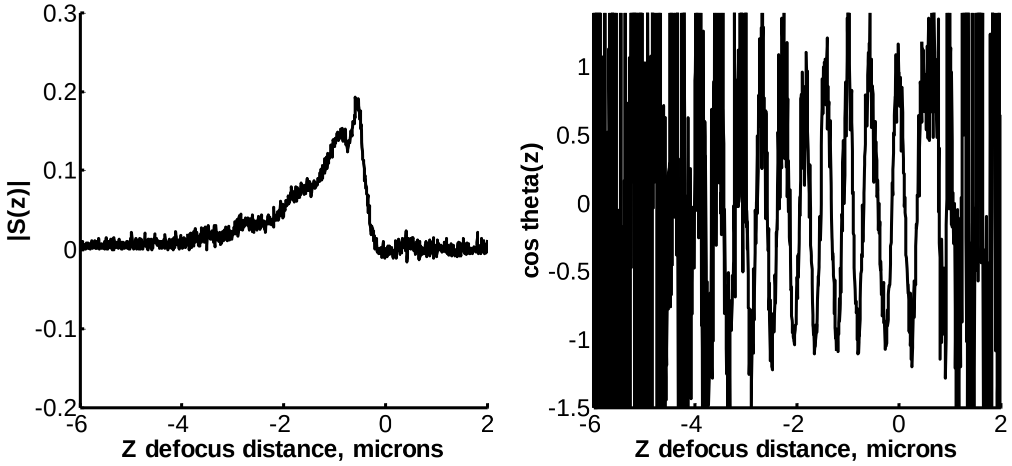


Figure 7.35 shows a) |*S(z)*| and b) *cosθ(z)* calculated using the A-SLM direct method with the noise level

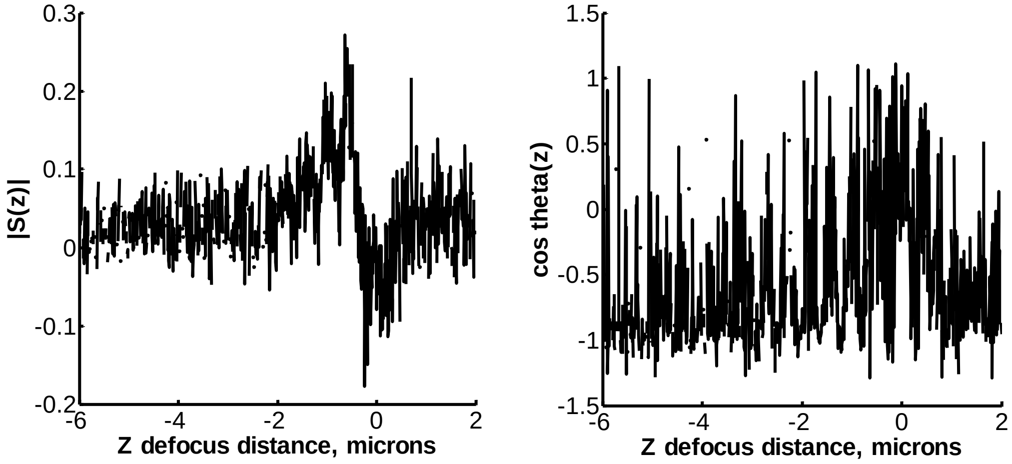


Figure 7.36 shows a) |*S(z)*| and b) *cosθ(z)* calculated using the A-SLM indirect method with the noise level

For the amplitude stepping, from the figure 7.35 and 7.36 we can see that the A-SLM direct method gave us a very good noise performance compared to the indirect method. The noise performance of the direct A-SLM method was comparable to the P-SLM method. As discussed earilier although the A-SLM direct method is good in term of noise performance, it does require a 100% fill factor A-SLM. In other words, the direct method is depicted an ideal scenario case for the amplitude stepping algorithm.

For the phase stepping technique, we can see from the result in figure 7.37 that that P-SLM technique was the most robust to noise, followed by the A-SLM direct method, the A-SLM indirect method respectively.

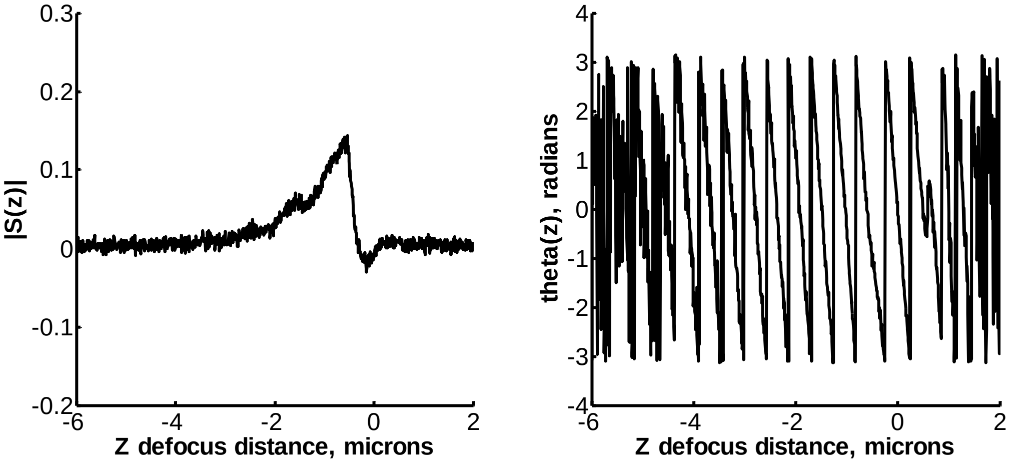
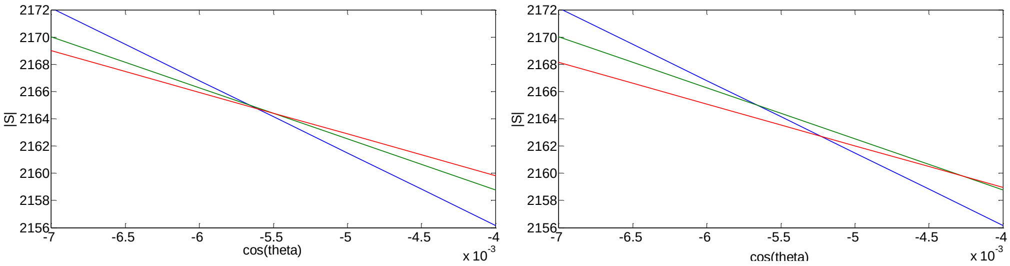


Figure 7.37 shows a) $|S(z)|$ and b) $\theta(z)$ calculated using the P-SLM method with the noise level

The next task is to investigate some possibilities to improve the noise performance of the algorithms. It is interesting to note that for the A-SLM indirect method at the absence of noise all the all the root locus curves passed through the same point, on the other hand when there was a presence of noise the root locus curves will not pass through the same point as depicted in figure 7.38.

In this section we will present a way to optimise the solution by 2D using geometric mean value weighing; this is similar to the way that physicists and civil engineers determine centriod of an object.



a) b)

Figure 7.38 Root locus plots at $z = 3.6$ microns a) noiseless case and b) noisy case

By having three curves so that we have 3 intersection pionts, the best point to represent this system is its centriod (this is a common civil engineering and classical physics term).

$$|S|_{centroid} = \frac{\int \cos(\theta)S(\cos(\theta))dcos(\theta)}{\int S(\cos(\theta))dcos(\theta)} \qquad [7.34]$$

$$\cos\,(\theta)_{centroid} = \frac{\int \cos(\theta)S(\cos(\theta))dS}{\int S(\cos(\theta))S} \qquad [7.35]$$

If the three curves are almost linear $|S|_{\text{centroid}}$ and $cos\theta_{\text{centroid}}$ can be approximated from

$$|S|_{centroid} = \frac{|S|_{12}+|S|_{13}+|S|_{23}}{3} \qquad [7.36]$$

$$\cos(\theta)_{centroid} = \frac{\cos(\theta)_{12}+\cos(\theta)_{13}+\cos(\theta)_{23}}{3} \qquad [7.37]$$

In addtion, the noise is proportional to the intensity of signal detected and the intensity of the ripples depend on the amplitude slm coefficient α_n; therefore we might as well include the α_n weighting factor into the averaging algorithm. The centriod approximation with weighting factor due to the α_n is given by:

$$|S|_{centroid} = \frac{(\alpha_1+\alpha_2)|S|_{12}+(\alpha_1+\alpha_3)|S|_{13}+(\alpha_2+\alpha_3)|S|_{23}}{2(\alpha_1+\alpha_2+\alpha_3)} \qquad [7.38]$$

$$cos\,\theta_{centroid} = \frac{(\alpha_1+\alpha_2)cos\,\theta_{12}+(\alpha_1+\alpha_3)\,cos\,\theta_{13}+(\alpha_2+\alpha_3)cos\,\theta_{23}}{2(\alpha_1+\alpha_2+\alpha_3)} \qquad [7.39]$$

Let us compare the performance of weighting and non-weighting approximations. We can see from the results in figure 7.39 that the weighting approximation can improve the stability of the results compared to the one without wieghting factor shown in figure 7.36 eariler. The figure shows that the overall noise performance of the weighting algorithm was better than than the none weighting factor case, more over the weighting factor case can reconstruct the *cosθ* term upto around -4 microns, whereas when there is no factor case we can recontruct the *cosθ* term upto only -2 microns.

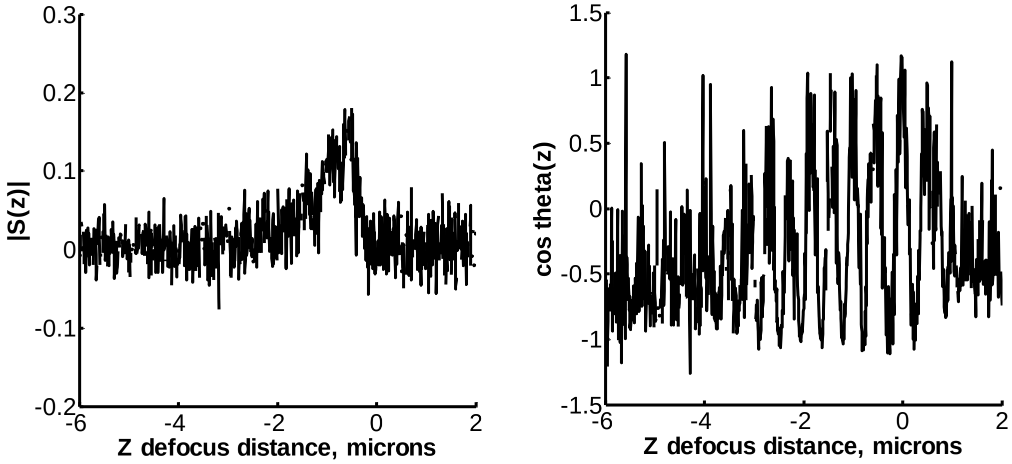


Figure 7.39 shows a) |*S(z)*| and b) *cosθ(z)* calculated using the A-SLM indirect method with weighting factor at the same noise level of 30.37dBV.

With only a single level of noise, we cannot conclude that the P-SLM method is better than the A-SLM methods. Therefore let us consider the case that the noise level is varied over the range of SNR input noise level of 0.5 (noise is greater than signal) up to 1000 (very low noise), which corresponds to -6.02 dBV to 60 dVB. We can then determine the SNR output from the |*S(z)*| calculated using the A-SLM direct and P-SLM methods by computing the Monte Carlo simulations as described by the following steps.

- *Step 1)* Calculate the |*S(z)*| for the noiseless case for the both A-SLM and P-SLM methods.
- *Step 2)* Vary the SNR input of the noise source the Gaussian noise function from -6.02dBV to 1000dBV.

- *Step 3)* Add the random Gaussian noise to *V(z)* curves. Then work out |*S(z)*| for the noisy case.
- *Step 4)* We can then determine the amount of noise for one random noise pattern by $Noise(z) = ||S(z)|_{noisy}{}^2 - |S(z)|_{noiseless}{}^2|$ [7.40]
- *Step 5)* Repeat steps 3 and 4 for a million times until the noise term in the equation [7.40] is stable.
- *Step 6)* Average the noise term by dividing the noise term in the equation [7.40] by 1,000,000.
- *Step 7)* Calculate $SNR_{output}(z)$ by $SNR_{output}\ (z) = |S(z)|_{noiseless}{}^2 / Noise(z)$
  Note that this is SNR in power ratio
- *Step 8)* Calculate the root mean square $SNR_{output}(z)$ over the z over the z defocus distance.

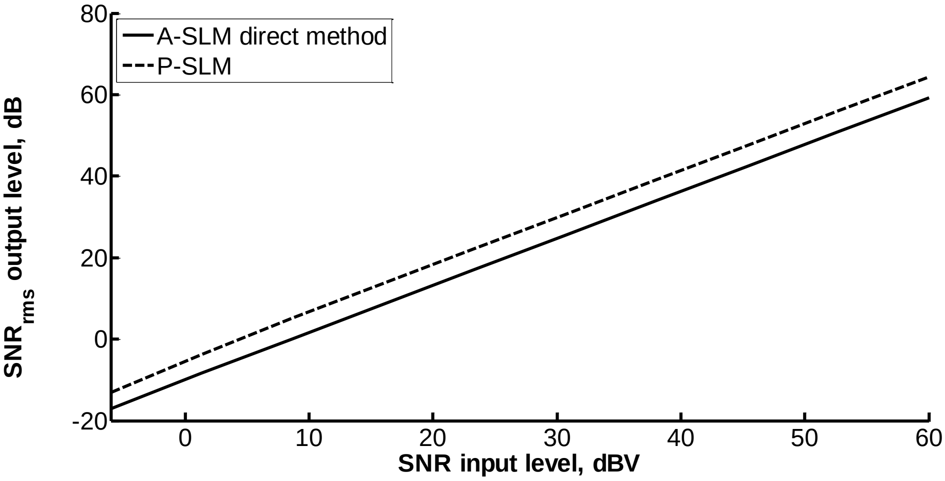


Figure 7.40 shows root mean square SNR of amplitude and phase stepping techniques in log scale as a function of SNR input

We can see from figure 7.40 that the noise performance of the P-SLM method is better than the A-SLM method over the input noise level range.

From figure 7.40, we can determine the slope of noise transfer function $N(SNR_{input})$, which is given by:

$$Slope\ of\ N(SNR_{input}) = \frac{SNR_{output}}{SNR_{input}} \qquad [7.41]$$

In other words, the slope of the noise transfer function $N(SNR_{input})$ are the gradients of the curves in figure 7.40. The noise transfer functions for the A-SLM direct method and the P-SLM method are 1.10 and 1.14 respectively. The noise performance for the P-SLM algorithm is better than the A-SLM by 6.43dB over the range of noise values tested.

7.6 How can the SP profile extraction algorithms be used to improve the measurement localisation?

As discussed in chapter 5 that the z defocus distance will affect the measurement localisation on the sample as a smaller *z* defocus scanning point will have a smaller footprint on the sample and, of course, it will have less crosstalk effect from the adjacent regions. In other words, measurements carried out closer to the focal point will give us the plasmonic angle θ_p closer to the uniform value.

Both the amplitude and phase stepping methods enable us to investigate this effect as we do not have to scan over the z defocus distance; whereas for the *V(z)* ripple period measurement, we need to measure at least a few ripples in order to calculate the ripple period and work out the plasmonic angle corresponding to the ripple period.

In this section, I will confirm that the measurement localisation of the SP for the lower z defocus distance is better than the longer z defocus distance. The measurement localisation criterion here is defined as how close we can measure the relative angle θ extracted using P-SLM technique at the centre of grating compared to its uniform cases; this is similar to the definition discussed in chapter 5. We can now define the accuracy of measurement η as:

$$\eta = \frac{\theta_{grating\ at\ the\ centre\ of\ n=1.5} - \theta_{grating\ at\ the\ centre\ of\ n=1.33}}{\theta_{uniform\ sample\ of\ n=1.5} - \theta_{bare\ gold\ with\ water\ of\ n=1.33}} \qquad [7.42]$$

The η term tells us about how closely we can resolve the SP relative angle θ in a confined region compared with the uniform sample case. In order to determine the measurement localisation at different z points; let us calculate η as a function of grating period at three z points; firstly *z=-0.48 microns* (at the beginning on the first ripple) and the second z point is *z=-0.92 microns* (the position in the middle between the first and the second ripples; this is used to represent the two consecutive peaks measurement method described in chapter 5) and the last point is *z=-1.35 microns* (which is the beginning position of the third ripple).

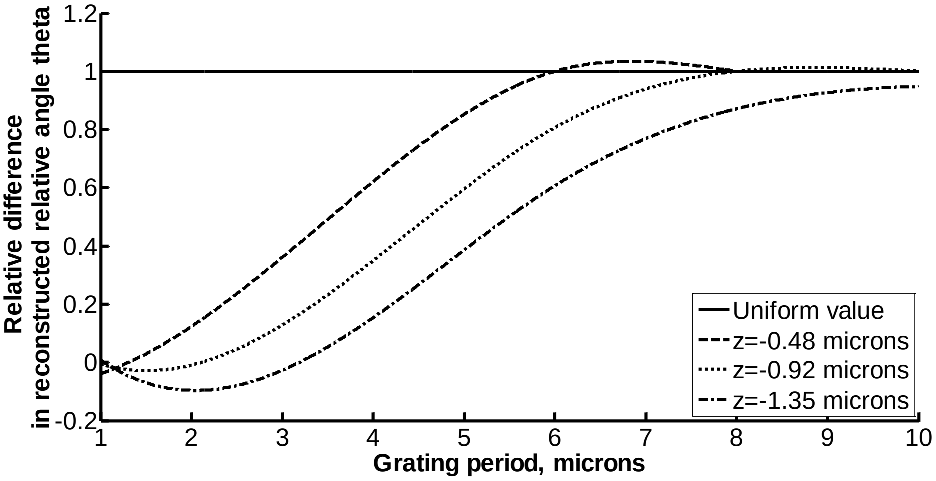


Figure 7.41 shows difference in recovered value η of reconstructed relative angle, θ, on coated (*10 nm* dielectric n=1.5) layer and bare layers versus grating period for incident x-polarisation. Solid curve: *uniform value*, dashed curve: *grating sample measured at z=-0.48 microns*, dotted curve: *grating sample measured at z=-0.92 microns*, dot-dashes curve: *grating sample measured at z=-1.35 microns.*

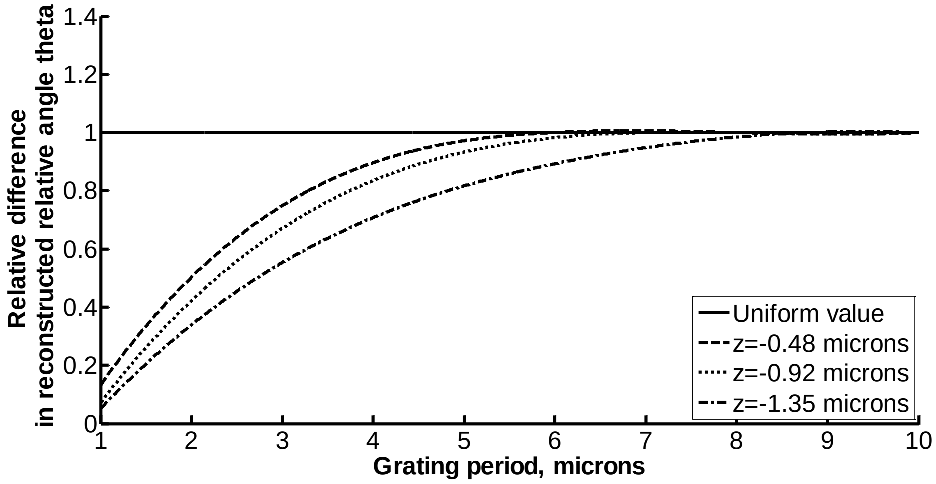


Figure 7.42 shows difference in recovered value η of reconstructed relative angle, θ, on coated (*10 nm* dielectric n=1.5) layer and bare layers versus grating period for incident y-polarisation. Solid curve: *uniform value*, dashed curve: *grating sample measured at z=-0.48 microns*, dotted curve: *grating sample measured at z=-0.92 microns*, dot-dashes curve: *grating sample measured at z=-1.35 microns.*

It is interesting to note that the η term calculated at *z=-0.92* microns is similar to the two consecutive peaks measurement described in chapter 5; this is because the *z=-*

0.92 microns point was actually at the centre of the first two peaks therefore it is a good position to represent the effect from the two ripples.

The figures 7.41 and 7.42 do not only confirm that the measurement localisation of the interferometric system does depend on the amount of crosstalk from the adjacent regions due to the z defocus position, but also clearly illustrate the use of amplitude and phase stepping techniques.

It is important to note that although we can see that the results in figure 7.41 and 7.42 enable us to quantify and study the SP effect at positions, which are close to the focal point, we have to note that the shape of the *V(z)* ripples around the first ripple of the two vector approximation may not be as stable as the other ripples as discussed in section 7.1.

7.6 Discussion and conclusion

In this chapter, we presented the new algorithms to extract an SPR propagation parameters, such as plasmonic angle θ_p and SP propagation length from the *V(z)* interference response using spatial light modulator. I have suggested some methods based on both A-SLM and P-SLM devices.

For the A-SLM, I have suggested two methods, which were direct and indirect methods. The direct method solved SPR parameters $|S(z)|$ and $\cos\theta(z)$ based on the concept of alternatively switching on and off the discrete pupil function, whereas the indirect method solved the SPR parameters by plotting root locus lines.

At the absence of noise, all the A-SLM can give us similar results. The limitations of the A-SLM algorithm is that it can only determine $|S(z)|$ and $\cos\theta(z)$, not the $\theta(z)$ term as we have ambigility in the cosine fucniton as discussed.

For the ideal case of 100% fill factor A-SLM, this makes the two vectors summation approximation hold and enables us to accurately model the *V(z)* full back focal plane integration calculation with two vector approximation. However, the main limitation of this algorithm seems to be the fact that it is necessary to have a very good fill factor SLM, otherwise the algorithm might fail. Although some more sophisticated techniques can be employed to improve the extracted values as discussed, this makes the size of the problem become more complicated.

In order to get around the problems, the P-SLM algorithm was investigated. We have demonstrated that the results obtained from the P-SLM were similar to the A-SLM and agreed well with the theoretical values for the noiseless simulations.

The phase modulation alone without an amplitude pupil function cannot give us a stable result as the ripples caused by discontinuities of the pupil function invalidate the two phasor approximation. We therefore still need an amplitude pupil function, we have illustrated that the amplitude pupil function can be provided together with phase modulation using only one P-SLM.

The main advantage of performing such calculations is that the algorithms allow us to quantify the measurement localisation and crosstalk at any particular z point without having to scan over a few *V(z)* ripples. By employing one of these methods, we can show that the z defocus closer to the focal point will have better measurement localisation performance because it has less crosstalk effect from the adjacent regions as the footprint of the focal spot size on the sample is determined by the largest value of defocus.

Chapter 8

Discussion, conclusion and future work

In this thesis, we have illustrated the use of Fresnel and rigorous diffraction theory to quantify SP microscope and sensor responses for different configurations and also developed a theoretical framework to predict the performance of non-interferometric and interferometric SP microscopes. In this chapter, I will summarise and highlight some of the key findings in this thesis and some possible future work.

8.1 Discussion and conclusion

Surface plasmons are a wave guided mode propagating along a metallic interface with decaying evanescent field perpendicular to the direction of propagation. This field is sensitive to refractive index attached to the metallic surface. The sensitivity of the SPs is attractive for measuring the chemical and biological change in nanometer layer of biological material. Therefore for imaging or sensing the properties of thin layer of chemical or biological materials the SPs are a very good candidate. As the properties of the SPs are significantly different from the optical waves propagating in free space; the microscopy techniques to optimise the SP performance differ from those for non-SPR microscopy.

Imaging performance

For non-interferometric microscopy, we have found that the non-interferometric microscope has a trade-off between lateral resolution and sensitivity. Images with high lateral resolution will have a poor contrast. The resolution depends on the propagation length of the SPs. The shorter propagation length will give a better resolution compared to the longer propagation length. It is necessary to provide an annular ring to excite SPs on the back focal plane and to avoid the contribution from non plasmonic angles on the back focal plane. If we image the sample without the mask, the image obtained will have virtually no contrast as the energy associated with SPs is small compared to the whole frequency spectrum on the back focal plane. The imaging performance of the SPs depends on the mask size and mask position. The position of the mask alters the apparent aspect ratio of the image. The non-interferometric system suffers from crosstalk, which can mean that the apparent aspect ratio differs from that of the sample. The contrast mechanism for the non-interferometric microscope is provided by different plasmonic angle on back focal plane θ_p. The contrast can be changed by altering the radial position of the annular mask; for example, if the sample is a grating with θ_{p1} and θ_{p2} corresponding to the two different material regions on the grating, if the mask is centered around the θ_{p1} the first material region will appear brighter and the second material region will appear darker, whereas if the mask is located around the θ_{p2}, the first material region will now appear darker and the other region will appear brighter instead. However, if the mask is at the centre of the two plasmons angles, images will have low contrast.

For the interferometric microscope, we have shown that the lateral resolution of the interferometric microscope is superior to the non-interferometric. The contrast mechanism of the interferometric come from the fact that the $V(z)$ ripple period is sensitive to refractive index changes of the sample. The lateral resolution of the image also depends on the amount of crosstalk. In other words, the lateral resolution will be better for low value of negative z defocus.. Images with contrast reversal can be observed as the defocus is changed. The advantage of the interferometric system compared to the non-interferometric system is that the interferometric microscope has shown its ability to provide us with high lateral resolution imaging system without losing too much contrast. We believe that the interferometric system provides a way to get around the trade-off between sensitivity and resolution in the non-interferometric system.

Sensitivity performance

For the non-interferometric setup, it is recommended to use the longest possible propagation length of SPs in order to have the highest sensitivity. The optical configuration is not an important factor for the non-interferometric system, whereas the propagation length of the SPs is the key characteristic that will affect the sensitivity of the sensor.

For the interferometric system the propagation length of the SPs is also an important feature but it affects the signal strength rather than the resolution. In this case there is an advantage of using the shorter propagation length as it can improve the ripple amplitude of the *V(z)* curve.

The optical configuration is a key factor for the sensitivity; the higher coupling index n_0 will give us the smaller SP angle. The smaller angle of SPs enables us to measure an SP effect in a more localized region, which is, of course, a key advantage of doing the interferometric system compared to the non-interferometric system. In other words, for the interferometric setup the performance of the sensor is now limited primarily by the optical configuration rather than the propagation length of the SPs. It is important to bear in mind that the real reason that they both have different sensitivity is because they operate with different mechanisms.

In this thesis, we have established a method to quantify the crosstalk and capability of the microscope to make an SP measurement in confined region. This is done by comparing the plasmonic angle measured at the centre of each material with the plasmonic angle of the uniform material case; this tells us how close to the uniform value we can measure SP parameter at a presence of crosstalk. We found that crosstalk is greater for the x polarisation compared to the y polarisation; and r polarisation gives intermediate results. The crosstalk is lower for shorter propagation length and also for the longer grating period. The *V(z)* measurement method appears to be sensitive not only along the direction of SPs propagation and also z direction as when we scan over the z defocus we might find that the size of the footprint on the sample is bigger than the size of the intended material. This will affect the accuracy of the z measurement. Therefore a high index couplant plays an important role to improve the localisation as the high refractive index couplant shifts the plasmon angle inside the back focal plane, so it reduces the footprint of the focal point on the sample $2z\tan(\theta_p)$.

In the study, we have addressed the accuracy of the $V(z)$ measurements; the measurement of $V(z)$ will be accurate if some criteria are met which are: firstly, the pupil function is designed to avoid high frequency components oscillation and secondly phases of the centre $(2k_{vec}z)$ of the back focal plane and the SPs angle $(2k_{vec}z\cos\theta)$ move at well defined angular velocity corresponding to the angle of incidence.

We have also suggested two techniques to extract SPs profile from $V(z)$ measurement, which are amplitude stepping and phase stepping algorithms. The phase stepping is, of course, more robust to noise, however, the amplitude stepping technique is also provided in order to investigate the feasibility of using a cheaper device to carry out a similar task as the phase modulator. The algorithms provide us with a new sensing and imaging mechanisms. The methods can give us a means to measure changes in index without scanning over the z defocus distance; this removes the need to perform mechanical scans. We have also shown that the SPs parameters obtained from the algorithm are accurate compared to theoretical value. The $|S(z)|$, $\theta(z)$ and $\theta(z)$ provide us a way to quantify the crosstalk effect without having to scan over a few ripples.

Recently, my colleagues and I have successfully demonstrated the use of confocal SP microscope to perform the $V(z)$ measurement. The results obtained from the confocal microscope are similar to those obtained from the heterodyne interferometric microscope. The differences between the heterodyne and confocal microscopes are that firstly for the confocal microscope the ripple amplitude is smaller than the heterodyne system as it detects $|V(z)|^2$ rather than *V(z)* and secondly the ripple amplitude also depends on the pinhole of the confocal microscope. For the smaller ripple amplitude issue, we have demonstrated the use of pupil function engineering to increase the ripple magnitude. By choosing the confocal pinhole so that it is small enough we can get the results similar to the heterodyne system.

8.2 Summary of what have been achieved during this study

In this study, I have developed a reliable study tool that is capable of investigating SPs effects on 1D and 2D nano structures. The software is not designed to explore only the SPs, but it is also capable of exploring other effects, such as ATR based systems and waveguides. The software is now employed in other research projects in our group, such as designing a structure to give parallel confocal focal spots for parallel confocal scanning microscope, asymmetrical plasmonic structure to excite SPs with a shorter propagation length.

We have developed a theoretical framework to understand the performance and mechanism of different SPs microscopes. We also provided a quantitative way to compare the performance of each SP microscope setup.

We have successfully demonstrated the use of a confocal SPR microscope to measure the *V(z)* response. The main advantage of using the confocal system compared to the heterodyne system is that it reduces the complexity of the optical setup by providing the reference beam together with the reflected beam, so called common path interferometer.

We have suggested and validated robust ways to accurately extract SPs profile from the *V(z)* measurement. The SPs profile enables us to measure the SPs response without having to scan over the z defocus distance; this will, of course, save the scanning time. Not only is the scan time improved, but also the technique enables one to measure the SP response with less crosstalk from the adjacent regions as we can now measure the SPs effect with a low negative z value as discussed in chapter 6.

8.3 Future work

In this section, I will provide some ideas that can be extended from the study. All ideas presented here are backed up by simulations in order to show the feasibility of the ideas.

8.3.1 V(ϕ)

One of the challenges of the interferometric detection method is to find methods that do not involve defocusing. We have discussed different methods to do this in chapter 6. Here is another possible method. This method requires a phase SLM to modulate the back focal plane. We now use this to simulate the precise effect of defocus. As mentioned in chapter 5 that when the sample is defocused, the phase of the reflected beam is shifted by phase factor given by:

$$\emptyset(sin\theta) = \alpha(1 - cos\theta) \qquad [8.1]$$

Where α is proportionality constant that we vary to mimic the effect of defocus, α is simply *2kz* and θ is the angle of incidence. Therefore, if we can modulate the phase of each position on the back focal plane to have a radially varying phase shift as shown in the equation 8.1, this means that we can reproduce the *V(z)* response without scanning the z defocus.

There are some technical issues concerning whether we are incident on the SLM in either the forward and reverse directions or whether the light interacts in both directions. For uniform samples this simply has the effect of doubling the phase shift imposed by the SLM, for non-uniform sample we may need to simulated the effect, the two situations will be a little different because the diffracted orders will be phase shifted on return by a different amount from the incident beam, simply because the return is at different angles.

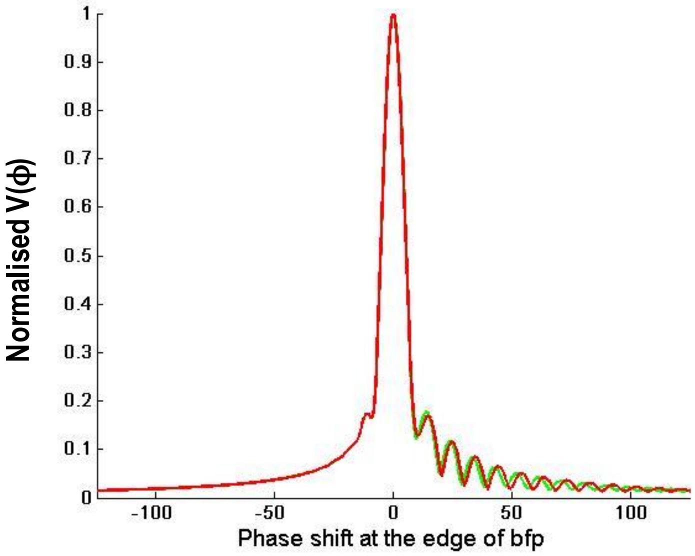


Figure 8.1 shows *V(ϕ)* for different curvatures on the back focal plane. The red one is uncoated and the green one is coated with a 10nm layer of *n=1.5*. The lens was 1.65 and water is the final medium.

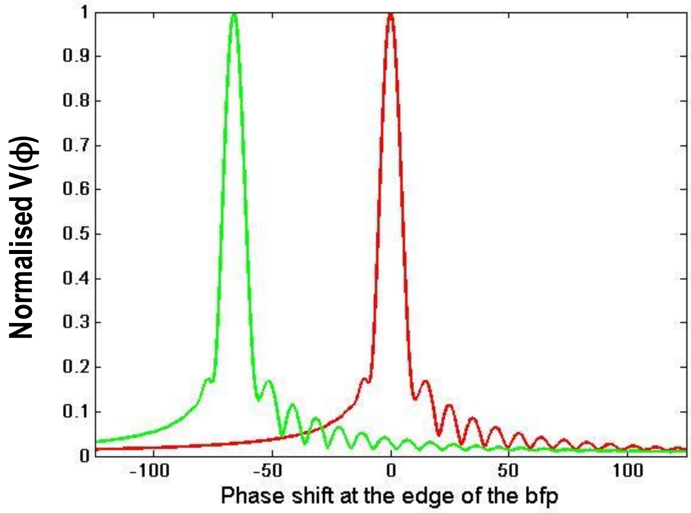


Figure 8.2 shows *V(ϕ)* curves for different phases. The green is for a z defocus of -3 microns. If we use a positive defocus it moves the same distance in the other direction.

The figures 8.1 and 8.2 demonstrated the feasibility of using the phase SLM to reproduce the same response as *V(z)* without scanning the z axis; this means the technique has better measurement speed and has less problem with the microphonic noise due to the moving stage sample stage. Let us consider how the method would operate. We fix the sample and change the value of *a* of fig. 8.2 will move so that a curve equivalent to the *V(z)* is produced without moving the sample.

8.3.2 Sensitivity tuneable using phase SLM

The sensitivity of the *V(z)* measurement depends on the *z* operating point. Some of z positions give us a very good sensitivity; on the other hand some of the positions do give a poor contrast. The contrast can be adjusted by providing additional phase shift to the SP angle on the back focal plane, this is equivalent to moving the *V(z)* curves around the operating point. The clear illustration of using the technique is that for example we know that the low value of z defocus will give us a good resolution however the contrast will be low as illustrated in figure 8.3. Of course, we can increase the contrast of the image by defocusing to another z position, but we cannot change the effects of the crosstalk.

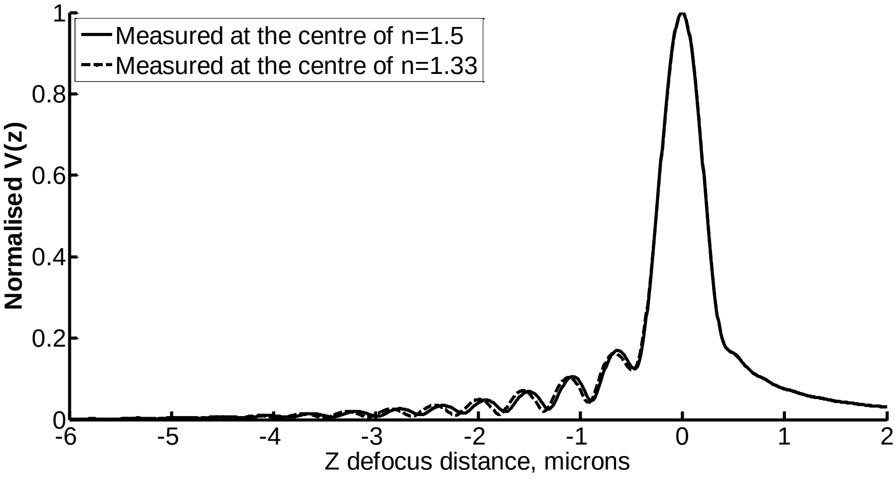


Figure 8.3 shows *V(z)* curves measured at centres of each material of grating sample 8 micron period; $NA=1.65$, $n_0=1.78$, $\lambda=633$ nm linear polarised wave.

From the two points *V(z)* mechanisms, we take a linescan image at $z=-\ 0.5$ microns the image will have definitely low contrast. Let us apply some phase shift distribution on the plasmonic angle and keep the centre part of the back focal plane phase fixed as the same way as explained in the previous chapter. By plotting the *V(phase shift)* as a function of phase stepping at the two centres of grating regions while keeping z fixed at -0.5 microns as shown in the figure 8.4 below.

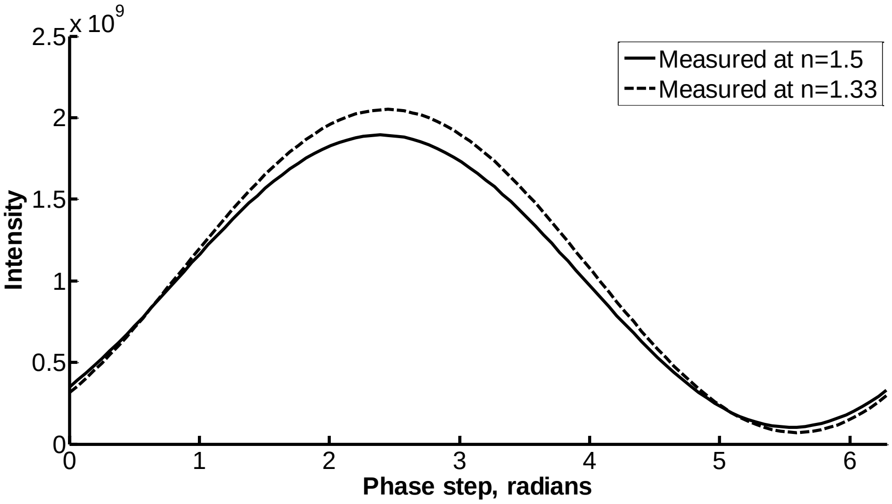


Figure 8.4 shows *V(phase step)* taken at *z=-0.5* microns

From figure 8.4, we can see that without phase stepping (*phase shift = 0 radian*) the two centres of grating have virtually no contrast; however with some phase step, such as 2.5 radians, the contrast of the image can be improved.

8.3.3 Asymmetrical surface plasmons

Asymmetrical surface plasmons (a-plasmons) mode is another plasmonic mode occurring with a thin metallic film such as 15-20 nm of gold. The requirements to excite the a-plasmons are a much thinner layer of gold and also allowing the sine of the angle of incidence to be much larger than unity; we will show later that this can be achieved with grating vector coupling. From figure 8.5, we can see that at incident *k-vectors* much greater than n_0 we get a large enhancement of the field, greater than 1. This is due to the excitation of another plasmon mode. We can see some evidence of the usual plasmon mode around 1.4 but this is weak because the gold layer is thin. By examining the phase response it shows that this is a plasmon mode. We see that the thinner the film the greater the *k*-vector, that is the shorter the wavelength.

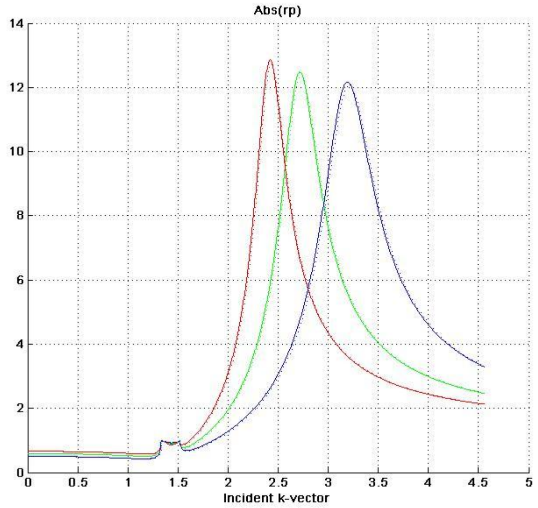


Figure 8.5 modulus of reflection coefficient for excitation of gold layers for layer thicknesses of 18nm (red), 15nm(green) and 12nm(blue).

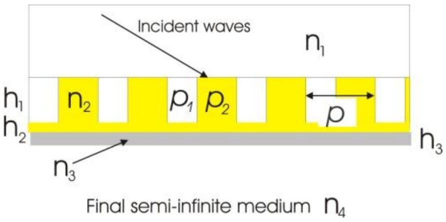


Figure 8.6 shows the gold grating for a-plasmons excitation

Figure 8.6 shows a picture of the sample we would like to produce to test the a-plasmon idea. Essentially, it is just a series of stripes with the dimensions mentioned in the image. The simulation predicts the following angle scan at 633 nm as shown in figure 8.7 below.

Figure 8.7 shows the reflection coefficient as a function of incident angle represented as $sin\theta$. The figure used for the simulation in figure 8.9 are $n_1=1.518$, n_2-gold at 633 nm, $n_3=1.34$, $n_4=1.33$. $h_1=10$ nm, $h_2=15$ nm, $h_3=1$ nm for the solid curve and 100 nm for dashed curve, $p_1=p_2=140$ nm. At large angles we see the dip at large angles; this is just the ordinary plasmon. The dip is pretty weak because we are more interested in the a-plasmon which appears at *20 degrees*. Now the normal plasmon is excited by the zero order excitation. Now look at the a-plasmon, this is the dip at 20 degrees. We still need to optimise the depth of the dip and the structure to optimise the a-plasmons effects. The a-plasmons mode has a shorter penetration depth compared to the normal SPs mode; this means that it will be useful for sensing refractive index change in a thin layer.

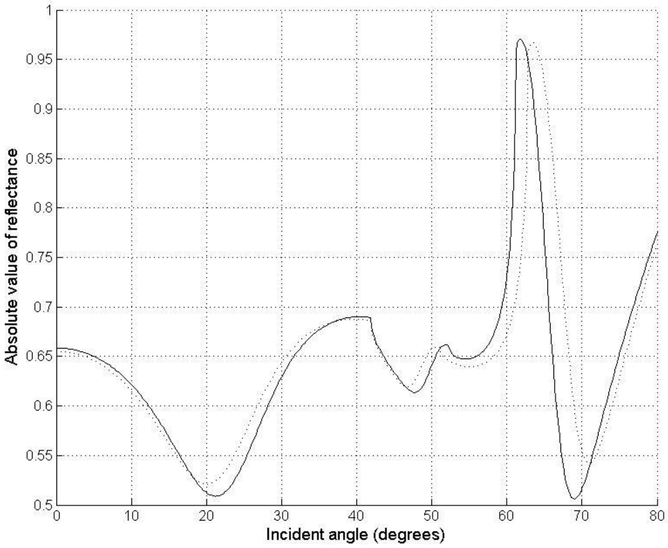


Figure 8.7 shows the reflection coefficient as a function of incident angle for $n_1=1.518$, n_2-gold at 633nm, $n_3=1.34$, $n_4=1.33$. $h_1=10nm$, $h_2=15nm$, $h_3=1nm$

8.3.4 Some 2D structure to enhance sensitivity of SP

Over the past few years, there are many attempts to improve the sensitivity of SPs by using hole arrays, pole arrays and other structured shapes. In chapter 2, we have shown that the RCWA calculation can be used to accurately model the rectangular nanohole array; the results obtained in chapter 2 agreed well with the results presented in (Yang & Ho 2009), where they used FDTD to model the structure. Therefore the next task is to investigate some 1D or 2D structures that can be used to enhance the sensitivity of SPs and try to develop a theoretical framework to encapsulate and understand their sensitivity mechanisms.

8.4 What are deliverables of the project?

In this section, I would like to end this thesis by summarising some of applications, where the knowledge studied in this thesis is applicable.

- Objective lens based SPR instrument will provide a viable quantitative tool for examining small refractive index changes on length scales not accessible to the prism based SPR methods.
- As discussed that the heterodyne interferometric system is quite a challenge in optical alignment, therefore the confocal SP microscope will provide a more compact and convenient to build. Indeed the confocal arrangement allows more processing on the sample on account of its simplicity. I hope the concept of the confocal SP microscopy will be extended and used more widely to study the effect of SP.
- For the SP detection array, the knowledge about crosstalk will be a key tool for designing the size of detection area so that its scale does not affect the accuracy of SP measurements.

References

Aldric S. 2011. Material Safety Data Sheet.

Anker JN, Hall WP, Lyandres O, Shah NC, Zhao J, Van Duyne RP. 2008. Biosensing with plasmonic nanosensors. *Nature Materials* 7:442-53

Argoul F, Monier K, Roland T, Elezgaray J, Berguiga L. 2010. High resolution surface plasmon microscopy for cell imaging. In *Biophotonics: Photonic Solutions for Better Health Care Ii*, ed. JTVVMDL Popp

Atalar A. 1978. An angular-spectrum approach to contrast in reflection acoustic microscopy. *Journal of Applied Physics* 49:5130-9

Atalar A. 1979. A physical model for acoustic signatures. *Journal of Applied Physics* 50:8237-9

Azzam RMA. 1977. *Ellipsometry and polarized light / R. M. A. Azzam and N. M. Bashara*. Amsterdam ; New York : New York :: North-Holland Pub. Co. ; sole distributors for the U.S.A. and Canada, Elsevier North-Holland

Barnes WL, Dereux A, Ebbesen TW. 2003. Surface plasmon subwavelength optics. *Nature* 424:824-30

Beketov GV, Shirshov YM, Shynkarenko OV, Chegel VI. 1998. Surface plasmon resonance spectroscopy: prospects of superstrate refractive index variation for separate extraction of molecular layer parameters. *Sensors and Actuators B: Chemical* 48:432-8

Berger CEH, Kooyman RPH, Greve J. 1994. Resolution in surface-plasmon microscopy. *Review of Scientific Instruments* 65:2829-36

Bergman DJ, Stockman MI. 2003. Surface plasmon amplification by stimulated emission of radiation: Quantum generation of coherent surface plasmons in nanosystems. *Physical Review Letters* 90

Berguiga L, Roland T, Fahys A, Elezgaray J, Argoul F. 2010. High resolution surface plasmon imaging of nanoparticles. In *Nanophotonics Iii*, ed. DLNJMOA Andrews

Berguiga L, Zhang S, Argoul F, Elezgaray J. 2007. High-resolution surface-plasmon imaging in air and in water: V(z) curve and operating conditions. *Optics Letters* 32:509-11

Born M, Wolf E. 1999. *Principles of optics : electromagnetic theory of propagation, interference and diffraction of light*. Cambridge; New York: Cambridge University Press

Bosenber.J. 1971. Photoelectrons from optically excited nonradiative surface plasma oscillations. *Physics Letters A* A 37:439-&

Bozhevolnyi SI. 1996. Localization phenomena in elastic surface-polariton scattering caused by surface roughness. *Physical Review B* 54:8177-85

Brekhovskikh LM, Evtushenko VA, Makarov SS, Pisarenko VF. 1960. Calculation of the vertical profile of sound velocity in a sea. *Doklady Akademii Nauk Sssr* 135:581-3

Brockman JM, Frutos AG, Corn RM. 1999. A multistep chemical modification procedure to create DNA arrays on gold surfaces for the study of protein-DNA interactions with surface plasmon resonance imaging. *Journal of the American Chemical Society* 121:8044-51

Cardona M. 1971. Fresnel reflection and surface plasmons. *Am. J. Phys.* 39:1277-&

Chen B-H, Wang Y-C, Lin J-H, Electromagnetics A. 2008. High sensitivity of phase-based surface plasmon resonance in nano-cylinder array. In *Piers 2008 Cambridge, Proceedings,* pp. 511-4

Chen WP, Ritchie G, Burstein E. 1976. Excitation of surface electromagnetic-waves in attenuated total reflection prism configurations. *Physical Review Letters* 37:993-7

Cooper MA. 2002. Optical biosensors in drug discovery. *Nature Reviews Drug Discovery* 1:515-28

de Bruijn HE, Kooyman RPH, Greve J. 1992. Choice of metal and wavelength for surface-plasmon resonance sensors: some considerations. *Appl. Opt.* 31:440_1-2

Dostalek J, Vaisocherova H, Homola J. 2005. Multichannel surface plasmon resonance biosensor with wavelength division multiplexing. *Sensors and Actuators B-Chemical* 108:758-64

Ekgasit S, Thammacharoen C, Yu F, Knoll W. 2005. Influence of the metal film thickness on the sensitivity of surface plasmon resonance biosensors. *Appl. Spectrosc.* 59:661-7

Elezgaray J, Roland T, Berguiga L, Argoul F. 2010. Modeling of the scanning surface plasmon microscope. *Journal of the Optical Society of America a-Optics Image Science and Vision* 27:450-7

Ermolaev AM, Rashba GI, Solyanik MA. 2011. Collective excitations of electron gas on the nanotube surface in a magnetic field. *Low Temperature Physics* 37:919-24

Eustis S, El-Sayed MA. 2006. Why gold nanoparticles are more precious than pretty gold: Noble metal surface plasmon resonance and its enhancement of the radiative and nonradiative properties of nanocrystals of different shapes. *Chemical Society Reviews* 35:209-17

Fanton JT, Opsal J, Willenborg DL, Kelso SM, Rosencwaig A. 1993. Multiparameter measurements of thin-films using beam-profile reflectometry. *Journal of Applied Physics* 73:7035-40

Flanagan MT, Pantell RH. 1984. Surface-plasmon resonance and immunosensors. *Electronics Letters* 20:968-70

Galush WJ, Shelby SA, Mulvihill MJ, Tao A, Yang P, Groves JT. 2009. A Nanocube Plasmonic Sensor for Molecular Binding on Membrane Surfaces. *Nano Letters* 9:2077-82

Giebel CB, S. Herminghaus,M. Riedel,P. Leiderer, U. Weiland,M. Bastmeyer. 1999. Imaging of Cell/Substrate Contacts of Living Cells with Surface Plasmon Resonance Microscopy. *Biophysical Journal* 76:8

Giebel KF, Bechinger C, Herminghaus S, Riedel M, Leiderer P, et al. 1999. Imaging of cell/substrate contacts of living cells with surface plasmon resonance microscopy. *Biophysical Journal* 76:509-16

Gong Q, Hsu SS. 1994. Aberration measurement using axial intensity. *Optical Engineering* 33:1176-86

Grabar KC, Freeman RG, Hommer MB, Natan MJ. 1995. Preparation and characterization of au colloid monolayers. *Analytical Chemistry* 67:735-43

Green RJ, Tasker S, Davies J, Davies MC, Roberts CJ, Tendler SJB. 1997. Adsorption of PEO-PPO-PEO triblock copolymers at the solid/liquid interface: A surface plasmon resonance study. *Langmuir* 13:6510-5

Grigorenko AN, Beloglazov AA, Nikitin PI, Kuhne C, Steiner G, Salzer R. 2000. Dark-field surface plasmon resonance microscopy. *Optics Communications* 174:151-5

Grigorenko AN, Nikitin PI, Kabashin AV. 1999. Phase jumps and interferometric surface plasmon resonance imaging. *Applied Physics Letters* 75:3917-9

Hecht B, Bielefeldt H, Novotny L, Inouye Y, Pohl DW. 1996. Local excitation, scattering, and interference of surface plasmons. *Physical Review Letters* 77:1889-92

Hell SW, Lindek S, Stelzer EHK. 1994. Enhancing the axial resolution in far-field light-microscopy - 2-photon 4pi confocal fluorescence microscopy. *Journal of Modern Optics* 41:675-81

Holmes RD, Somekh MG. 1994. Extended-focus phase imaging with an interferometric confocal microscope. *Appl. Opt.* 33:654-61

Homola J. 2008. Surface plasmon resonance sensors for detection of chemical and biological species. *Chemical Reviews* 108:462-93

Homola J, Koudela I, Yee SS. 1999a. Surface plasmon resonance sensors based on diffraction gratings and prism couplers: sensitivity comparison. *Sensors and Actuators B: Chemical* 54:16-24

Homola J, Lu HB, Yee SS. 1999b. Dual-channel surface plasmon resonance sensor with spectral discrimination of sensing channels using dielectric overlayer. *Electronics Letters* 35:1105-6

Hu C, Liu, D. 2010. High-performance grating coupled surface plasmon resonance sensor based on Al-Au Bimetallic layer. *Modern Appl. Sci*:4

Huang D-W, Ma Y-F, Sung M-J, Huang C-P. 2010. Approach the angular sensitivity limit in surface plasmon resonance sensors with low index prism and large resonant angle. *Optical Engineering* 49

Huang YH, Ho HP, Wu SY, Kong SK. 2012. Detecting Phase Shifts in Surface Plasmon Resonance: A Review. *Advances in Optical Technologies* 2012

Ichimura I, Hayashi S, Kino GS. 1997. High-density optical recording using a solid immersion lens. *Applied Optics* 36:4339-48

Ilett C, Somekh MG, Briggs GAD. 1984. Acoustic microscopy of elastic discontinuities. *Proceedings of the Royal Society of London Series a-Mathematical Physical and Engineering Sciences* 393:171-&

Inagaki T, Kagami K, Arakawa ET. 1982. Photoacoustic study of surface plasmons in metals. *Appl. Opt.* 21:949-54

Iwata T, Mizutani Y. 2010. Ellipsometric measurement technique for a modified Otto configuration used for observing surface-plasmon resonance. *Optics Express* 18:14480-7

Jain PK, Huang X, El-Sayed IH, El-Sayed MA. 2008. Noble Metals on the Nanoscale: Optical and Photothermal Properties and Some Applications in Imaging, Sensing, Biology, and Medicine. *Accounts of Chemical Research* 41:1578-86

Jamil MMA, Denyer MCT, Youseffi M, Britland ST, Liu S, et al. 2008. Imaging of the cell surface interface using objective coupled widefield surface plasmon microscopy. *Journal of Structural Biology* 164:75-80

Jiang H, Sabarinathan J. 2010. Effects of Coherent Interactions on the Sensing Characteristics of Near-Infrared Gold Nanorings. *Journal of Physical Chemistry C* 114:15243-50

Johnson PB, Christy RW. 1972. Optical Constants of the Noble Metals. *Physical Review B* 6:4370-9

Kabashin AV, Evans P, Pastkovsky S, Hendren W, Wurtz GA, et al. 2009a. Plasmonic nanorod metamaterials for biosensing. *Nature Materials* 8:867-71

Kabashin AV, Patskovsky S, Grigorenko AN. 2009b. Phase and amplitude sensitivities in surface plasmon resonance bio and chemical sensing. *Opt. Express* 17:21191-204

Kano H, Knoll W. 1998. Locally excited surface-plasmon-polaritons for thickness measurement of LBK films. *Optics Communications* 153:235-9

Kano H, Knoll W. 2000. *A scanning microscope employing localized surface-plasmon-polaritons as a sensing probe*. Amsterdam, PAYS-BAS: Elsevier. 11-5 pp.

Kano H, Mizuguchi S, Kawata S. 1998. Excitation of surface-plasmon polaritons by a focused laser beam. *J. Opt. Soc. Am. B* 15:1381-6

Khavasi A, Mehrany K, Rashidian B. 2007. Three-dimensional diffraction analysis of gratings based on Legendre expansion of electromagnetic fields. *Journal of the Optical Society of America B-Optical Physics* 24:2676-85

Kikuta H, Yoshida H, Iwata K. 1995. Ability and limitation of effective-medium theory for subwavelength gratings. *Optical Review* 2:92-9

Kim DG, Choi WK, Choi YW, Dagli N. 2007. Triangular resonator based on surface plasmon resonance of attenuated reflection mirror. *Electronics Letters* 43:1365-7

Konopsky VN. 2000. Operation of scanning plasmon near-field microscope with gold and silver tips in tapping mode: demonstration of subtip resolution. *Optics Communications* 185:83-93

Kretschm.E, Raether H. 1968. Radiative decay of non radiative surface plasmons excited by light. *Zeitschrift Fur Naturforschung Part a-Astrophysik Physik Und Physikalische Chemie* A 23:2135-&

Kwiecien P, Richter I. 2011. *Efficient Three Dimensional Aperiodic Rigorous Coupled Wave Analysis Technique*

Lahav A, Auslender M, Abdulhalim I. 2008. Sensitivity enhancement of guided-wave surface-plasmon resonance sensors. *Optics Letters* 33:2539-41

Lan TH, Chung YK, Tien CH. 2011. Broad Detecting Range of Objective-Based Surface Plasmon Resonance Sensor via Multilayer Structure. *Jpn. J. Appl. Phys.* 50

Lee K-S, Son JM, Jeong D-Y, Lee TS, Kim WM. 2010. Resolution Enhancement in Surface Plasmon Resonance Sensor Based on Waveguide Coupled Mode by Combining a Bimetallic Approach. *Sensors* 10:11390-9

Liedberg B, Lundstrom I, Stenberg E. 1993. Principles of biosensing with an extended coupling matrix and surface-plasmon resonance. *Sensors and Actuators B-Chemical* 11:63-72

Liu N, Weiss T, Mesch M, Langguth L, Eigenthaler U, et al. 2010. Planar Metamaterial Analogue of Electromagnetically Induced Transparency for Plasmonic Sensing. *Nano Letters* 10:1103-7

Mansfield SM, Kino GS. 1990. Solid immersion microscope. *Applied Physics Letters* 57:2615-6

McGuirk M, Carniglia CK. 1977. An angular spectrum representation approach to the Goos-Hänchen shift. *J. Opt. Soc. Am.* 67:103-7

Migdall AL, Roop B, Zheng YC, Hardis JE, Xia GJ. 1990. Use of heterodyne-detection to measure optical transmittance over a wide-range. *Applied Optics* 29:5136-44

Moh KJ, Yuan XC, Bu J, Zhu SW, Gao BZ. 2008. Surface plasmon resonance imaging of cell-substrate contacts with radially polarized beams. *Optics Express* 16:20734-41

Moharam MG, Grann EB, Pommet DA, Gaylord TK. 1995a. Formulation for stable and efficient implementation of the rigorous coupled-wave analysis of binary gratings. *Journal of the Optical Society of America a-Optics Image Science and Vision* 12:1068-76

Moharam MG, Pommet DA, Grann EB, Gaylord TK. 1995b. Stable implementation of the rigorous coupled-wave analysis for surface-relief gratings - enhanced transmittance matrix approach. *Journal of the Optical Society of America a-Optics Image Science and Vision* 12:1077-86

Mulvaney P. 1996. Surface plasmon spectroscopy of nanosized metal particles. *Langmuir* 12:788-800

Nelson BP, Grimsrud TE, Liles MR, Goodman RM, Corn RM. 2001. Surface plasmon resonance imaging measurements of DNA and RNA hybridization adsorption onto DNA microarrays. *Analytical Chemistry* 73:1-7

Nemova G, Kabashin AV, Kashyap R. 2008. Surface plasmon-polariton Mach-Zehnder refractive index sensor. *J. Opt. Soc. Am. B* 25:1673-7

Nenninger GG, Piliarik M, Homola J. 2002. Data analysis for optical sensors based on spectroscopy of surface plasmons. *Measurement Science & Technology* 13:2038-46

Nishiuma S, Handa Y, Imamura T, Ogino M, Yamada T, et al. 2008. Localized surface plasmon resonant metal nanostructures as refractive index sensors. *Jpn. J. Appl. Phys.* 47:1828-32

Notcovich AG, Zhuk V, Lipson SG. 2000a. Surface plasmon resonance phase imaging. *Applied Physics Letters* 76:1665-7

Notcovich AG, Zhuk V, Lipson SG, Osa. 2000b. Surface plasmon resonance phase imaging. In *Biomedical Topical Meetings, Technical Digest*, pp. 444-6

Novotny L, Hecht B, Pohl DW. 1997. Interference of locally excited surface plasmons. *Journal of Applied Physics* 81:1798-806

Offside MJ, Somekh MG, See CW. 1989. Common path scanning heterodyne optical profilometer for absolute phase measurement. *Applied Physics Letters* 55:2051-3

Otto A. 1968. Excitation of nonradiative surface plasma waves in silver by method of frustrated total reflection. *Zeitschrift Fur Physik* 216:398-&

Pang L, Hwang GM, Slutsky B, Fainman Y. 2007. Spectral sensitivity of two-dimensional nanohole array surface plasmon polariton resonance sensor. *Applied Physics Letters* 91

Parmon W, Bertoni HL. 1979. Ray interpretation of the material signature in the acoustic microscope. *Electronics Letters* 15:684-6

Pechprasarn S, Somekh MG. 2012. Surface plasmon microscopy: resolution, sensitivity and crosstalk. *Journal of Microscopy* 246:287-97

Pendry JB. 2000. Negative refraction makes a perfect lens. *Physical Review Letters* 85:3966-9

Porto JA, Garcia-Vidal FJ, Pendry JB. 1999. Transmission resonances on metallic gratings with very narrow slits. *Physical Review Letters* 83:2845-8

Raether H. 1988. Surface-plasmons on smooth and rough surfaces and on gratings. *Springer Tracts in Modern Physics* 111:1-133

Reitz JR. 1993. *Foundations of electromagnetic theory*. Addison-Wesley: Reading, Mass. ;

Ritchie RH, Arakawa ET, Cowan JJ, Hamm RN. 1968. Surface-plasmon resonance effect in grating diffraction. *Physical Review Letters* 21:1530-&

Roh S, Chung T, Lee B. 2010. Overview of plasmonic sensors and their design methods. In *Advanced Sensor Systems and Applications Iv*, ed. BLYWABXFXZL Culshaw

Roh S, Chung T, Lee B. 2011. Overview of the Characteristics of Micro- and Nano-Structured Surface Plasmon Resonance Sensors. *Sensors* 11:1565-88

Roland T, Berguiga L, Elezgaray J, Argoul F. 2010. Scanning surface plasmon imaging of nanoparticles. *Physical Review B* 81

Rothenhausler B, Knoll W. 1988. Surface-plasmon microscopy. *Nature* 332:615-7

Rothenhäusler B, Rabe J, Korpiun P, Knoll W. 1984. On the decay of plasmon surface polaritons at smooth and rough Ag-air interfaces: A reflectance and photo-acoustic study. *Surface Science* 137:373-83

Schasfoort RBM, Tudos AJ. 2008. *Handbook of surface plasmon resonance*: Royal Society of Chemistry

See CW, Somekh MG, Holmes RD. 1996. Scanning optical microellipsometer for pure surface profiling. *Applied Optics* 35:6663-8

Shatalin SV, JuŠKaitis R, Tan JB, Wilson T. 1995. Reflection conoscopy and micro-ellipsometry of isotropic thin film structures. *Journal of Microscopy* 179:241-52

Shen J, Liu S, Cao R, Fan X, Du J, et al. 2011. Magnetic surface plasmon-induced tunable photonic bandgaps in two-dimensional magnetic photonic crystals. *Applied Physics a-Materials Science & Processing* 105:789-93

Smolyaninov, II, Davis CC, Elliott J, Zayats AV. 2005a. Resolution enhancement of a surface immersion microscope near the plasmon resonance. *Optics Letters* 30:382-4

Smolyaninov, II, Elliott J, Zayats AV, Davis CC. 2005b. Far-field optical microscopy with a nanometer-scale resolution based on the in-plane image magnification by surface plasmon polaritons. *Physical Review Letters* 94

Somekh M. 2007. Surface Plasmon and Surface Wave Microscopy. *Springer series : Optical Imaging and Microscopy* 87:53

Somekh M, Pechprasarn S. 2010. Surface plasmon microscopy resolution vs sensitivity. *IEEEXplore Photonics Global Conference (PGC), 2010*

Somekh MG. 1992. Depth discrimination in scanned heterodyne microscope systems. *Journal of Microscopy-Oxford* 168:131-51

Somekh MG, Bertoni HL, Briggs GAD, Burton NJ. 1985. A Two-Dimensional Imaging Theory Of Surface Discontinuities With The Scanning Acoustic Microscope. *Proceedings of the Royal Society of London Series a-Mathematical Physical and Engineering Sciences* 401:29-51

Somekh MG, Hsu K, Pitter MC. 2011. Effect of processing strategies on the stochastic transfer function in structured illumination microscopy. *Journal of the Optical Society of America a-Optics Image Science and Vision* 28:1925-34

Somekh MG, Liu SG, Velinov TS, See CW. 2000a. High-resolution scanning surface-plasmon microscopy. *Applied Optics* 39:6279-87

Somekh MG, Liu SG, Velinov TS, See CW. 2000b. Optical V(z) for high-resolution 2 pi surface plasmon microscopy. *Optics Letters* 25:823-5

Somekh MG, See CW, Goh J. 2000c. Wide field amplitude and phase confocal microscope with speckle illumination. *Optics Communications* 174:75-80

Somekh MG, Stabler G, Liu S, Zhang J, See CW. 2009. Wide-field high-resolution surface-plasmon interference microscopy. *Optics Letters* 34:3110-2

Stabler G, Somekh MG, See CW. 2004. High-resolution wide-field surface plasmon microscopy. *Journal of Microscopy-Oxford* 214:328-33

Stoilov G, Dragostinov T. 1997. Phase-stepping interferometry: Five-frame algorithm with an arbitrary step. *Optics and Lasers in Engineering* 28:61-9

Talaat H, Bucaro JA, Huang W-S, Macdiarmid AG. 1985. Photoacoustic detection of plasmon surface polaritons in heavily doped polyacetylene films. *Synthetic Metals* 10:245-53

Tamir T, Burke JJ, Stegeman GI. 1985. Excitation of surface-plasmon modes along thin metal-films. *Journal of the Optical Society of America a-Optics Image Science and Vision* 2:P45-P

Tan. 2011. High resolution angle scanning surface plasmon resonance microscopy and its application to bio-molecular interactions. *Ph.D thesis, University of Nottingham*

Tanaka T, Yamamoto S. 2003. Laser-Scanning Surface Plasmon Polariton Resonance Microscopy with Multiple Photodetectors. *Appl. Opt.* 42:4002-7

Terris BD, Mamin HJ, Rugar D, Studenmund WR, Kino GS. 1994. Near-field optical-data storage using a solid immersion lens. *Applied Physics Letters* 65:388-90

Velinov T, Somekh MG, Liu S. 1999. Direct far-field observation of surface-plasmon propagation by photoinduced scattering. *Applied Physics Letters* 75:3908-10

Wan DH, Chen HL, Tseng SC, Wang LA, Chen YP. 2010. One-Shot Deep-UV Pulsed-Laser-Induced Photomodification of Hollow Metal Nanoparticles for High-Density Data Storage on Flexible Substrates. *ACS Nano* 4:165-73

Wink T, vanZuilen SJ, Bult A, vanBennekom WP. 1997. Self-assembled monolayers for biosensors. *Analyst* 122:R43-R50

Wizinowich PL. 1990. Phase shifting interferometry in the presence of vibration: a new algorithm and system. *Appl. Opt.* 29:3271-9

Wood RW. 1902. On a remarkable case of uneven distribution of light in a diffraction grating spectrum. *Philos. Mag.* 4:396-402

Wu X, Zhang J, Chen J, Zhao C, Gong Q. 2009. Refractive index sensor based on surface-plasmon interference. *Optics Letters* 34:392-4

Yang T, Ho HP. 2009. Computational investigation of nanohole array based SPR sensing using phase shift. *Optics Express* 17:11205-16

Yeatman E, Ash EA. 1987. Surface-plasmon microscopy. *Electronics Letters* 23:1091-2

Yeatman EM. 1996. Resolution and sensitivity in surface plasmon microscopy and sensing. *Biosensors & Bioelectronics* 11:635-49

Yoshita M, Sasaki T, Baba M, Akiyama H. 1998. Application of solid immersion lens to high-spatial resolution photoluminescence imaging of GaAs quantum wells at low temperatures. *Applied Physics Letters* 73:635-7

Zhang B, Pechprasarn S, Zhang J, Somekh MG. 2012. Confocal surface plasmon microscopy with pupil function engineering. *Opt. Express* 20:7388-97

Zhang DG, Moh KJ, Yuan XC. 2010. Surface plasmon-coupled emission from shaped PMMA films doped with fluorescence molecules. *Optics Express* 18:12185-90

Zhang J, Pitter MC, Liu S, See C, Somekh MG. 2006. Surface-plasmon microscopy with a two-piece solid immersion lens: bright and dark fields. *Applied Optics* 45:7977-86

Zhang J, See CW, Somekh MG. 2007. Imaging performance of widefield solid immersion lens microscopy. *Applied Optics* 46:4202-8

Zhou H, Sheppard CJR. 1997. Aberration measurement in confocal microscopy: Phase retrieval from a single intensity measurement. *Journal of Modern Optics* 44:1553-61

Appendix A

Rigorous wave coupled analysis 1D and 2D grating theory

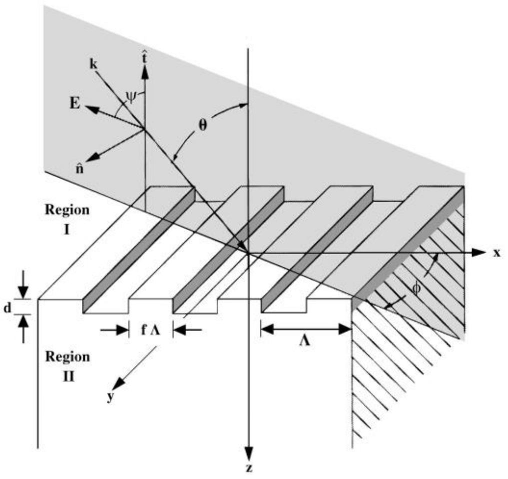


Figure A.1 shows variables definition (Moharam et al 1995a)

For a plane of incidence with arbitrary angle of incidence θ, azimuthal angle ϕ and electric field angle with respect to plane of incidence ψ, unit vector pointing direction of electric field can be written as:

$$\boldsymbol{u}=[\cos\psi\cos\theta\cos\phi-\sin\psi\sin\phi, -\cos\psi\cos\theta\sin\phi-\sin\psi\cos\phi, \cos\psi\sin\theta] \quad \text{[A.1]}$$

Electric field incidence can be then written as an amplitude product of the unit vector described in [A.1] as:

$$E_{incidence}=\boldsymbol{u}\, e^{-jk_0 n_I(\sin\theta\cos\phi\,\hat{\boldsymbol{x}}+\sin\theta\sin\phi\,\hat{\boldsymbol{y}}+\cos\theta\,\hat{\boldsymbol{z}})} \qquad \text{[A.2]}$$

where $k_0$is the free space wave vector and $n_I$is refractive index of the first layer material.

There are three regions in this calculation, namely homogeneous layer region I (z<0), grating region ($0 \leq z \leq d$) and homogeneous layer region II (z>d). The electric solution for the three regions can be expressed as:

For z<0, $\boldsymbol{E_I} = \boldsymbol{E_{incidence}} + \sum_i \boldsymbol{R_i} e^{-jk_0 n_I (k_{xi}\hat{\boldsymbol{x}} + k_y \hat{\boldsymbol{y}} + k_{I,zi}\hat{\boldsymbol{z}})}$ [A.3]

where $\boldsymbol{E_I}$ is electric field vector reflected to the region I, $\boldsymbol{R_i}$ reflected electric field vector of the ith diffracted order. k_{xi} is the wave vector along x axis of the ith diffracted order. k_{xi} is the wave vector along y axis. $k_{I,zi}$ is the wave vector along z axis of the ith diffracted order.

$$k_{xi} = k_0 (n_I sin\theta\ cos\phi - i\frac{\lambda_0}{\Lambda}) \qquad \text{[A.4]}$$

$$k_y = k_0 n_I sin\theta\ sin\phi \qquad \text{[A.5]}$$

$$k_{I,zi} = \sqrt{(k_0 n_I)^2 - {k_{xi}}^2 - {k_y}^2} \qquad \text{[A.6]}$$

where λ_0 is free space wavelength and Λ is grating period.

For z>d, $\boldsymbol{E_{II}} = \sum_i \boldsymbol{T_i} e^{-jk_0 n_1 (k_{xi} x + k_y y + k_{II,zi}(z-d))}$ [A.7]

where $\boldsymbol{E_{II}}$ is electric field vector transmitted to the region II, $\boldsymbol{T_i}$ transmitted electric field vector of the i^{th} diffracted order. $k_{II,zi}$ is the wave vector along z axis of the ith diffracted order.

$$k_{II,zi} = \sqrt{(k_0 n_{II})^2 - {k_{xi}}^2 - {k_y}^2} \qquad \text{[A.8]}$$

For $0 \leq z \leq d$, $\boldsymbol{E_g} = \sum_i (S_{xi}(z)\boldsymbol{x} + S_{yi}(z)\boldsymbol{y} + S_{zi}(z)\boldsymbol{z}) e^{-jk_0 n_I (k_{xi} x + k_y y)}$

[A.9]

$$\boldsymbol{H} = -j\sqrt{\frac{\varepsilon_0}{\mu_0}} \sum_i (U_{xi}(z)\boldsymbol{x} + U_{yi}(z)\boldsymbol{y} + U_{zi}(z)\boldsymbol{z}) e^{-jk_0 n_I (k_{xi} x + k_y y)}$$

[A.10]

where S_{xi}, $S_{yi}$and $S_{zi}$are the normalized electric field amplitudes of the *i*th diffracted order. U_{xi}, $U_{yi} and\ U_{zi}$ are the normalized magnetic field amplitudes of the *i*th diffracted order.

The Maxwell's equations for the grating region can be written as:

$$\boldsymbol{\nabla} \times \boldsymbol{E_g} = -j\omega\mu_0\boldsymbol{H_g} \qquad [A.11]$$

$$\boldsymbol{\nabla} \times \boldsymbol{H_g} = j\omega\varepsilon_0\varepsilon_r(x)\boldsymbol{E_g} \qquad [A.12]$$

where $\varepsilon_r(x)$ is relative permittivity at x harmonic number.

$$\varepsilon_r(x) = \sum_{h=-n}^{n} \varepsilon_h e^{j\frac{2\pi h}{\Lambda}} \qquad [A.13]$$

Where ε_h is the hth Fourier component of the relative permittivity in the grating region, which is complex for lossy or non-symmetric dielectric grating and *n* is the number of diffracted order used.

We can then substitute [A.9] and [A.10] into [A.11] and [A.12] and obtain the following matrices:

$$\begin{bmatrix} \partial^2\boldsymbol{S_y}/\partial z'^2 \\ \partial^2\boldsymbol{S_x}/\partial z'^2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{K_x}^2 + \boldsymbol{DE} & \boldsymbol{K_y}[\boldsymbol{E}^{-1}\boldsymbol{K_x}\boldsymbol{E} - \boldsymbol{K_x}] \\ \boldsymbol{K_x}[\boldsymbol{E}^{-1}\boldsymbol{K_y}\boldsymbol{E} - \boldsymbol{K_y}] & \boldsymbol{K_y}^2 + \boldsymbol{BE} \end{bmatrix} \begin{bmatrix} \boldsymbol{S_y} \\ \boldsymbol{S_x} \end{bmatrix} \qquad [A.14]$$

$$\begin{bmatrix} \partial^2\boldsymbol{U_y}/\partial z'^2 \\ \partial^2\boldsymbol{U_x}/\partial z'^2 \end{bmatrix} = \begin{bmatrix} \boldsymbol{K_y}^2 + \boldsymbol{EB} & [\boldsymbol{K_x} - \boldsymbol{E}\boldsymbol{K_x}\boldsymbol{E}^{-1}]\boldsymbol{K_y} \\ [\boldsymbol{K_y} - \boldsymbol{E}\boldsymbol{K_y}\boldsymbol{E}^{-1}]\boldsymbol{K_x} & \boldsymbol{K_x}^2 + \boldsymbol{ED} \end{bmatrix} \begin{bmatrix} \boldsymbol{U_y} \\ \boldsymbol{U_x} \end{bmatrix} \qquad [A.15]$$

Where $z' = k_0 z$. $\boldsymbol{S_x}$ is a vector consisting of S_{xi} values. $\boldsymbol{S_y}$ is a vector consisting of S_{yi} values. $\boldsymbol{K_x}$ is a diagonal matrix with the (i,i) element being equal to $\frac{k_{xi}}{k_0}$. $\boldsymbol{K_y}$ is a diagonal matrix with the (i,i) element being equal to $\frac{k_{yi}}{k_0}$. $\boldsymbol{E}$ is the matrix formed by the permittivity harmonic components, with the (i,p) element being equal to $\varepsilon(i-p)$.

$$\boldsymbol{B} = \boldsymbol{K_x}\boldsymbol{E}^{-1}\boldsymbol{K_x} - \boldsymbol{I} \qquad [A.16]$$

$$\boldsymbol{D} = \boldsymbol{K_y}\boldsymbol{E}^{-1}\boldsymbol{K_y} - \boldsymbol{I} \qquad [A.17]$$

For the 1D grating case, ***Ky*** is simply an unity matrix, so the equations [A.14] and [A.15] can be then reduced to.

$$\left[\frac{\partial^2\boldsymbol{U_x}}{\partial z'^2}\right] = [\boldsymbol{K_y}^2\boldsymbol{I} + \boldsymbol{A}][\boldsymbol{U_x}] \qquad [A.18]$$

$\left[\frac{\partial^2 \boldsymbol{S_x}}{\partial z'^2}\right] = [\boldsymbol{K_y}^2\boldsymbol{I} + \boldsymbol{BE}][\boldsymbol{S_x}]$ [A.19]

Where $\boldsymbol{A} = \boldsymbol{K_x}^2 - \boldsymbol{E}$

We can then solve the Eigen-value problem for the matrices in eq[A.14] and [A.15] for the 2D case or [A.18] and [A.19] for the 1D case.

The terms $U_{xi}, U_{yi}, S_{xi}\, , S_{yi}$ can be then reduced expanded as Fourier expansion:

$U_{xi}(z) = \sum_{m=1}^{n} w_{1,i,m}\{-c_{1,m}{}^{+}e^{-k_0 q_{1,m} z} + c_{1,m}{}^{-}e^{k_0 q_{1,m}(z-d)}\}$ [A.20]

$S_{xi}(z) = \sum_{m=1}^{n} w_{2,i,m}\{-c_{2,m}{}^{+}e^{-k_0 q_{2,m} z} + c_{2,m}{}^{-}e^{k_0 q_{2,m}(z-d)}\}$ [A.21]

$S_{yi}(z) =$

$\sum_{m=1}^{n} v_{11,i,m}\{-c_{1,m}{}^{+}e^{-k_0 q_{1,m} z} + c_{1,m}{}^{-}e^{k_0 q_{1,m}(z-d)}\} +$

$\sum_{m=1}^{n} v_{12,i,m}\{-c_{2,m}{}^{+}e^{-k_0 q_{2,m} z} + c_{2,m}{}^{-}e^{k_0 q_{2,m}(z-d)}\}$ [A.22]

$U_{yi}(z) =$

$\sum_{m=1}^{n} v_{21,i,m}\{-c_{1,m}{}^{+}e^{-k_0 q_{1,m} z} + c_{1,m}{}^{-}e^{k_0 q_{1,m}(z-d)}\} +$

$\sum_{m=1}^{n} v_{22,i,m}\{-c_{2,m}{}^{+}e^{-k_0 q_{2,m} z} + c_{2,m}{}^{-}e^{k_0 q_{2,m}(z-d)}\}$ [A.23]

Where $w_{1,i,m}$ and $q_{1,m}$ are the elements of the eigenvector matrix $\boldsymbol{W_1}$ of equation [A.14] for 2D case and [A.18] for 2D case and the positive square root of the eigenvalues respectively. $w_{2,i,m}$ and $q_{2,m}$ are the elements of the eigenvector matrix $\boldsymbol{W_2}$ of equation [A.15] for 2D case and [A.19] for 2D case and the positive square root of the eigenvalues respectively. $c_{1,m}{}^{+}$, $c_{1,m}{}^{-}$, $c_{2,m}{}^{+}$ and $c_{2,m}{}^{-}$ are unknown constants to be determined by boundary conditions. The quantities $v_{11,i,m}, v_{21,i,m}, v_{12,i,m}, v_{22,i,m}$ are the elements of the matrices $\boldsymbol{V_{11}}, \boldsymbol{V_{12}}, \boldsymbol{V_{21}}$ and $\boldsymbol{V_{22}}$, which are expressed as:

$\boldsymbol{V_{11}} = \boldsymbol{A}^{-1}\boldsymbol{W_1}\boldsymbol{Q_1}$ [A.24]

$\boldsymbol{V_{12}} = (k_y/k_0)\boldsymbol{A}^{-1}\boldsymbol{K_x}\boldsymbol{W_2}$ [A.25]

$\boldsymbol{V_{21}} = (k_y/k_0)\boldsymbol{B}^{-1}\boldsymbol{K_x}\boldsymbol{E}^{-1}\boldsymbol{W_1}$ [A.26]

$\boldsymbol{V_{22}} = \boldsymbol{B}^{-1}\boldsymbol{W_2}\boldsymbol{Q_2}$ [A.27]

Where $\boldsymbol{Q_1}$ and $\boldsymbol{Q_2}$ are diagonal matrices with the diagonal element $q_{1,m}$ and $q_{2,m}$ respectively.

By matching the tangential electric field and magnetic fields at the surface of grating *z=0*, the matching conditions are given by:

$$\sin\psi\ \delta_{i0} + \boldsymbol{R_{s,i}} = \cos\varphi_i\, S_{xi}(0) \qquad [A.28]$$

$$j\left[sin\psi\, n_I\, cos\theta - \left(\frac{k_{I,zi}}{k_0}\right)\boldsymbol{R_{s,i}}\right] = -[cos\varphi_i\, U_{xi}(0) + sin\varphi_i\, U_{yi}(0)] \qquad [A.29]$$

$$cos\psi\, cos\theta - j\left[\frac{k_{I,si}}{k_0 {n_I}^2}\right]\boldsymbol{R_{p,i}} = cos\varphi_i S_{xi}(0) + sin\varphi_i S_{yi}(0) \qquad [A.30]$$

$$-jn_I cos\psi + \boldsymbol{R_{p,i}} = -[\cos\varphi_i\, U_{yi}(0) - sin\varphi_i U_{xi}(0)] \qquad [A.31]$$

Where $\boldsymbol{R_{s,i}}$ and $\boldsymbol{R_{p,i}}$ are the components of the amplitude of electric and the magnetic field vectors normal to the diffraction plane respectively and $\varphi_i = tan^{-1}(k_y/k_{xi})$. Equations [A.28] to [A.31] can be then rewritten in a matrix from as:

$$\begin{bmatrix} sin\psi\, \delta_{i0} \\ jsin\psi\, n_I\, cos\theta\, \delta_{i0} \\ -jcos\psi\, n_I \delta_{i0} \\ cos\psi\, cos\theta\, \delta_{i0} \end{bmatrix} + \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \\ -\boldsymbol{jY_I} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I} \\ \boldsymbol{0} & -\boldsymbol{jZ_I} \end{bmatrix} \begin{bmatrix} \boldsymbol{R_s} \\ \boldsymbol{R_p} \end{bmatrix} =$$

$$\begin{bmatrix} \boldsymbol{V_{ss}} & \boldsymbol{V_{sp}} & \boldsymbol{V_{ss}X_1} & \boldsymbol{V_{sp}X_2} \\ \boldsymbol{W_{ss}} & \boldsymbol{W_{sp}} & -\boldsymbol{W_{ss}X_1} & -\boldsymbol{W_{sp}X_2} \\ \boldsymbol{W_{ps}} & \boldsymbol{W_{pp}} & -\boldsymbol{W_{ps}X_1} & -\boldsymbol{W_{pp}X_2} \\ \boldsymbol{V_{ps}} & \boldsymbol{V_{pp}} & \boldsymbol{V_{ps}X_1} & \boldsymbol{V_{pp}X_2} \end{bmatrix} \begin{bmatrix} \boldsymbol{c_1^+} \\ \boldsymbol{c_1^-} \\ \boldsymbol{c_2^+} \\ \boldsymbol{c_2^-} \end{bmatrix} \quad [A.32]$$

Where $\boldsymbol{X_1}$ and $\boldsymbol{X_2}$ are diagonal matrices with the diagonal elements $e^{(-k_0 q_{1,m} d)}$ and $e^{(-k_0 q_{2,m} d)}$, respectively.

$$\boldsymbol{V_{ss}} = \boldsymbol{F_c V_{11}} \qquad [A.33]$$

$$\boldsymbol{W_{pp}} = \boldsymbol{F_c V_{22}} \qquad [A.34]$$

$$\boldsymbol{V_{ss}} = \boldsymbol{F_c V_{11}} \qquad [A.35]$$

$$\boldsymbol{W_{ss}} = \boldsymbol{F_c W_1} + \boldsymbol{F_s V_{21}} \quad [A.36]$$

$$\boldsymbol{V_{pp}} = \boldsymbol{F_c W_2} + \boldsymbol{F_s V_{12}} \quad [A.37]$$

$$\boldsymbol{V_{sp}} = \boldsymbol{F_c V_{12}} - \boldsymbol{F_c W_2} \qquad [A.38]$$

$$\boldsymbol{W_{ps}} = \boldsymbol{F_c V_{21}} + \boldsymbol{F_s W_1} \quad [A.39]$$

$$\boldsymbol{W_{sp}} = \boldsymbol{F_s V_{22}} \qquad [A.40]$$

$\boldsymbol{V_{ps}} = \boldsymbol{F_s V_{11}}$ [A.41]

Where $\boldsymbol{F_c}$ and $\boldsymbol{F_s}$ are diagonal matrices with the diagonal elements of $cos\varphi_i$ and $sin\varphi_i$ respectively.

For the case $z=d$, the matching conditions are given by:

$\cos\varphi_i\, S_{yi}(d) - sin\varphi_i S_{xi}(d) = \boldsymbol{T_{s,i}}$ [A.42]

$-[\cos\varphi_i\, U_{xi}(d) - sin\varphi_i U_{yi}(d)] = j(k_{I,zi}/k_0)\boldsymbol{T}_{s,i}$ [A.43]

$-[\cos\varphi_i\, U_{yi}(d) + sin\varphi_i U_{xi}(d)] = \boldsymbol{T_{p,i}}$ [A.44]

$\cos\varphi_i\, S_{xi}(d) + sin\varphi_i S_{yi}(d) = j(k_{I,zi}/k_0 n_{II}^2)\boldsymbol{T}_{p,i}$ [A.45]

Where $\boldsymbol{T_{s,i}}$ and $\boldsymbol{T_{p,i}}$ are the components of the amplitude of the electric and magnetic fields vectors normal to the diffraction plane. The equations [A.42] to [A.45] can be then rewritten as:

$$\begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{jY_{II}} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I} \\ \boldsymbol{0} & \boldsymbol{jZ_{II}} \end{bmatrix} \begin{bmatrix} \boldsymbol{T_s} \\ \boldsymbol{T_p} \end{bmatrix} = \begin{bmatrix} \boldsymbol{V_{ss}X_1} & \boldsymbol{V_{sp}X_2} & \boldsymbol{V_{ss}} & \boldsymbol{V_{sp}} \\ \boldsymbol{W_{ss}X_1} & \boldsymbol{W_{sp}X_2} & -\boldsymbol{W_{ss}} & -\boldsymbol{W_{sp}} \\ \boldsymbol{W_{ps}X_1} & \boldsymbol{W_{pp}X_2} & -\boldsymbol{W_{ps}} & -\boldsymbol{W_{pp}} \\ \boldsymbol{V_{ps}X_1} & \boldsymbol{V_{pp}X_2} & \boldsymbol{V_{ps}} & \boldsymbol{V_{pp}} \end{bmatrix} \begin{bmatrix} \boldsymbol{c_1^+} \\ \boldsymbol{c_1^-} \\ \boldsymbol{c_2^+} \\ \boldsymbol{c_2^-} \end{bmatrix} \qquad [A.46]$$

Where $\delta_{i0} = 1$ for $i = 0$ and $\delta_{i0} = 0$ for $i \neq 0$. $\boldsymbol{Y_I}$ and $\boldsymbol{Y_{II}}$ are diagonal matrices with the diagonal elements $k_{I,zi}/k_0$ and $k_{II,zi}/k_0$ respectively. $\boldsymbol{Z_I}$ and $\boldsymbol{Z_{II}}$ are diagonal matrices with the diagonal elements $k_{I,zi}/k_0 n_I^2$ and $k_{II,zi}/k_0 n_{II}^2$ respectively.

The equation [A.32] and [A.46] can be solved simultaneous and give us the electric field amplitude (TE) and magnetic field amplitude (TM) normal to the diffraction planes for reflection and transmission.

Appendix B

Validation for all RCWA cases presented in this thesis

All the grating cases used in this thesis are listed in the table below. The number of diffracted orders in the table is 2 sided diffracted orders *(2N+1)* and the maximum error indicates the maximum error on the back focal plane calculation; this number is determined by calculating the full bfp of two consecutive numbers of diffracted orders to see whether the numerical solution from RCWA approaches a convergence value as discussed chapter 3.

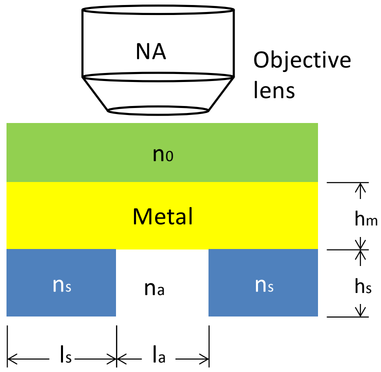


Figure B.1 shows symbols used in the table and their defnitions

NA	**n_0**	**Metal**	**h_m**	**n_s**	**n_a**	**l_s (nm)**	**l_a (nm)**	**h_s (nm)**	**Number of diffracted orders**	**Maximum error**
1.45	1.52	gold	50	2.023	1	1	7	20	61	1.78E-05
1.45	1.52	gold	50	2.023	1	2	6	20	53	3.96E-05
1.45	1.52	gold	50	2.023	1	3	5	20	51	4.93E-05
1.45	1.52	gold	50	2.023	1	4	4	20	51	4.38E-05
1.45	1.52	gold	50	2.023	1	5	3	20	51	6.36E-05
1.45	1.52	gold	50	2.023	1	6	2	20	53	7.80E-05
1.45	1.52	gold	50	2.023	1	7	1	20	61	9.30E-05
1.45	1.52	Aluminium	15	2.023	1	6	2	20	53	8.71E-05
1.65	1.78	gold	36	1.5	1.33	0.5	0.5	10	31	5.10E-05
1.65	1.78	gold	36	1.5	1.33	2	2	10	31	4.91E-05
1.65	1.78	gold	36	1.5	1.33	3	3	10	31	4.20E-06
1.65	1.78	gold	36	1.5	1.33	4	4	10	33	4.05E-05
1.65	1.78	gold	36	1.5	1.33	5	5	10	35	3.79E-05

1.65	1.78	gold	36	1.5	1.33	6	6	10	37	1.67E-05
1.65	1.78	gold	36	1.5	1.33	7	7	10	45	8.50E-05
1.65	1.78	gold	36	1.5	1.33	8	8	10	53	6.53E-05
1.65	1.78	gold	36	1.5	1.33	9	9	10	61	5.19E-05
1.65	1.78	gold	36	1.5	1.33	10	10	10	71	7.21E-05
1.65	1.78	gold	36	1.5	1.33	11	11	10	81	8.20E-06
1.65	1.78	gold	36	1.5	1.33	12	12	10	91	9.28E-05
1.65	1.78	gold	36	1.5	1.33	13	13	10	101	3.50E-06
1.65	1.78	gold	36	1.5	1.33	14	14	10	111	2.27E-05
1.65	1.78	gold	36	1.5	1.33	15	15	10	121	5.00E-05
1.65	1.78	gold	46	1.5	1.33	0.5	0.5	10	31	1.43E-05
1.65	1.78	gold	46	1.5	1.33	2	2	10	31	2.53E-05
1.65	1.78	gold	46	1.5	1.33	3	3	10	31	9.09E-05
1.65	1.78	gold	46	1.5	1.33	4	4	10	33	5.38E-05
1.65	1.78	gold	46	1.5	1.33	5	5	10	35	8.59E-05
1.65	1.78	gold	46	1.5	1.33	6	6	10	37	5.24E-05
1.65	1.78	gold	46	1.5	1.33	7	7	10	45	8.50E-05
1.65	1.78	gold	46	1.5	1.33	8	8	10	53	2.39E-05
1.65	1.78	gold	46	1.5	1.33	9	9	10	61	1.16E-05
1.65	1.78	gold	46	1.5	1.33	10	10	10	71	6.48E-05
1.65	1.78	gold	46	1.5	1.33	11	11	10	81	1.08E-05
1.65	1.78	gold	46	1.5	1.33	12	12	10	91	2.86E-05
1.65	1.78	gold	46	1.5	1.33	13	13	10	101	9.90E-06
1.65	1.78	gold	46	1.5	1.33	14	14	10	111	2.87E-05
1.65	1.78	gold	46	1.5	1.33	15	15	10	121	9.13E-05
1.49	1.52	gold	36	1.5	1.33	0.5	0.5	10	31	5.14E-05
1.49	1.52	gold	36	1.5	1.33	2	2	10	35	2.65E-05
1.49	1.52	gold	36	1.5	1.33	3	3	10	37	3.43E-05
1.49	1.52	gold	36	1.5	1.33	4	4	10	39	4.92E-05
1.49	1.52	gold	36	1.5	1.33	5	5	10	41	7.92E-05
1.49	1.52	gold	36	1.5	1.33	6	6	10	45	2.57E-05
1.49	1.52	gold	36	1.5	1.33	7	7	10	53	6.19E-05
1.49	1.52	gold	36	1.5	1.33	8	8	10	61	9.97E-05
1.49	1.52	gold	36	1.5	1.33	9	9	10	65	1.40E-06
1.49	1.52	gold	36	1.5	1.33	10	10	10	71	4.59E-05
1.49	1.52	gold	36	1.5	1.33	11	11	10	81	5.21E-05
1.49	1.52	gold	36	1.5	1.33	12	12	10	91	8.78E-05
1.49	1.52	gold	36	1.5	1.33	13	13	10	101	1.52E-05
1.49	1.52	gold	36	1.5	1.33	14	14	10	111	1.86E-05
1.49	1.52	gold	36	1.5	1.33	15	15	10	121	9.30E-05
1.49	1.52	gold	46	1.5	1.33	0.5	0.5	10	31	7.85E-05
1.49	1.52	gold	46	1.5	1.33	2	2	10	35	1.60E-05

1.49	1.52	gold	46	1.5	1.33	3	3	10	37	9.37E-05
1.49	1.52	gold	46	1.5	1.33	4	4	10	39	6.60E-06
1.49	1.52	gold	46	1.5	1.33	5	5	10	41	3.81E-05
1.49	1.52	gold	46	1.5	1.33	6	6	10	45	6.93E-05
1.49	1.52	gold	46	1.5	1.33	7	7	10	53	2.80E-06
1.49	1.52	gold	46	1.5	1.33	8	8	10	61	8.88E-05
1.49	1.52	gold	46	1.5	1.33	9	9	10	65	9.17E-05
1.49	1.52	gold	46	1.5	1.33	10	10	10	71	4.39E-05
1.49	1.52	gold	46	1.5	1.33	11	11	10	81	2.29E-05
1.49	1.52	gold	46	1.5	1.33	12	12	10	91	6.10E-05
1.49	1.52	gold	46	1.5	1.33	13	13	10	101	9.26E-05
1.49	1.52	gold	46	1.5	1.33	14	14	10	111	6.36E-05
1.49	1.52	gold	46	1.5	1.33	15	15	10	121	5.50E-06
1.49	1.52	gold	36	1.5	1.3	5	5	10	41	9.93E-05
1.49	1.52	gold	36	1.5	1.35	5	5	10	41	2.96E-05
1.49	1.52	gold	36	1.5	1.4	5	5	10	41	2.41E-05
1.49	1.52	gold	36	1.5	1.45	5	5	10	41	3.86E-05
1.49	1.52	gold	36	1.5	1.5	5	5	10	41	2.62E-05
1.49	1.52	gold	36	1.5	1.55	5	5	10	41	2.25E-05
1.49	1.52	gold	36	1.5	1.6	5	5	10	41	3.69E-05
1.49	1.52	gold	36	1.5	1.65	5	5	10	41	1.75E-05
1.49	1.52	gold	36	1.5	1.7	5	5	10	41	7.69E-05
1.49	1.52	gold	46	1.5	1.3	5	5	10	41	5.84E-05
1.49	1.52	gold	46	1.5	1.35	5	5	10	41	3.15E-05
1.49	1.52	gold	46	1.5	1.4	5	5	10	41	9.05E-05
1.49	1.52	gold	46	1.5	1.45	5	5	10	41	8.99E-05
1.49	1.52	gold	46	1.5	1.5	5	5	10	41	4.35E-05
1.49	1.52	gold	46	1.5	1.55	5	5	10	41	7.10E-06
1.49	1.52	gold	46	1.5	1.6	5	5	10	41	6.45E-05
1.49	1.52	gold	46	1.5	1.65	5	5	10	41	8.39E-05
1.49	1.52	gold	46	1.5	1.7	5	5	10	41	9.04E-05
1.65	1.78	gold	36	1.5	1.3	5	5	10	41	6.21E-05
1.65	1.78	gold	36	1.5	1.35	5	5	10	41	4.25E-05
1.65	1.78	gold	36	1.5	1.4	5	5	10	41	6.00E-07
1.65	1.78	gold	36	1.5	1.45	5	5	10	41	9.71E-05
1.65	1.78	gold	36	1.5	1.5	5	5	10	41	5.31E-05
1.65	1.78	gold	36	1.5	1.55	5	5	10	41	3.41E-05
1.65	1.78	gold	36	1.5	1.6	5	5	10	41	5.92E-05
1.65	1.78	gold	36	1.5	1.65	5	5	10	41	8.92E-05
1.65	1.78	gold	36	1.5	1.7	5	5	10	41	4.54E-05
1.65	1.78	gold	46	1.5	1.3	5	5	10	41	6.95E-05
1.65	1.78	gold	46	1.5	1.35	5	5	10	41	9.15E-05

1.65	1.78	gold	46	1.5	1.4	5	5	10	41	1.12E-05
1.65	1.78	gold	46	1.5	1.45	5	5	10	41	2.24E-05
1.65	1.78	gold	46	1.5	1.5	5	5	10	41	4.89E-05
1.65	1.78	gold	46	1.5	1.55	5	5	10	41	7.99E-05
1.65	1.78	gold	46	1.5	1.6	5	5	10	41	4.42E-05
1.65	1.78	gold	46	1.5	1.65	5	5	10	41	1.70E-05
1.65	1.78	gold	46	1.5	1.7	5	5	10	41	8.84E-05

Appendix C

List of metal choices for SP excitation

Material	Complex refractive index corresponding to the wavelengths of				typical thickness
	405 nm	532 nm	633 nm	830 nm	
Al	0.50047+i4.8978	0.938777+i6.4195	1.44947+i7.5387	2.71988+i8.2122	10-20 nm
Cu	1.18+i2.21	1.061429+i2.5927	0.24869+i3.4128	0.26146+i5.2871	15-25 nm
Ag	0.173+i2.0102	0.12927+i3.2012	0.13461+i3.9882	0.14604+i5.5259	40-55 nm
Au	1.64973+i1.9568	0.467+i2.4075	0.19683+i3.0905	0.18927+i5.4182	45-55 nm

Table C.1 shows complex refractive index corresponding to the wavelengths of 405 nm, 532 nm, 633 nm and 830 nm and their typical thicknesses for SP excitation.

Appendix D

Noise simulations to compare sensitivity between intensity detection and phase detection

At the absence of noise, intensity and phase detections can give us similar performance. However at the presence of noise the performance of the two detection methods give a significantly different performance as explained and reported by (Kabashin et al 2009b). Therefore in this section, I will provide a set of simulations to assess the performance of the two detection methods and determine the thinnest thickness that can be detected when they face with noises, such as photon shot noise noise due to vibrations and other noise sources. In this section, we will consider an overall noise effect of the system by simulating reflection responses with a random root mean square noise (RMS noise) of the reflection coefficients. The RMS noise level was modelled with a Gaussian distribution with mean value of μ and variance of σ^2 of electrical voltage level, *Gaussian*(μ,σ^2).

The signal to noise ratio for the electrical voltage is defined as:

$$SNR_{electrical\ signal} = \frac{\mu^2}{\sigma^2} \qquad \text{[A4.1]}$$

Note that the $\sqrt{SNR_{electrical\ signal}}$ term is directly proportional to signal to noise ratio (SNR) of the optical signal (Somekh *et al* 2011). It can be shown that when the variance σ^2 approaches the mean value μ; this condition is the condition for shot noise limit, which is the best case scenario for the noise simulation.

In this section, we consider effect of the noise on the both methods, where the mean is simply its reflection coefficients and variance $\sigma^2=0.03\mu^2$; this equivalents to the SNR of 33 (30.37 dBV).

The simulations were carried out based on the Kretchmann's configuration as shown in the figure D.1.

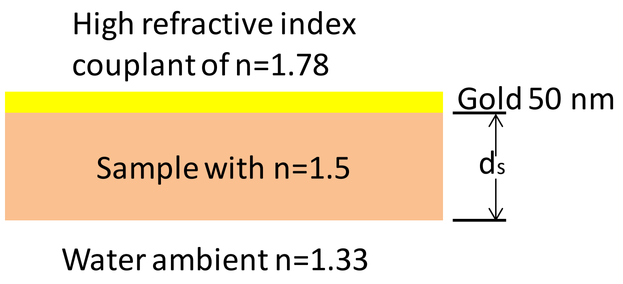


Figure D.1 shows Kretchmann's configuration modelled in this noise simulation

The thickness sensitivity of the system can be determined by varying the sample thickness d_s. At the absence of noise the amplitude detection and phase detection can give us the similar performance as shown in figure D.2.

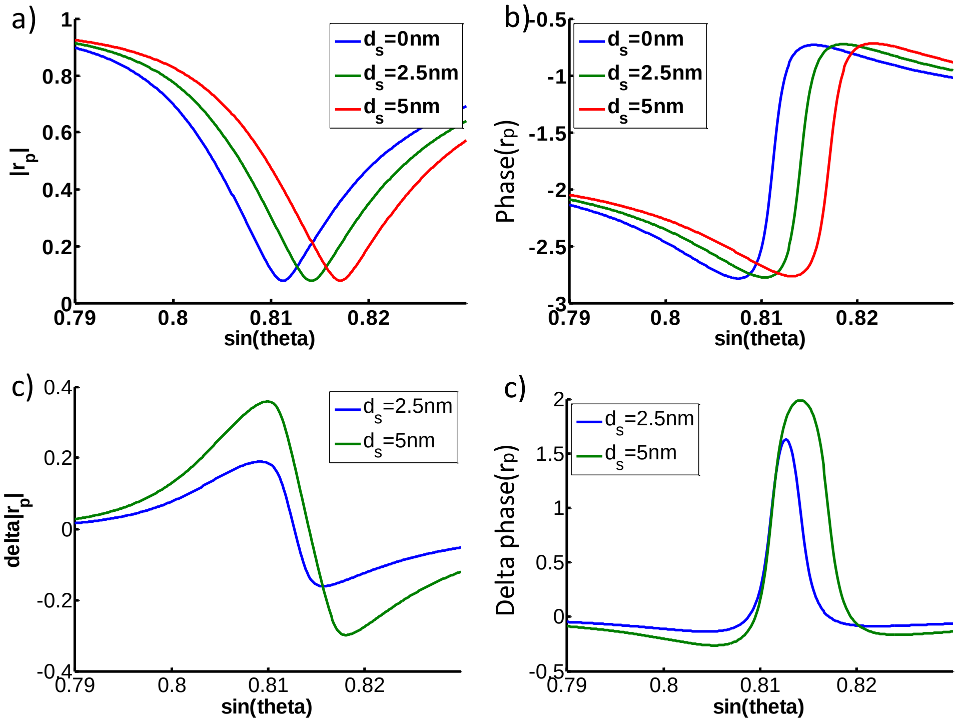
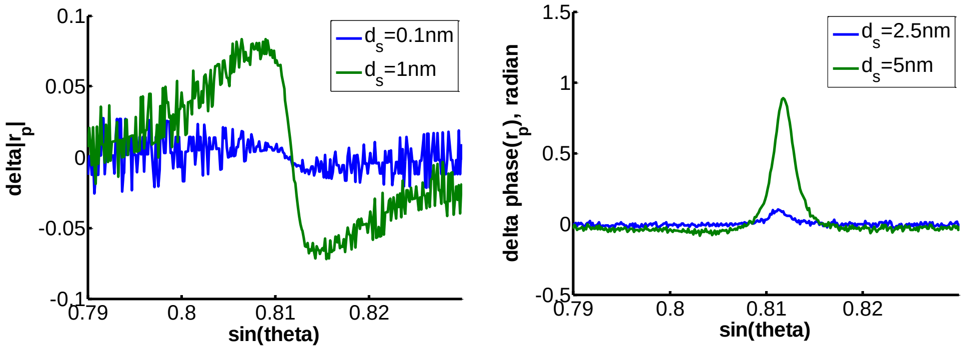


Figure D.2 shows reflection coefficients for $d_s=0nm$, $d_s=0.25nm$ and $d_s=0.5nm$

a) $|r_p|$ b) $phase(r_p)$ c) $\Delta|r_p| = |r_{p\ coated}|_{|}-|r_{p\ bare\ gold}|$ and d) $\Delta phase(r_p)=phase(r_{p\ coated)}-phase(r_{p\ bare\ gold})$

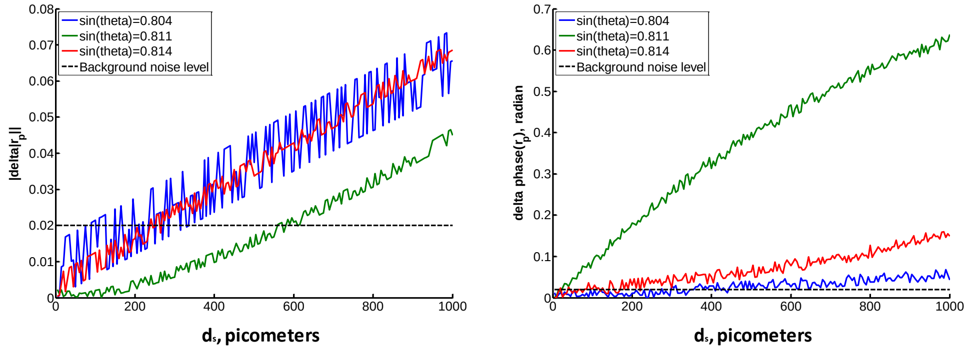
However if there is a presence of noise sources, the detection performance of the two methods are not the same as shown in figure D.3. From figure D.3, it can be seen that the amplitude detection cannot distinguish the detected signal from its noisy background level for the sample thickness of 0.1 nm, whereas the phase detection can still give a decent SNR level.



a) b)

Figure D.3 shows a) $\Delta|r_p| = |r_{p\ coated}|_{|}-|r_{p\ bare\ gold}|$ and b) $\Delta phase(r_p)=phase(r_{p\ coated)}-phase(r_{p\ bare\ gold})$ for 0.1 nm and 1 nm samples

Therefore the next task is to determine the minimum thickness that can be detected using the amplitude detection and phase detection and also compare the performance of each detection points. For the amplitude detection, $sin\theta=0.804$ (the maximum Δ|rp| in the figure D.2c) and $sin\theta=0.814$ (the minimum Δ|rp| in the figure D.2c) are the operating points included in this simulation. For the phase detection, the maximum phase change occurs at the plasmonic angle $sin\theta=0.811$. The detection performance of these three operating points as a function of thickness of the sample is shown in figure D.4.



a) b)

Figure D.4 shows a) $\Delta|r_p| = |r_{p\ coated}|_{|}-|r_{p\ bare\ gold}|$ and b) $\Delta phase(r_p)=phase(r_{p\ coated)}-phase(r_{p\ bare\ gold})$ as a function of d_s for the three operating points $sin\theta=0.804$, $sin\theta=0.811$ and $sin\theta=0.814$.

Figure D.4 tells us that the phase detection is more sensitive to the change in the sample thickness than the amplitude detection and it also has a higher SNR. The only disadvantage of the phase detection is that it is required an interferometer to detect the phase change. Table D.1 shows the thinnest thickness that can be detected by the amplitude and phase measurements at the present of the Gaussian noise.

sinθ operating points	Amplitude detection	Phase detection
0.804	0.335 nm	0.445 nm
0.811	0.615 nm	0.025 nm
0.814	0.265 nm	0.182 nm

Table D.1 shows the thinnest thickness that can be detected by the amplitude and phase measurements at the present of the Gaussian noise.